

CENTRO DE INVESTIGACIÓN CIENTÍFICA Y DE EDUCACIÓN
SUPERIOR DE ENSENADA



DIVISIÓN DE FÍSICA APLICADA

Departamento de Óptica

ESTUDIOS NUMÉRICOS DE ESPARCIMIENTO INVERSO DE ONDAS
ELECTROMAGNÉTICAS

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el grado de
Doctor en Ciencias presenta:

DEMETRIO MACÍAS GUZMÁN

Ensenada, Baja California, México. Febrero de 2003

RESUMEN de la tesis de DEMETRIO MACÍAS GUZMÁN, presentada como requisito parcial para la obtención del grado de DOCTOR EN CIENCIAS en ÓPTICA, con orientación en ÓPTICA FÍSICA. Ensenada, Baja California, México, Febrero de 2003

ESTUDIOS NUMÉRICOS DE ESPARCIMIENTO INVERSO DE ONDAS
ELECTROMAGNÉTICAS

Aprobado por:

Dr. Eugenio R. Méndez Méndez,

Director de Tesis

El tema central de esta tesis es el desarrollo de algoritmos numéricos para la reconstrucción del perfil de superficies rugosas, utilizando datos de esparcimiento en el campo lejano. Específicamente, se han propuesto dos algoritmos de inversión, ambos con algunas variantes. El primero de ellos está basado en principios de empataamiento de frentes de onda y hace uso de la información de amplitud y fase del campo esparcido por la superficie. El otro algoritmo utiliza la intensidad esparcida en el campo lejano y aborda el problema de esparcimiento inverso como un problema de optimización no lineal. Los datos de entrada para ambos algoritmos fueron generados en forma numérica utilizando un método riguroso basado en el teorema integral de Green.

Palabras clave: esparcimiento de luz, esparcimiento inverso, microscopía confocal, interferometría, computación evolutiva.

ABSTRACT of the Thesis of DEMETRIO MACÍAS GUZMÁN, presented in partial fulfilment of the requirements for the degree of DOCTOR IN SCIENCES in OPTICS, with major in PHYSICAL OPTICS. Ensenada, Baja California, Mexico, Febrero de 2003

ESTUDIOS NUMÉRICOS DE ESPARCIMIENTO INVERSO DE ONDAS
ELECTROMAGNÉTICAS

Abstract approved by:

Dr. Eugenio R. Méndez Méndez,

Thesis advisor

The subject of this thesis is the development of numerical algorithms for the reconstruction of the profiles of one-dimensional rough surfaces using far-field scattered data. Specifically two inversion algorithms have been proposed, one of them is based on wave-front matching principles and uses information of the amplitude and phase of the scattered field. The other algorithm uses far-field scattered intensity and approaches the inverse problem as a non linear optimization problem. The input data for the two algorithms was generated numerically using a rigorous method based on Green's Integral Theorem.

Key words: Light scattering, inverse scattering, confocal microscopy, interferometry, evolutionary computation.

Este trabajo está especialmente dedicado

A *Gabriela (Gaviota) Fumagalli*, por todo lo maravilloso que su llegada ha traído consigo.

A mi madre *Luz Margarita Guzmán* y a mis hermanos *Juan, Paloma, Eugenia* y *Gabriela*, quienes a pesar de la distancia siempre han estado presentes.

A la memoria de mi padre *Eugenio Macías*.

I want to especially dedicate this work to the memory of my dearest friend *Patricia "Pat" Higgings*, whose unconditional friendship and wonderful sense of humor made of 2000 an unfogettable year.

Agradecimientos

No solamente por sus valiosas críticas y observaciones durante el desarrollo de este trabajo, sino también por sus enseñanzas a lo largo de mi estancia en este centro de investigación, quiero expresar mi más sincero agradecimiento a los miembros de mi comité de tesis, *Dr. Eugenio R. Méndez M., Dr. Neil C. Bruce, Dr. Anatolii Khomenko F., Dr. Gustavo Olague y Dr. Víctor Ruíz.*

A *Ileana y a Claudio* por su amistad, por los interminables "recreos" y también por los asados.

A *Blanca* por el privilegio de su amistad.

A las secretarias de la jefatura *Olga, Ana y Carmen*, mil gracias por todo!

A todos aquellos que en algún momento compartieron conmigo esta experiencia.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) el apoyo económico brindado durante mis estudios de posgrado.

Índice General

I	Introducción	7
II	Teoría	11
II.1	El problema de esparcimiento directo	12
II.1.1	Aproximación de Kirchhoff (o del Plano Tangente)	23
II.1.2	Teorías perturbativas, Hipótesis de Rayleigh	25
II.2	El problema de esparcimiento inverso: Antecedentes	29
II.2.1	Reconstrucción de superficies rugosas utilizando información de amplitud y de fase	29
II.2.2	Reconstrucción de superficies rugosas utilizando información de intensidad	37
II.3	Representación numérica de superficies rugosas unidimensionales	40
II.3.1	Método espectral	41
II.3.2	Aproximación de Shoenberg	43
III	El problema de esparcimiento inverso utilizando datos de amplitud y fase	52
III.1	Algoritmo de empatamiento de frentes de onda (AEFO)	53
III.1.1	Reconstrucción de una superficie rugosa utilizando el AEFO	58
III.1.2	Aplicación del AEFO en sistema estratificado con tres capas	76
III.2	Conclusiones parciales	90
IV	El esparcimiento inverso como un problema de optimización	91
IV.1	Algoritmos de optimización	92
IV.1.1	Método Simplex	95
IV.1.2	Método de Recocido Simulado	100
IV.1.3	Algoritmos Evolutivos	102
IV.2	Recuperación de perfiles utilizando la representación espectral	115
IV.2.1	Método Simplex	117
IV.2.2	Recocido Simulado	118
IV.2.3	Algoritmo de inversión evolutivo sin recombinación	123

IV.3 Recuperación de perfiles utilizando funciones B-spline	130
IV.3.1 Algoritmo de inversión híbrido $(\mu/\rho, \lambda) + \text{Simplex}$	131
IV.4 Conclusiones Parciales	137
V Discusión y Conclusiones	139
V.1 Discusión	140
V.2 Resumen y conclusiones	149
A El problema de la unicidad al utilizar datos de intensidad	155
B Esquema de reducción de esparcimiento múltiple	157
C Publicaciones y presentaciones en congresos	163
Bibliografía	166

Índice de Figuras

Figura		Página
1	Geometría del problema de esparcimiento.	13
2	Teorema de Green.	15
3	Representación numérica de una superficie aleatoria rugosa	42
4	Base de funciones Spline generada, a partir de las ecuaciones (69) y (70) para la secuencia de nodos 77.	47
5	Función Spline Básica cúbica formada por la unión, en los nodos, de cuatro polinomios cúbicos.	48
6	Superficie aleatoria representada como una curva spline (línea sólida). El polígono de control está dibujado con una línea punteada.	50
7	Geometría para ilustrar el principio de operación del algoritmo de empataamiento de frentes de onda.	57
8	(a) Perfil utilizado para la generación de los datos de esparcimiento (b) Primera y segunda derivadas del perfil.	59
9	Datos de esparcimiento en función de los números de onda q y k . (a) Amplitud, (b) Fase.	60
10	Mapas de la función $F_{real}^{p,s}(\xi, \eta)$ utilizando los datos de la figura 9:(a) Polarización p, (b) polarización s. El perfil original se muestra con una línea continua y contraste invertido.	61
11	Mapa de la función $F_{conf}^{p,s}(\xi, \eta)$ utilizando los datos de la figura 9:(a) Polarización p, (b) polarización s. El perfil original se muestra con una línea continua.	62
12	Perfil estimado a partir de los máximos de la función $F_{real}^{s-p}(\xi, \eta)$	64
13	Reconstrucciones obtenidas al utilizar a) Plata, b) vidrio. En ambos casos se consideró polarización p.	65
14	Ilustración del campo incidente cuando se tienen: (a) 151, (b) 61 y (c) 30 ángulos de incidencia, respectivamente.	70

15	Mapa de $F_{real}^p(\xi, \eta)$ para la superficie perfectamente conductora mostrada en la figura 8, considerando (a) 151,(b) 61 y (c) 30 ángulos de incidencia, respectivamente.	71
16	Amplitud y fase de los datos de esparcimiento, (a) sin ruido y (b) con ruido.	75
17	Mapa de $F_{real}^p(\xi, \eta)$ utilizando los datos de esparcimiento de la figura 16,(a) Datos sin ruido y (b) datos con ruido.	75
18	Geometría del problema de esparcimiento para el caso de superficies estratificadas.	77
19	Mapa de $F_{real}^{p,s}(\xi, \eta)$ correspondiente a un medio estratificado constituido por una película de plata recubierta con vidrio. Los perfiles originales se muestran con líneas blancas. Para la superficie superior se consideró una desviación estándar de alturas $\delta = 0.1\lambda$ y una longitud de correlación $a = 1.0\lambda$. Para la superficie inferior se consideró un ancho $w = 4.0\lambda$ y una altura $h = -1.0\lambda$. (a) Polarización s, (b) polarización p.	81
20	Mapa de $F_{real}^{p,s}(\xi, \eta)$ correspondiente a un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 1.7689$ y un sustrato dieléctrico con $\epsilon_s(\omega) = 4.00$. Ambas superficies son realizaciones de un proceso aleatorio gaussiano. Sus desviaciones estándar de alturas son $\delta_p = 0.1\lambda$ y $\delta_s = 0.3\lambda$ y sus longitudes de correlación son $a_p = 1.0\lambda$ y $a_s = 3.0\lambda$, respectivamente.(a) Polarización s, (b) polarización p.	83
21	a) Haz sin aberraciones enfocado en el origen, b) Haz enfocado a través de una una interface plana.	84
22	Mapa de $F_{real}^{p,s}(\xi, \eta)$ cuando se corrige la aberración esférica en un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 2.25$ y un sustrato metálico con $\epsilon_s(\omega) = -17.25 + i0.498$.(a)Polarización s, (b) polarización p	88
23	Mapa de $F_{real}^{p,s}(\xi, \eta)$ cuando se corrige la aberración esférica en un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 1.7689$ y un sustrato dieléctrico con $\epsilon_s(\omega) = 4.00$. (a) Polarización s, (b) polarización p.	88
24	Mapas de la función $F_{real}^{s-p}(\xi, \eta)$ considerando el plano promedio de la interface superior y corrigiendo la aberración esférica. (a) Medio estratificado correspondiente a la figura (22), (b)medio estratificado correspondiente a la figura (23).	89
25	Ilustración de las operaciones del simplex para una función de dos variables. 96	
26	Diagrama de bloques de Algoritmo Evolutivo Canónico.	104
27	Perfil original a reconstruir.	116

28	Patrón de esparcimiento generado por la superficie de la figura 27, al ser iluminada con una onda plana a incidencia normal. La línea punteada corresponde al patrón de esparcimiento generado, bajo las mismas condiciones de iluminación, por el perfil dibujado con círculos en la figura 34(a), cuya descripción se encuentra en la sección IV.2.3.	117
29	Reconstrucción de perfiles utilizando el método simplex incluyendo el perfil original en la población inicial. Reconstrucciones obtenidas después de (a) 300 iteraciones, (b) 500 iteraciones y (c) 1000 iteraciones. .	119
30	Reconstrucciones obtenidas con el recocido simulado para a) $T_1 = 0$, b) $T_2 = 50$ y c) $T_3 = 100$. La inicialización fue la misma para las tres corridas.	121
31	Diagrama de bloques de las estrategias evolución utilizadas.	124
32	Esquema de mutación utilizado con la representación espectral.	125
33	Perfiles reconstruidos con estrategias de evolución,(a) utilizando la estrategia (μ, λ) , b) utilizando la estrategia $(\mu + \lambda)$.El perfil original se muestra con la línea continua.	127
34	a) Perfil reconstruido, b) Perfil reconstruido reflejado con respecto a ambos ejes de simetría. El perfil original se muestra con una línea continua.	128
35	Reconstrucción de perfiles con datos ruidosos utilizando la estrategia de evolución (μ, λ) .(a) Reconstrucción con ruido al 5% del valor máximo de intensidad esparcida, b)Reconstrucción con ruido al 10% del valor máximo de intensidad esparcida. El perfil original se muestra con una línea continua.	130
36	Perfil empleado para la generación de los datos de esparcimiento. . . .	132
37	Patrón de intensidad esparcida por la superficie mostrada en la figura 36, para incidencia normal. La línea punteada muestra el patrón de esparcimiento producido por la superficie en la figura (39b)(línea continua) bajo las mismas condiciones de iluminación.	133
38	Gráfica de convergencia para la mejor solución encontrada por cada una de los algoritmos estudiados. Se tiene, respectivamente, (a) Estrategia $(\mu/\rho + \lambda)$ +Simplex, (b)Estrategia- $(\mu/\rho, \lambda)$ +Simplex, (c)Estrategia- (μ, λ) +Simplex y (d) Estrategia- $(\mu + \lambda)$ + Simplex.	135
39	Reconstrucción del perfil utilizando: (a) La estrategia $(\mu/\rho + \lambda)$ -ES+ Downhill simplex, b) La estrategia $(\mu/\rho, \lambda)$ -ES+ Downhill simplex. El perfil original se muestra con círculos y las reconstrucciones están dibujadas con una línea punteada para las dos estrategias de evolución y con una línea continua para el método simplex.	136
40	Reconstrucciones obtenidas utilizando las estrategias: (c) (μ, λ) + simplex y (d) $(\mu + \lambda)$ + simplex . El perfil original se muestra con círculos. Se utiliza una línea punteada para las reconstrucciones correspondientes a las estrategias de evolución y una línea continua para las reconstrucciones obtenidas con el Método Simplex.	137

41	Reconstrucción de una superficie aleatoria perfectamente conductora : (a) Mapa de la función $F_{real}^{s-p}(\xi, \eta)$ empleando el algoritmo de empata- miento de frentes de onda, b) Perfiles estimados con los algoritmos de inversión propuestos.	141
42	Reconstrucción de un surco triangular con pendientes de 45° utilizando: (a) Mapa de la función $F_{real}^{s-p}(\xi, \eta)$ empleando el algoritmo de empata- miento de frentes de onda, b) Perfiles estimados con los algoritmos de inversión propuestos.	143
43	Reconstrucción de un surco triangular con pendientes de 55° . (a) Per- files estimados con el AEFO y la estrategia híbrida $(\mu/\rho, \lambda)+$ simplex, considerando dos inicializaciones distintas. (b)Comparación entre el pa- trón de esparcimiento producido por el perfil original(línea continua) y los patrones generados por las reconstrucciones mostradas en la gráfica (a) de esta figura.	145
44	Información de amplitud y de fase esparcida por (a) el surco triangular dibujado con una línea continua en la figura 43a, (b) la reconstrucción dibujada con cruces en la figura 43a y (c) la reconstrucción dibujada con una línea punteada en la figura 43a.	147
45	Reconstrucciones obtenidas con la estrategia $(\mu/\rho, \lambda)$ para (a) polariza- ción s y (b) polarización p.	150

Capítulo I

Introducción

El esparcimiento de ondas electromagnéticas, producido cuando éstas inciden en una superficie rugosa, ha sido ampliamente estudiado desde hace ya varias décadas. Mediante la aplicación de diferentes teorías, como por ejemplo la aproximación de Kirchhoff, el teorema de extinción o teorías perturbativas, se han realizado trabajos que permiten explicar y entender diferentes aspectos del fenómeno, tales como la dependencia de la intensidad promedio de la luz esparcida por una superficie como función del ángulo de esparcimiento [DeSanto, 1979] o el retroesparcimiento reforzado [McGurn *et al.*, 1985; Méndez y O'Donnell, 1987]. En particular, para entender este último fenómeno es necesario modelar situaciones en las que existe esparcimiento múltiple, lo cual es difícil de realizar con métodos analíticos, por lo que normalmente se emplean técnicas numéricas [Thorsos, 1988; Maradudin *et al.*, 1990].

Por otro lado, el problema inverso, que versa sobre la reconstrucción del perfil de una superficie a partir de datos de esparcimiento, también ha sido estudiado, aunque

no tan extensivamente. Sin embargo, debido a su complejidad, ofrece una veta rica en posibilidades de estudio. Las líneas de investigación más exploradas sobre problemas inversos en óptica se concentran en la recuperación de parámetros estadísticos de la superficie [Welford, 1977; Chandley, 1976; Elson y Bennet, 1979; Stover y Serafi, 1984; Malyshkin *et al.*, 1997]. Sin embargo, aunque más escasos, existen también trabajos en los que se intenta recuperar el perfil superficial. En 1980, A. Roger y D. Maystre [Roger y Maystre, 1980] reportaron un trabajo en el que intentaban recuperar el perfil de rejillas de difracción partiendo de datos del campo lejano. Para superficies no periódicas, están los trabajos de Wombell y DeSanto [Wombell y DeSanto, 1991a; Wombell y DeSanto, 1991b], basados en teorías perturbativas y en la aproximación de Kirchhoff. Más recientemente Quartel y Sheppard [Quartel y Sheppard, 1996a; Quartel y Sheppard, 1996b], propusieron un método aproximado para resolver el fenómeno de esparcimiento inverso con un algoritmo que simula el proceso de formación de imágenes en un microscopio confocal. Este novedoso punto de vista permite hacer uso de la información de esparcimiento obtenida al iluminar a la superficie desde varias direcciones.

En esta tesis se estudian dos algoritmos de inversión para la recuperación del perfil. El primero basado en un principio de empatamiento de frentes de onda, que hace uso de datos de amplitud (módulo y fase) del campo lejano esparcido, generado al iluminar la superficie a diferentes ángulos de incidencia. Este algoritmo está relacionado con los algoritmos propuestos por Quartel y Sheppard [Quartel y Sheppard, 1996a; Quartel y Sheppard, 1996b]. La segunda técnica de inversión propuesta es de naturaleza muy

distinta, pues emplea datos de intensidad esparcida en el campo lejano y basa su funcionamiento en el planteamiento del problema de esparcimiento inverso como un problema de optimización no lineal. Cabe señalar que este algoritmo hace uso de principios heurísticos y de herramientas computacionales para la determinación de la solución en lo que puede catalogarse como una estrategia evolutiva.

La estructura del presente trabajo puede dividirse en cuatro bloques fundamentales que corresponden a los capítulos II, III, IV y V.

En el capítulo II se plantea el problema bajo estudio y se introducen algunos conceptos teóricos necesarios para visualizar las posibles maneras de resolver el problema directo y apreciar la complejidad del problema inverso. En la primera sección se plantea el problema de esparcimiento directo y se describen brevemente tres de las metodologías que se han empleado para su solución. Posteriormente, se lleva a cabo una revisión de la literatura referente al problema inverso. También se propone la posibilidad de resolver el problema inverso como un problema de optimización y se define un funcional objeto, cuya minimización reconstruye la superficie buscada. Finalmente, la tercera sección está enfocada a la descripción de los esquemas de representación numérica de superficies rugosas, que serán necesarios para la implementación de los algoritmos que se proponen en este trabajo.

El capítulo III está dedicado a descripción del algoritmo de empatamiento de frentes de onda (AEFO), basado en el empleo de datos de la amplitud esparcida. Se estudia su desempeño para el caso de superficies perfectamente conductoras, dieléctricas y metálicas. A su vez, se propone y estudia una estrategia de muestreo que permite reducir la

redundancia de información, determinando el número mínimo de ángulos de incidencia y de esparcimiento necesarios para la reconstrucción. Otro aspecto importante que se estudia es la sensibilidad del algoritmo en presencia de ruido en los datos y a efectos de esparcimiento múltiple. Un caso interesante que también es estudiado en este capítulo es la reconstrucción de las superficies internas de un medio estratificado. Dado que los primeros resultados presentan algunos problemas asociados a la refracción que sufren las ondas incidentes antes de llegar a la superficie bajo estudio, se propone un esquema para corregirlos y se presentan las reconstrucciones obtenidas. Finalmente, se presentan las conclusiones del capítulo.

En el capítulo IV se plantea el problema inverso como un problema de optimización no lineal. En la primera sección se discuten brevemente los conceptos básicos de la teoría de optimización. A su vez, con base en las características del funcional objeto definido en el capítulo II, se eligen tres algoritmos de optimización y se describen los principios sobre los cuales se fundamenta la operación de cada uno de ellos. Las secciones restantes muestran los resultados obtenidos con cada uno de los algoritmos de inversión propuestos a lo largo del capítulo. Al final se enuncian las conclusiones parciales del capítulo.

El quinto capítulo muestra, en su primera sección, una comparación entre el desempeño de dos de los algoritmos propuestos para tres casos particulares. El AEFO y el basado en estrategias evolutivas. Se discuten los resultados obtenidos para, en la segunda sección, dar paso a las conclusiones generales del trabajo y a una propuesta para la continuación de esta línea de investigación.

Capítulo II

Teoría

El objetivo de este capítulo es presentar una revisión de algunos aspectos teóricos relevantes sobre el problema que se estudia en este trabajo y discutir los métodos más promisorios propuestos en la literatura para resolver el problema inverso. La primera sección se enfoca en la descripción de tres aproximaciones clásicas al problema de esparcimiento directo. En la segunda sección se plantea el problema de esparcimiento inverso a estudiar y se discuten las características esenciales de los algoritmos de inversión reportados en la literatura. Se plantea la posibilidad de abordar el problema de esparcimiento inverso formulándolo como un problema de optimización. Finalmente, en la tercera sección se describen los esquemas de representación de superficies rugosas unidimensionales que serán utilizados en el resto del trabajo.

II.1 El problema de esparcimiento directo

En esta sección se considera el problema de esparcimiento de luz por una superficie rugosa. Inicialmente se define el problema físico a estudiar y también se introduce la notación que será utilizada. Posteriormente, se describe un método riguroso para la determinación del campo esparcido por una superficie rugosa. Dado que el método está basado en la solución de ecuaciones integrales, se conoce como el método de la ecuación integral. Finalmente, se presentan dos métodos aproximados de solución del problema directo para el caso particular de una interface entre el vacío y un conductor perfecto. El primer método está basado en la llamada aproximación de Kirchhoff o del plano tangente y el segundo es la teoría perturbativa de amplitudes pequeñas. Debido a que, por limitaciones de memoria y de velocidad en las computadoras actuales, solamente resulta práctico abordar problemas rigurosos con superficies unidimensionales (superficies que son invariantes en una dirección), la revisión se limitará a este caso.

La geometría del problema se presenta en la figura 1. Se considera el esparcimiento de luz por una superficie rugosa unidimensional, cuyo perfil está definido por la ecuación $x_3 = \zeta(x_1)$. A su vez, se supone que las regiones semiinfinitas $x_3 > \zeta(x_1)$ y $\zeta(x_1) > x_3$ contienen medios homogéneos e isotrópicos caracterizados por sus constantes dieléctricas $\epsilon_0(\omega)$ y $\epsilon_I(\omega)$, respectivamente. El eje x_3 es perpendicular al plano promedio de la superficie y ésta, a su vez, es invariante en la dirección x_2 . El plano de incidencia es el plano $x_1 - x_3$. Se supone también que la región $x_3 > \zeta(x_1)$ como el medio de incidencia.

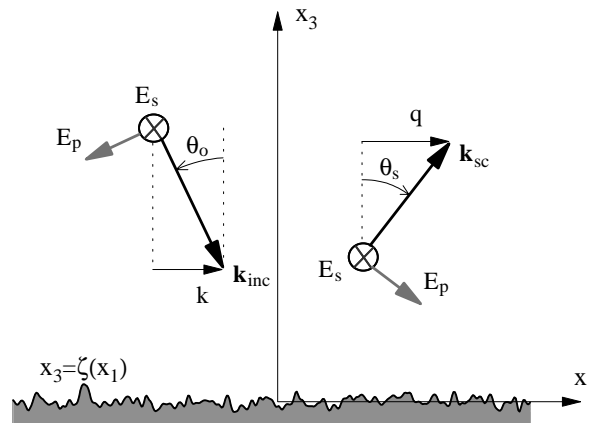


Figura 1: Geometría del problema de esparcimiento.

Con base en la geometría del problema, para campos invariantes a lo largo de x_2 es clara la posibilidad de emplear un tratamiento escalar [Born y Wolf, 1980]. Es decir, el campo incidente se puede especificar por la componente x_2 del vector de campo eléctrico o magnético, según el tipo de polarización. Suponiendo ondas monocromáticas, el campo es de la forma

$$\Psi(\mathbf{r}; t) = (0, \psi(\mathbf{r}), 0) \exp\{-i\omega t\}, \quad (1)$$

donde $\mathbf{r} = (x_1, x_3)$. Si se considera el estado elemental de polarización TM o p, la componente escalar $\psi(\mathbf{r})$ representará la componente x_2 del campo magnético $H_2(\mathbf{r})$. Por otro lado, si se considera el estado de polarización TE o s, entonces $\psi(\mathbf{r})$ representará la componente x_2 del campo eléctrico $E_2(\mathbf{r})$. En lo que resta de este trabajo, se supondrá una dependencia temporal $\exp\{-i\omega t\}$, pero únicamente se hará referencia explícita a ella cuando sea necesario para darle claridad a la exposición.

Con referencia en la figura 1, se denotará por $\psi^>(\mathbf{r})$ al campo en la región $x_3 > \zeta(x_1)$ y por $\psi^<(\mathbf{r})$ al campo en la región $\zeta(x_1) > x_3$.

En la región $x_3 > \zeta(x_1)$, $\psi^>(\mathbf{r})$ satisface la ecuación homogénea de Helmholtz

$$\left(\nabla + \epsilon_0(\omega) \left(\frac{\omega}{c} \right)^2 \right) \psi^>(\mathbf{r}) = 0. \quad (2)$$

Similarmente, en la región $\zeta(x_1) > x_3$ se tiene que

$$\left(\nabla + \epsilon_I(\omega) \left(\frac{\omega}{c} \right)^2 \right) \psi^<(\mathbf{r}) = 0. \quad (3)$$

En las expresiones (2) y (3), ω es la frecuencia angular del campo y c es la velocidad de la luz.

El punto de partida para la formulación analítica del problema de esparcimiento directo es el teorema integral de Green [Jackson, 1965] que establece, para dos funciones escalares arbitrarias $u(\mathbf{r}')$ y $g(\mathbf{r}')$ bien comportadas dentro de un volumen v' , se cumple que

$$\int_{v'} (u(\mathbf{r}') \nabla'^2 g(\mathbf{r}') - g(\mathbf{r}') \nabla'^2 u(\mathbf{r}')) dv' = \oint_{\Sigma'} (u(\mathbf{r}') \nabla' g(\mathbf{r}') - g(\mathbf{r}') \nabla' u(\mathbf{r}')) \cdot \mathbf{d}\Sigma', \quad (4)$$

donde Σ' representa una superficie arbitraria cerrada que contiene al volumen v' , $\mathbf{d}\Sigma'$ es un vector normal a la superficie, que apunta hacia afuera del volumen v' , y cuya magnitud es igual al elemento diferencial $d\Sigma'$.

Inicialmente, considerando la geometría mostrada en la figura 2, se aplica el teorema integral de Green en la región $x_3 > \zeta(x_1)$. En esta etapa de la exposición, es conveniente introducir la llamada función de Green $G_0(\mathbf{r} | \mathbf{r}')$ del medio $x_3 > \zeta(x_1)$, que es solución

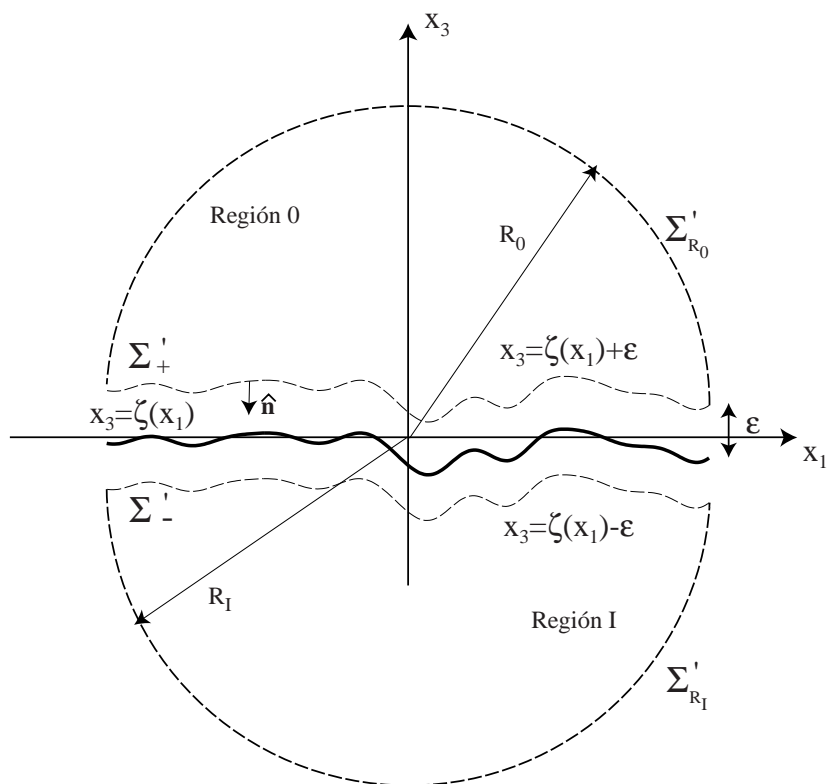


Figura 2: Teorema de Green.

de la ecuación

$$\left(\nabla'^2 + \epsilon_0(\omega) \left(\frac{\omega}{c} \right)^2 \right) G_0(\mathbf{r} | \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}') \quad (5)$$

Ahora, se eligen las funciones $u(\mathbf{r})$ y $g(\mathbf{r})$ de manera que $u(\mathbf{r}) = \psi^>(\mathbf{r})$ y que $g(\mathbf{r} | \mathbf{r}') = G_0(\mathbf{r} | \mathbf{r}')$. Despejando $\nabla'^2\psi^>(\mathbf{r}')$ y $\nabla'^2G_0(\mathbf{r} | \mathbf{r}')$ de las ecuaciones (2) y (5), respectivamente, y sustituyendo en el lado izquierdo de la expresión matemática del teorema de Green [ecuación (4)] se tiene que

$$\int_{v'} [u(\mathbf{r}')\nabla'^2g(\mathbf{r}') - g(\mathbf{r}')\nabla'^2u(\mathbf{r}')] dv' = \int_{v'} \psi^>(\mathbf{r}')\delta(\mathbf{r}' - \mathbf{r}) dv' = \begin{cases} \psi^>(\mathbf{r}), & \text{si } \mathbf{r} \in v' \\ 0, & \text{si } \mathbf{r} \notin v' \end{cases} \quad (6)$$

De aquí puede verse la conveniencia de haber elegido la función $G_0(\mathbf{r} | \mathbf{r}')$ definida por la ecuación (5).

Ahora se considera el lado derecho de la expresión del teorema de Green [ecuación (4)]. La superficie cerrada Σ' puede descomponerse como la suma de varias superficies (figura 2) tales que $\Sigma' = \Sigma'_{R_0} \cup \Sigma'_{l_1} \cup \Sigma'_{l_2} \cup \Sigma'_+$, donde Σ'_{R_0} representa como la superficie de un semi-cilindro paralelo al eje x_2 , Σ'_{l_1} y Σ'_{l_2} son las áreas correspondientes a las secciones transversales del semi-cilindro y Σ'_+ es el área de una superficie imaginaria justo arriba del perfil superficial $x_3 = \zeta(x_1)$.

Dada la invariancia de las funciones evaluadas a lo largo de la dirección x_2 y al hecho de que la normal a las superficies Σ'_{l_1} y Σ'_{l_2} es paralela a esta dirección, las contribuciones correspondientes se anulan. Entonces, para el campo por encima de la

superficie puede escribirse

$$\begin{aligned} \psi^>(\mathbf{r}) = & \frac{1}{4\pi} \lim_{R_0 \rightarrow \infty} \int_{-\infty}^{\infty} \left(\psi^>(\mathbf{r}') \frac{\partial}{\partial n'_+} G_0(\mathbf{r}' | \mathbf{r}) - G_0(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'_+} \psi^>(\mathbf{r}') \right) dS'_{R_0} \\ & + \frac{1}{4\pi} \int_{-\infty}^{\infty} \left(\psi^>(\mathbf{r}') \frac{\partial}{\partial n'_+} G_0(\mathbf{r}' | \mathbf{r}) - G_0(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'_+} \psi^>(\mathbf{r}') \right) dS'_+, \end{aligned} \quad (7)$$

donde se ha hecho uso de la definición de derivada direccional $\hat{\mathbf{n}} \cdot \nabla = \frac{\partial}{\partial n}$ a lo largo de un vector unitario $\hat{\mathbf{n}}$, normal a la superficie, cuyo sentido apunta hacia afuera del volumen v' . A su vez, dS'_{R_0} y dS'_+ son elementos diferenciales de longitud de arco que corresponden, respectivamente, al semicírculo de radio R_0 contenido en el plano $x_1 - x_3$ y al perfil $x_3 = \zeta(x_1)$.

El campo total $\psi^>(\mathbf{r})$ en la región $x_3 > \zeta(x_1)$ consiste de dos contribuciones, representadas por las integrales del lado derecho de la ecuación (7). La contribución de la integral sobre la semi circunferencia de radio infinito debe dar como resultado el campo incidente $\psi^>(\mathbf{r})_{inc}$ (en ausencia de la superficie, sería la única contribución). De este modo, la ecuación (7) se escribe en forma compacta como:

$$\psi^>(\mathbf{r}) = \psi^>(\mathbf{r})_{inc} + \psi^>(\mathbf{r})_{sc}, \quad (8)$$

donde

$$\psi^>(\mathbf{r})_{sc} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left(\psi^>(\mathbf{r}') \frac{\partial}{\partial n'_+} G_0(\mathbf{r}' | \mathbf{r}) - G_0(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'_+} \psi^>(\mathbf{r}') \right) dS'_+. \quad (9)$$

A partir de la expresión (9) es claro que la determinación del campo esparcido invo-

lucra el conocimiento del campo $\psi^>(\mathbf{r}')$ evaluado en la superficie y su derivada normal $\frac{\partial}{\partial n'}\psi^>(\mathbf{r}')$. Estas dos cantidades son desconocidas y el problema de esparcimiento directo consiste, esencialmente, en la determinación de ellas. El método de la ecuación integral se basa en el establecimiento y solución de dos ecuaciones integrales acopladas para determinar el campo y su derivada normal en la superficie. Procediendo de manera análoga a se hizo para la región $x_3 = \zeta(x_1)$, es posible encontrar una expresión para el campo total $\psi^<(\mathbf{r})$ en la región $x_3 < \zeta(x_1)$. Para esto, debe considerarse la función de Green del medio I, que es solución de la ecuación

$$\left(\nabla'^2 + \epsilon_I(\omega) \left(\frac{\omega}{c} \right)^2 \right) G_I(\mathbf{r} | \mathbf{r}') = -4\pi\delta(\mathbf{r} - \mathbf{r}'), \quad (10)$$

Después de un poco de álgebra se encuentra que

$$\Theta(\zeta(x_1) - x_3) \psi^<(\mathbf{r}) = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left(\psi^<(\mathbf{r}') \frac{\partial}{\partial n'_-} G_I(\mathbf{r}' | \mathbf{r}) - G_I(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'_-} \psi^<(\mathbf{r}') \right) dS'_-, (11)$$

donde dS'_- es la diferencial de longitud de arco correspondiente a la curva $x_3 = \zeta(x_1)$ y Θ representa la función escalón de Heaviside. Su presencia en el lado izquierdo de la ecuación (11) es representativa del hecho de que si el punto de observación \mathbf{r} está en v'_- , la integral de volumen del teorema de Green dá el campo $\psi^<(\mathbf{r})$, y si está fuera, dá cero.

Las ecuaciones para determinar las funciones fuente se obtienen considerando que

\mathbf{r} se acerca a la superficie por arriba. También es necesario emplear las condiciones de continuidad para las componentes tangenciales de los campos a través de la frontera.

Se tiene que

$$\psi^>(x_1, x_3 | \omega) \Big|_{x_3=\zeta(x_1)} = \psi^<(x_1, x_3 | \omega) \Big|_{x_3=\zeta(x_1)} \quad (12)$$

y

$$\frac{1}{\nu_0} \frac{\partial}{\partial n} \psi^>(x_1, x_3 | \omega) \Big|_{x_3=\zeta(x_1)} = \frac{1}{\nu_I} \frac{\partial}{\partial n} \psi^<(x_1, x_3 | \omega) \Big|_{x_3=\zeta(x_1)}, \quad (13)$$

donde $\nu_{\mathcal{R}} = 1$ para polarización s, y $\nu_{\mathcal{R}} = \epsilon_{\mathcal{R}}(\omega)$ para polarización p. A su vez, la derivada normal puede definirse en términos del perfil de la superficie, como:

$$\frac{\partial}{\partial n} = [1 + (\zeta'(x_1)^2)]^{-1/2} \left(-\zeta'(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right) = \frac{1}{\phi(x_1)} \frac{\partial}{\partial N}, \quad (14)$$

donde $\frac{\partial}{\partial N}$ es el operador derivada normal no unitaria. Finalmente, después de algunas manipulaciones algebraicas, y de utilizar las relaciones

$$\frac{\partial}{\partial n} = -\frac{\partial}{\partial n_+} = \frac{\partial}{\partial n_-} \quad (15)$$

y

$$dS'_{\pm} = \sqrt{1 + (\zeta'(x'_1))^2} dx'_1 = \phi(x'_1) dx'_1. \quad (16)$$

Al acercar el punto de observación a la superficie, se llega a

$$\varphi(x_1 | w) = \psi(x_1 | w)_{inc} + \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} dx'_1 [\varphi(x'_1 | \omega) H^{(0)}(x_1 | x'_1) - \chi(x'_1 | \omega) L^{(0)}(x_1 | x'_1)] \quad (17)$$

y

$$0 = -\lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 [\varphi(x'_1 | \omega) H^{(1)}(x_1 | x'_1) - \nu_{\mathcal{R}} \chi(x'_1 | \omega) L^{(1)}(x_1 | x'_1)], \quad (18)$$

en donde $\varphi(x'_1 | \omega)$ y $\chi(x'_1 | \omega)$ son dos funciones fuente definidas por las expresiones:

$$\varphi(x_1 | \omega) = \psi^>(x_1, \zeta(x_1)) \Big|_{x_3 = \zeta(x_1) + \epsilon} \quad (19)$$

y

$$\chi(x_1 | \omega) = \left[-\zeta'(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] \psi^>(x_1, x_3) \Big|_{x_3 = \zeta(x_1) + \epsilon}. \quad (20)$$

A su vez,

$$H^{(0)}(x_1 | x'_1) = \left[-\zeta'_1(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_0(x_1, x'_1 | x_3, x'_3) \Big|_{\substack{x'_3 = \zeta_1(x'_1) \\ x_3 = \zeta_1(x_1) + \epsilon}}, \quad (21)$$

$$L^{(0)}(x_1 | x'_1) = G_0(x_1, x'_1 | x_3, x'_3) \Big|_{\substack{x'_3 = \zeta_1(x'_1) \\ x_3 = \zeta_1(x_1) + \epsilon}}, \quad (22)$$

$$H^{(1)}(x_1 | x'_1) = \left[-\zeta'_1(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_I(x_1, x'_1 | x_3, x'_3) \Big|_{\substack{x'_3 = \zeta_1(x'_1) \\ x_3 = \zeta_1(x_1) + \epsilon}} \quad (23)$$

y

$$L^{(1)}(x_1 | x'_1) = G_I(x_1, x'_1 | x_3, x'_3) \Big|_{\substack{x'_3 = \zeta_1(x'_1) \\ x_3 = \zeta_1(x_1) + \epsilon}}. \quad (24)$$

Si se conocen las funciones fuente, con base en la ecuación (9), es posible calcular el campo esparcido en cualquier punto por arriba de la superficie. Sin embargo, para calcular el campo en la zona lejana a ésta, es conveniente expresarlo como una superposición de ondas planas. Para esto se introduce la representación de Weyl [Mandel y Wolf, 1995] para la función de Green $G_0(x_1, x_3 | x'_1, x'_3)$

$$G_0(x_1, x_3 | x'_1, x'_3) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} \frac{2\pi i}{\alpha_0(q)} \exp \{i [q(x_1 - x'_1) + \alpha_0(q)(x_3 - x'_3)]\}, \quad (25)$$

con

$$\alpha_0(q) = \sqrt{\left(n_0(\omega) \frac{\omega}{c}\right)^2 - q^2}.$$

La ecuación (25) expresa a la función de Green como una superposición de ondas planas. Al sustituirla en la ecuación (9), el campo esparcido por la superficie puede expresarse de la forma:

$$\psi_{p,s}^>(x_1, x_3)_{sc} = \int_{-\infty}^{\infty} \frac{dq}{2\pi} R_{p,s}(q|k) \exp \{iqx_1 + i\alpha_0(q)x_3\}, \quad (26)$$

donde $R_{p,s}(q|k)$ es la amplitud de esparcimiento, o espectro angular del campo esparcido, dada por

$$\begin{aligned} R_{p,s}(q|k) &= \frac{i}{2\alpha_0(q)} \int_{-\infty}^{\infty} dx'_1 \exp \{-iqx'_1 - i\alpha_0(q)\zeta(x'_1)\} \times \\ &\times [i(q\zeta'(x'_1) - \alpha_0(q)) \varphi(x'_1|\omega) - \chi(x'_1|\omega)]. \end{aligned} \quad (27)$$

Las funciones fuente pueden determinarse resolviendo el sistema de ecuaciones integrales acopladas (17) y (18). Es necesario tener en cuenta que las funciones de Hankel y sus derivadas tienen singularidades importantes alrededor del origen, por lo que no es posible intercambiar el orden del operador límite con el de integración. Un procedimiento para resolver numéricamente este sistema se encuentra descrito en forma detallada en la referencia [Maradudin *et al.*, 1990].

Dada la dificultad para resolver analíticamente el problema, la determinación de las funciones fuente se realiza en forma numérica. Además, como puede observarse en la ecuación (20), la relación de la función fuente con el perfil de la superficie es complicada.

El problema se simplifica para el caso particular en el que la superficie es un conductor perfecto, ya que solamente se requiere determinar una función fuente, la otra es idénticamente igual a cero, debido a que el campo eléctrico en la superficie de un conductor perfecto es nulo. En este caso, la amplitud esparcida puede escribirse en forma genérica como:

$$R_{p,s}(q|k) = \frac{i}{2\alpha_0(q)} \int_{-\infty}^{\infty} dx'_1 \exp \{-iqx'_1 - i\alpha_0(q)\zeta(x'_1)\} [\beta((q, \zeta')F(x'_1|\omega))]. \quad (28)$$

El factor $\beta_{p,s}(q, \zeta')$ depende de la polarización y está dado por

$$\beta_{p,s}(q, \zeta') = \begin{cases} i[q\zeta'(x'_1) - \alpha_0(q)] & \text{polarización p} \\ 1 & \text{polarización s} \end{cases}. \quad (29)$$

A su vez, $F_{p,s}(x'_1|\omega)$ representa la función fuente, dada por

$$F_{p,s}(x'_1|\omega) = \begin{cases} \varphi(x'_1|\omega) & \text{polarización p} \\ \chi(x'_1|\omega) & \text{polarización s} \end{cases} \quad (30)$$

Finalmente, la ecuación (17) para un conductor perfecto adquiere la forma

$$\varphi_{p,s}(x'_1 | \omega) = \psi(x_1 | w)_{inc} + \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} dx'_1 [F_{p,s}(x'_1|\omega)\beta_{p,s}(q, \zeta')] \quad (31)$$

II.1.1 Aproximación de Kirchhoff (o del Plano Tangente)

La aproximación de Kirchhoff es una aproximación local para el valor del campo o su derivada normal en la superficie. Se basa en la suposición de que las condiciones de frontera para el campo en cada punto del perfil se pueden aproximar por aquellas que corresponden a un plano infinito tangente a éste. Para que la aproximación sea buena debe cumplirse que $\lambda \ll a$. Es decir, que la longitud de correlación a de la superficie $x_3 = \zeta(x_1)$ debe ser mucho mayor que la longitud de onda del campo incidente.

Por el momento se supondrá que la superficie es iluminada por una onda plana. La componente no nula del vector de campo, ya sea magnético o eléctrico, está dada por la ecuación

$$\psi(x_1, x_3)_{inc} = \psi_0 \exp \{ ikx_1 - i\alpha_0(k)x_3 \}, \quad (32)$$

donde $\alpha_0(k) = \sqrt{(\omega/c)^2 - k^2}$ y ω es la frecuencia del campo incidente. Los ángulos θ_0

y θ_s están relacionados con los números de onda k y q , respectivamente, por

$$k = \frac{\omega}{c} \sin \theta_0, \quad q = \frac{\omega}{c} \sin \theta_s. \quad (33)$$

En la Aproximación de Kirchhoff, las funciones fuente estarán dadas por las expresiones [Beckmann y Spizzichino, 1963; Ogilvy, 1991]

$$\varphi^{AK}(x'_1, \zeta(x'_1)) = (1 + R)\psi(x'_1, \zeta(x'_1))_{inc} \quad (34)$$

y

$$\chi^{AK}(x'_1, \zeta(x'_1)) = (1 - R)\frac{\partial}{\partial N'}\psi(x'_1, \zeta(x'_1))_{inc}, \quad (35)$$

en donde R es el coeficiente de reflexión de Fresnel y $\frac{\partial}{\partial N'}$ es el operador derivada normal, definido en la ecuación (14). Sustituyendo las ecuaciones (32), (34) y (35) en la ecuación (27), después de acomodar términos se llega a siguiente la expresión para la amplitud compleja esparcida

$$R_{p,s}^{AK}(q|k) = \frac{i}{2\alpha_0(q)} \int_{-\infty}^{\infty} dx'_1 [i(q\zeta'(x'_1) - \alpha_0(q))(1 + R) + i(k\zeta'(x'_1) + \alpha_0(k))(1 - R)] \\ \times \exp \{i v_x x'_1 - i v_z \zeta(x'_1)\}, \quad (36)$$

donde $v_x = k - q$ y $v_z = \alpha_0(q) + \alpha_0(k)$. Este resultado se puede simplificar integrando por partes la ecuación (36). El procedimiento para un medio general se complica, pues el coeficiente de reflexión es función de la posición. Entonces, se considera el caso

particular de un conductor perfecto, para el cual se tendrá que $R = \pm 1$ dependiendo del tipo de polarización empleada; el signo positivo corresponde a la polarización p, mientras que el negativo a la s. Después de integrar por partes la ecuación (36), la expresión resultante tiene la forma

$$R_{p,s}^{AK}(q|k) = F(\theta_i, \theta_s) \int_{-\infty}^{\infty} dx'_1 \exp \{iv_x x'_1 - iv_z \zeta(x'_1)\} \quad (37)$$

donde $F(\theta_0, \theta_s)$ factor angular dado por:

$$F(\theta_0, \theta_s) = \pm \frac{1 + \cos(\theta_0) \cos(\theta_s) - \sin(\theta_0) \sin(\theta_s)}{\cos(\theta_0) + \cos(\theta_s)} \quad (38)$$

El signo del factor angular $F(\theta_0, \theta_s)$ dependerá de las condiciones de frontera elegidas. Es decir, para polarización p será negativo, mientras que para polarización s será positivo.

Es necesario mencionar que, formalmente, la ecuación (37) consta de dos términos. Uno de ellos es el que se muestra dicha ecuación, mientras que el otro, está relacionado con efectos de borde y es despreciable cuando el tamaño de la superficie es mucho mayor que la longitud de onda del campo incidente [Beckmann y Spizzichino, 1963].

II.1.2 Teorías perturbativas, Hipótesis de Rayleigh

El problema directo puede resolverse de manera aproximada con la teoría perturbativa de pequeña amplitud. La versión más utilizada se basa en la llamada hipótesis de

Rayleigh, que permite la determinación (en forma aproximada) de la amplitud de esparcimiento $R(q | k)$, sin necesidad de especificar las funciones fuente $\varphi(x'_1 | \omega)$ y $\chi(x'_1 | \omega)$. La amplitud de esparcimiento $R(q | k)_{p,s}$ es calculada mediante una expansión en serie de potencias, cuyo parámetro pequeño es representativo del cociente entre las alturas de la superficie y la longitud de onda del campo incidente.

Aunque de naturaleza distinta, como en el caso de la aproximación de Kirchhoff, existen limitaciones referentes a la aplicación de la hipótesis de Rayleigh, que condicionan la convergencia de la serie resultante. De manera general, para que esta aproximación pueda utilizarse, las rugosidades de la superficie deben cumplir la condición $|\alpha_0(k)\zeta(x_1)| \ll 1$ [Thorsos y Jackson, 1989]; es decir, deben ser pequeñas en comparación con la longitud de onda del campo incidente.

Antes de enunciar la Hipótesis de Rayleigh, es conveniente recordar que la representación de un campo en términos de su espectro angular está restringida a geometrías de medio espacio o a regiones limitadas por planos paralelos. En ambos casos, el medio dentro de la región debe ser homogéneo y sin fuentes [Mandel y Wolf, 1995]. La Hipótesis de Rayleigh parte del supuesto de que es válido representar al campo reflejado por la superficie, en la región $x_3 = \zeta(x_1)$, como

$$\psi_R(x_1, x_3) = \int_{-\infty}^{\infty} \frac{dq}{2\pi} R(q | k) \exp i[qx_1 + \alpha_0(q)x_3], \quad (39)$$

cuando en principio, esta representación solamente es válida a partir de $x_3 = \zeta_{max}(x_1)$.

Con el fin de ilustrar la determinación de la amplitud reflejada $R(q | k)$, utilizando

el método de Rayleigh y teoría perturbativa, se considera el caso de un conductor perfecto iluminado con una onda plana y polarización s. Sustituyendo las ecuaciones (32) y (39) en la ecuación (12) se tiene que

$$\psi_0 \exp \{ikx'_1 - i\alpha_0(k)\zeta(x'_1)\} = - \int_{-\infty}^{\infty} \frac{dq}{2\pi} R(q | k) \exp i[qx'_1 + \alpha_0(q)\zeta(x'_1)] \quad (40)$$

Es claro que la ecuación anterior viola la restricción impuesta para la representación de un campo como una superposición de ondas planas, ya que la evaluación se lleva a cabo directamente en el perfil y no en un plano. Multiplicando la ecuación (40) por $\exp \{-ipx'_1\}$ e integrando con respecto a x'_1 , se llega a:

$$\psi_0 \int_{-\infty}^{\infty} dx'_1 \exp \{i(k-p)x'_1 - i\alpha_0(k)\zeta(x'_1)\} \quad (41)$$

$$= - \int_{-\infty}^{\infty} dx'_1 \int_{-\infty}^{\infty} \frac{dq}{2\pi} R(q | k) \exp i[(q-p)x'_1 + \alpha_0(q)\zeta(x'_1)]. \quad (42)$$

Definiendo

$$I(\gamma | Q) = \int_{-\infty}^{\infty} dx'_1 \exp \{-i[Qx'_1 + \gamma\zeta(x'_1)]\} \quad (43)$$

se llega a la versión de una ecuación reducida de Rayleigh para un conductor perfecto:

$$\psi_0 [I(\alpha_0(k) | p-k)] = - \int_{-\infty}^{\infty} \frac{dq}{2\pi} [I(-\alpha_0(q) | p-q)] R(q | k). \quad (44)$$

Utilizando una expansión en serie de Taylor del tipo,

$$\exp -i\gamma\zeta(x_1) = 1 + \frac{-i\gamma}{1!}\zeta(x_1) + \frac{(-i\gamma)^2}{2!}\zeta^2(x_1) + \dots, \quad (45)$$

la función $I(\gamma | Q)$ queda

$$\begin{aligned} I(\gamma | Q) &= \int_{-\infty}^{\infty} dx_1 \exp \{-iQx_1\} \left[1 + \frac{-i\gamma}{1!}\zeta(x_1) + \frac{(-i\gamma)^2}{2!}\zeta^2(x_1) + \dots \right] \\ &= 2\pi\delta(Q) + \frac{-i\gamma}{1!}\hat{\zeta}^{(1)}(Q) + \frac{(-i\gamma)^2}{2!}\hat{\zeta}^{(2)}(Q) + \dots, \quad (46) \end{aligned}$$

donde $\hat{\zeta}^n(Q)$ representa la transformada de Fourier de $\zeta^n(x_1)$. Sustituyendo en la ecuación (44) se obtiene

$$\begin{aligned} -\psi_0 \left[2\pi\delta(p-k) + \frac{-i\gamma}{1!}\hat{\zeta}^{(1)}(p-k) + \frac{(-i\gamma)^2}{2!}\hat{\zeta}^{(2)}(p-k) + \dots \right] = \\ - \int_{-\infty}^{\infty} \frac{dq}{2\pi} [R_s^{(0)}(q|k) + R_s^{(1)}(q|k) + R_s^{(2)}(q|k)] \times \\ \times \left(2\pi\delta(p-q) + \frac{-i\gamma}{1!}\hat{\zeta}^{(1)}(p-q) + \frac{(-i\gamma)^2}{2!}\hat{\zeta}^{(2)}(p-q) + \dots \right), \quad (47) \end{aligned}$$

donde $R_s^{(n)}(q|k)$ representa la contribución a la amplitud de esparcimiento de orden $\zeta^n(x_1)$.

Finalmente, después de igualar los términos semejantes y de invocar las propiedades de la función delta, se llega a que la amplitud esparcida en la región $x_3 \geq \zeta_{max}(x_1)$ está

dada por la serie

$$R_s(q | k) = \psi_0 \left[-2\pi\delta(p - k) + 2i\alpha_0(k)\hat{\zeta}^{(1)}(p - k) + \dots \right] \quad (48)$$

Es importante mencionar que el primer término representa la amplitud especular. A su vez, el segundo término relaciona la amplitud esparcida con la transformada de Fourier de la superficie.

II.2 El problema de esparcimiento inverso: Antecedentes

En la sección anterior se describió en que consiste el problema de esparcimiento directo y se expusieron brevemente algunos métodos de solución. En esta sección se presenta una introducción al problema de esparcimiento inverso y se lleva a cabo una pequeña revisión del tema.

II.2.1 Reconstrucción de superficies rugosas utilizando información de amplitud y de fase

Un tipo de problema de esparcimiento inverso que ha sido ampliamente estudiado en óptica consiste en la recuperación de algunas propiedades promedio de la superficie [Chandley y Welford, 1975; Chandley, 1976; Welford, 1977; Elson y Bennet, 1979; Stover y Serati, 1984]. Cantidades típicas de interés pueden ser la desviación estándar de

alturas $\delta [\langle \zeta^2(x_1) \rangle - \langle \zeta(x_1) \rangle^2]^{1/2}$ y la longitud de correlación a , que son representativos de la escala vertical y horizontal en la superficie. Los paréntesis cuadrados denotan el promedio sobre un conjunto (ensemble) de realizaciones de una superficie aleatoria. El problema de la reconstrucción del perfil de una superficie determinista o de una realización particular de una aleatoria es más complejo. Sin embargo, existen algunos trabajos sobre el tema, como por ejemplo el método propuesto por Roger y D. Maystre [Roger y Maystre, 1980], enfocado principalmente a rejillas de difracción perfectamente conductoras, en el que los autores establecen dos problemas inversos a resolver, dependiendo del tipo de información con el que se cuente. El primero es un problema de reconstrucción en el que se busca hallar el perfil de una rejilla a partir del conocimiento de su eficiencia, medida experimentalmente. El segundo problema inverso consiste en la síntesis y básicamente se busca encontrar, si es que existe, una rejilla de difracción cuya eficiencia a sido definida *a priori*.

Uno de los primeros trabajos en los que se intenta reconstruir el perfil de una superficie con rugosidades aleatorias fue reportado por [Wombell y DeSanto, 1991a]. Los autores proponen un algoritmo de reconstrucción basado en la Hipótesis de Rayleigh, para recuperar el perfil de una superficie aleatoria unidimensional perfectamente conductora, cuando ésta es iluminada por una onda plana con polarización s (condiciones de frontera de Dirichlet). El primer paso de su método es suponer conocidas las amplitudes complejas

$$F_+(k) = \int_{-\infty}^{\infty} dx \chi(x_1 | \omega) \exp \{-i [kx_1 - \alpha_0(k)\zeta(x_1)]\} \quad (49)$$

y

$$F_-(q) = \int_{-\infty}^{\infty} dx \chi(x_1 | \omega) \exp \{-i [qx_1 + \alpha_0(q)\zeta(x_1)]\}, \quad (50)$$

donde $F_+(k)$ es un término fuente relacionado con el campo incidente y su derivada, evaluados sobre una línea paralela al plano promedio $x_3 = 0$ situada a una distancia arbitraria $x_3 = Z_0$ arriba de la superficie. A su vez, la función $F_-(q) = R_s(q | k)$ es el espectro angular del campo esparcido.

Posteriormente, suponiendo $k = q$, al sumar y restar entre sí las ecuaciones (49) y (50) se obtiene

$$\Gamma_+(k) = \frac{F_+(k) + F_-(k)}{2} = \int_{-\infty}^{\infty} dx_1 \chi(x_1 | \omega) \cos \{\alpha_0(k)\zeta(x_1)\} \exp \{-ikx_1\}, \quad (51)$$

$$\Gamma_-(k) = \frac{F_+(k) - F_-(k)}{2i} = \int_{-\infty}^{\infty} dx_1 \chi(x_1 | \omega) \sin \{\alpha_0(k)\zeta(x_1)\} \exp \{-ikx_1\}. \quad (52)$$

Suponiendo que $|\alpha_0(k)\zeta(x_1)| \ll 1$, las ecuaciones (51) y (52) se reescriben como

$$\Gamma_+(k) \approx \int_{-\infty}^{\infty} dx_1 \chi(x_1 | \omega) \exp \{-ikx_1\} \quad (53)$$

$$\Gamma_-(k) \approx \alpha_0(k) \int_{-\infty}^{\infty} dx_1 \chi(x_1 | \omega) \zeta(x_1) \exp \{-ikx_1\} \quad (54)$$

Finalmente, las ecuaciones anteriores pueden interpretarse como las transformadas de Fourier de la función fuente $\chi(x_1 | \omega)$ y del producto de ésta con el perfil de la superficie. Calculando la transformada de Fourier inversa de las funciones $\Gamma_+(k)$ y $\Gamma_-(k)$, denotadas respectivamente por $\gamma_+(x_1)$ y $\gamma_-(x_1)$, y efectuando el cociente entre ellas,

se obtiene que

$$\zeta_r(x_1) = \frac{\gamma_-(x_1)}{\gamma_+(x_1)}, \quad (55)$$

donde $\zeta_r(x_1)$ representa una estimación del perfil original. Debido a que las funciones $\Gamma_-(k)$ y $\Gamma_+(k)$ no cuentan con la información de ondas evanescentes, la reconstrucción estará limitada a detalles mayores o comparables con la longitud de onda. Sin embargo, la limitación más importante que presenta el método reside en el hecho de haber utilizado una aproximación de rugosidades pequeñas.

Wombell y DeSanto propusieron un método alternativo, basado en la aproximación de Kirchhoff, para reconstruir el perfil de la superficie [Wombell y DeSanto, 1991b]. De esta forma, ya no se tiene una limitación directa con respecto a las alturas de las rugosidades. Al igual que con el método anterior, se supone una superficie rugosa unidimensional perfectamente conductora. Para el campo incidente se considera una onda plana con polarización s (condiciones de frontera de Dirichlet).

El punto de partida para este método de inversión es suponer que la amplitud esparcida por la superficie en el campo lejano $R_{p,s}^{AK}$ es conocida. Para llevar a cabo su análisis, Wombell y DeSanto reescriben la ecuación (37) como

$$R_s^{AK}(p|l) = \left[\frac{p^2 + l^2}{l} \right] D(\theta_s, \theta_0), \quad (56)$$

con

$$D(\theta_s, \theta_0) = \int_{-\infty}^{\infty} dx'_1 \exp \left\{ i \left(\frac{\omega}{c} \right) px'_1 \right\} \exp \left\{ -i \left(\frac{\omega}{c} \right) l \zeta(x'_1) \right\}. \quad (57)$$

Las transformaciones

$$p = \sin(\theta_s) - \sin(\theta_0) \quad \text{y} \quad l = \cos(\theta_s) - \cos(\theta_0), \quad (58)$$

se obtienen a partir de las relaciones (33). De esta forma, es claro que el factor angular $F(\theta_0, \theta_s)$ de la ecuación (37) puede escribirse como $\left[\frac{p^2 + l^2}{l} \right]$.

A partir de la ecuación (57) se observa que al variar el ángulo de incidencia θ_0 , o el de esparcimiento θ_s , se tendrá una variación en p y en l . Sin embargo, si existiese alguna forma de variar p y mantener fija a l , entonces se tendría una relación de transformada de Fourier entre una exponencial compleja cuyo argumento es el perfil $\zeta(x_1)$ y la función $D(\theta_0, \theta_s)$. Es decir,

$$D(\theta_0, \theta_s) = \mathcal{F} \left\{ \exp \left\{ -i \left(\frac{\omega}{c} \right) l \zeta(x'_1) \right\} \right\} \quad (59)$$

El resultado anterior es interesante, ya que el perfil original $\zeta(x_1)$ podría estimarse aplicando el operador transformada de Fourier inversa a la ecuación (59), de manera que

$$P(x_1, l) = \mathcal{F}^{-1} \{ D(\theta_0, \theta_s) \} = \exp \left\{ -i \left(\frac{\omega}{c} \right) l \zeta(x'_1) \right\}. \quad (60)$$

Si el perfil estimado se denota como $r(x_1, l)$, al calcular la fase de la expresión (60) se tiene

$$r^\pm(x, l) = \frac{1}{kl} \tan^{-1} \left\{ \frac{-\mathcal{I}m \{ P^\pm(x, l) \}}{\mathcal{R}e \{ P^\pm(x, l) \}} \right\} \quad (61)$$

donde $\mathcal{R}e \{ \}$ y $\mathcal{I}m \{ \}$ denotan la parte real e imaginaria de la función $P^\pm(x, l)$, respectivamente.

Aunque en principio simple, este método posee limitaciones importantes que deben considerarse. Una de ellas está relacionada con la adecuada elección del parámetro l para que los valores correspondientes a p generen ángulos de incidencia y de esparcimiento físicamente posibles. A su vez, el algoritmo de inversión requiere que las superficies a reconstruir se encuentren dentro de los límites de validez de la aproximación de Kirchhoff. Es decir, las pendientes deben ser suaves. Finalmente, al ser necesario calcular la fase para la reconstrucción del perfil, es posible la existencia de ambigüedades que requieran de un algoritmo adicional para evitarlas.

Dos algoritmos de inversión particularmente interesantes fueron reportados por Quartel y Sheppard [Quartel y Sheppard, 1996a; Quartel y Sheppard, 1996b]. En estos trabajos, los autores consideran la posibilidad de utilizar la información de esparcimiento generada cuando la superficie se ilumina desde varias direcciones. Los métodos fundamentan su operación en el proceso de formación de imágenes en el microscopio confocal, y en el microscopio confocal interferométrico, trabajando en modo perfilométrico.

La teoría está basada en la aproximación de Kirchhoff. En la primera publicación [Quartel y Sheppard, 1996a], los autores utilizan un formalismo similar al de [Wombell y DeSanto, 1991b], partiendo de la ecuación (57) reescrita de la siguiente forma:

$$R^{KA}(p, l) = \left[\frac{p^2 + l^2}{l} \right] \int_{-\infty}^{\infty} dx'_1 \int_{-\infty}^{\infty} dx'_3 \delta [x'_3 - \zeta(x'_1)] \times \\ \times \left(\exp \left\{ -i \left(\frac{\omega}{c} \right) l x'_3 \right\} \right) \exp \left\{ -i \left(\frac{\omega}{c} \right) p x'_1 \right\} \quad (62)$$

Quartel y Sheppard, por medio de una analogía, interpretan la ecuación (62) como la función de esparcimiento $S(p, l)$ descrita en [Sheppard, 1986] y que a su vez juega un papel fundamental en el cálculo de la amplitud de la imagen de un microscopio confocal

$$U(x_1, x_3) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(p, l) S(p, l) \exp \left\{ i \frac{\omega}{c} [px_1 + lx_3] \right\} dp dl, \quad (63)$$

donde $c(p, l)$ es la función de transferencia coherente del sistema óptico. En su análisis, [Quartel y Sheppard, 1996a] argumentan que debido a que su objetivo no es simular el funcionamiento de un microscopio confocal sino únicamente recuperar, de la mejor manera posible, el perfil de una superficie rugosa, la elección de la función $c(p, l)$ puede ser en forma arbitraria. Entonces, a partir de la ecuaciones (62) y (63), eligen

$$c(p, l) = \begin{cases} \frac{l}{p^2 + l^2} & \text{para } p^2 + l^2 < 4 \text{ y } l > 2 \cos \alpha, \\ 0 & \text{para cualquier otro caso.} \end{cases} \quad (64)$$

Donde α es la abertura numérica.

La imagen del objeto se generará mediante un barrido axial, i.e. en la dirección x_3 , para cada uno de los puntos del plano promedio $x_3 = 0$. El resultado es un mapa de las intensidades, en función de las coordenadas, que sigue la forma del perfil original. Debe notarse que la función delta en la ecuación (62) representa la imagen ideal de la superficie $\zeta(x'_1)$.

En una publicación posterior [Quartel y Sheppard, 1996b], estos autores proponen la reconstrucción de perfiles unidimensionales haciendo uso de microscopio confocal

interferométrico. Para llevar a cabo su objetivo, Quartel y Sheppard suponen conocidas las intensidades de salida correspondientes a cada uno de los brazos de un interferómetro de Michelson. Posteriormente, en forma electrónica, estas mediciones son procesadas para obtener una señal proporcional a

$$I = \mathcal{R}e \{UR^*\}, \quad (65)$$

donde $U(x_1, x_3)$ está dada por la ecuación (63) y $R = R_0 \exp \{i\phi\}$ es la amplitud del haz de referencia. El resultado del proceso de inversión, al igual que en el caso del microscopio confocal, es un mapa de la función (65) conformado por una estructura estructura de franjas cuya forma asemeja al perfil original.

Una vez formado el mapa de intensidades (microscopio confocal), Quartel y Sheppard [Quartel y Sheppard, 1996a] utilizan un algoritmo de detección de picos de intensidad para estimar el perfil. Por otro lado, en [Quartel y Sheppard, 1996b] se propone la utilización de un algoritmo de interpolación parabólica, ya que la detección de máximos arroja perfiles con discontinuidades debidas, principalmente, a que ocasiones el máximo no se encuentra en la franja de más alto contraste que sigue la forma del perfil sino en una adyacente.

II.2.2 Reconstrucción de superficies rugosas utilizando información de intensidad

Aunque notables, los algoritmos descritos anteriormente presentan limitaciones importantes. Una de ellas reside en la dificultad de su implementación experimental. En un trabajo reciente, [Destouches *et al.*, 2001] reportan resultados experimentales obtenidos al emplear un algoritmo de inversión basado en la aproximación de Kirchhoff, para la reconstrucción del perfil de rejillas de difracción. En su publicación, estos autores mencionan que una de las razones por las que no utilizaron el método de Wombell y DeSanto [Wombell y DeSanto, 1991b] fue, precisamente, su dificultad de implementación. Otra limitación importante en los algoritmos de inversión mencionados es que necesariamente requieren la información de la fase para llevar a cabo la reconstrucción.

Es interesante entonces considerar la posibilidad de reconstruir el perfil de una superficie rugosa partiendo de datos de intensidad, pues esta es una situación cuya realización experimental resulta más sencilla. Sin embargo, este nuevo enfoque presenta aspectos importantes que deben tomarse en cuenta. Para ilustrar la complejidad del problema se considera un superficie perfectamente conductora iluminada por una onda plana [ecuación (32)] con polarización *s*. La amplitud esparcida por la superficie está dada por la ecuación (28), por lo tanto, la intensidad en el campo lejano será:

$$I_s(q|k) = \frac{1}{4\alpha_0^2(k)} \left| \int_{-\infty}^{\infty} dx_1 [F(x_1|\omega) \exp \{-i\alpha_0(q)\zeta(x_1)\}] \exp \{-iqx_1\} \right|^2. \quad (66)$$

Si se supone, por el momento, que la relación entre el perfil de la superficie $\zeta(x_1)$ y la función fuente $F(x_1|\omega)$ es conocida y que la variación de $\alpha_0(q)$ puede ser despreciada, la intensidad esparcida puede interpretarse como la transformada de Fourier de la cantidad entre los paréntesis cuadrados. Es claro que la información de fase de la amplitud compleja esparcida se ha perdido y este hecho complica, aún más, la recuperación del perfil. Es conocido que para el caso de objetos unidimensionales, el problema de la reconstrucción de una función a partir del módulo cuadrado de su transformada de Fourier no tiene una solución única [Walther, 1963; O'Neill y Walther, 1963]. Esto quiere decir, dentro del contexto de problema que se estudia en este trabajo, que un mismo patrón de esparcimiento en el campo lejano puede ser generado por varias superficies.

El problema de la reconstrucción de un objeto real, a partir del módulo cuadrado de su transformada de Fourier, ha sido estudiado extensivamente dentro del contexto de la interferometría astronómica de speckle [Dainty, 1984]. Dentro de los algoritmos propuestos para la solución de este problema se encuentran los trabajos de [Gerchberg y Saxton, 1971; Gerchberg y Saxton, 1972] para objetos unidimensionales y el propuesto Fienup [Fienup, 1978]. En este último se emplea un método iterativo para minimizar el error existente entre el módulo de la transformada de Fourier del objeto original y un módulo estimado que se genera en cada iteración del algoritmo. El resultado presentado por Fienup es la reconstrucción, en forma directa, de un objeto bidimensional y es importante enfatizar que para este caso, aunque no se ha demostrado analíticamente la existencia de una solución única al problema, existen resultados experimentales que

permiten suponer la unicidad de la solución [Dainty, 1975].

Cabe mencionar que el método propuesto por Fienup, aunque en su trabajo no se hace referencia explícita a este hecho, puede considerarse un método de optimización no lineal. Siguiendo esta filosofía, el problem inverso estudiado en esta tesis puede abordarse utilizando este enfoque.

Formulación del problema inverso como un problema de optimización

Se ha visto que la dependencia entre la intensidad esparcida en el campo lejano y el perfil de la superficie es complicada. Sin embargo, el problema directo puede resolverse rigurosamente en forma numérica [Maradudin *et al.*, 1990] y este hecho abre la posibilidad de definir una función objetivo que haga uso de estos datos y cuyo óptimo sea el perfil de la superficie que quiere recuperarse. Suponiendo que se cuenta con información de intensidad esparcida en el campo lejano $I^{(m)}(q | k)$, la semejanza entre un perfil propuesto $z_c(x_1)$ y el perfil original $\zeta(x_1)$ puede evaluarse mediante la comparación directa que resulta de calcular la diferencia entre el patrón de intensidad medido $I^{(m)}(q|k)$ y el patrón $I^{(c)}(q|k)$, resultado de resolver el problema directo considerando el perfil $z_c(x_1)$. El objetivo es encontrar una superficie tal que se satisfaga la condición $I^{(c)}(q|k) = I^{(m)}(q|k)$. Cuando esta igualdad se cumple, y si la solución es única, podría suponerse que el perfil ha sido recuperado.

Con el fin de reducir el espacio de búsqueda y, para disminuir el número de posibles soluciones, la definición de la función objetivo utiliza información obtenida de iluminar a la superficie con varios ángulos de incidencia. Se define la función de costo de la

forma

$$f(\zeta(x_1), k_j) = \sum_{j=1}^{N_{ang}} \int |I^{(m)}(q|k_j) - I^{(c)}(q|k_j)| dq, \quad (67)$$

donde N_{ang} es el número de direcciones desde las cuales se ilumina a la superficie, k'_j s corresponde a los diferentes ángulos de incidencia. La integral en la ecuación (67) establece una medida de la diferencia entre los patrones de intensidad $I^{(m)}(q|k_j)$ y $I^{(c)}(q|k_j)$ en la vecindad de q . Entonces, la función $f(\zeta(x_1), k_j)$ puede interpretarse como la suma de los errores resultantes de la comparación entre los patrones de esparcimiento generados por la superficie original $\zeta(x_1)$ y una superficie $z^{(c)}(x_1)$, cuando ésta es iluminada a diferentes ángulos de incidencia. En principio, puede suponerse que si el valor de la ecuación (67) es mínimo, ambas superficies son iguales. Es claro que los parámetros a optimizar (variables objeto) son definidos por las alturas de la superficie y que, en el trabajo numérico éstos deben ser discretizados. Es decir, que el conjunto de parámetros a optimizar son las alturas muestreadas de la superficie. Por lo tanto, el siguiente paso consiste en hallar un esquema de representación adecuado.

II.3 Representación numérica de superficies rugosas unidimensionales

Para implementar los algoritmos estudiados en este trabajo es necesario, tanto para generar datos de esparcimiento como para resolver el problema inverso, utilizar versiones muestreadas de las superficies cuyos perfiles deben reconstruirse. El objetivo de esta

sección es describir los dos posibles esquemas de representación utilizados para generar y representar superficies rugosas de manera discreta.

II.3.1 Método espectral

Debido a que una gran parte de los estudios, teóricos y experimentales, sobre esparcimiento de luz por superficies corrugadas aleatoriamente consideran superficies gaussianas, es interesante considerar el caso en que las superficies a reconstruir en este trabajo pertenecen a esta clase estadística. Las superficies a reconstruir constituyen entonces una sola realización de un proceso aleatorio gaussiano, con media nula, desviación estándar igual a uno y función de correlación gaussiana. Los parámetros estadísticos elegidos para caracterizar a la superficie son la desviación estándar de alturas δ y la longitud de correlación a .

Las superficies pertenecientes a esta clase pueden generarse en forma numérica utilizando el método espectral expuesto en [Thorsos, 1988; Maradudin *et al.*, 1990]. Se tiene que el conjunto de valores

$$\zeta_n = \frac{\delta}{\sqrt{L}} \sum_{j=-N/2}^{N/2-1} \frac{[M_j + iN_j]}{\sqrt{2}} \sqrt{g(|q_j|)} \exp\{iq_j\chi_n\}, \quad (68)$$

representa una realización discreta de un perfil superficial $\zeta_n = \zeta(\chi_n)$, que puede ser convenientemente generada empleando una transformada rápida de Fourier, donde χ_n son los puntos muestreados, definidos a lo largo de la dirección x_1 como $\chi_n = -L/2 + n\Delta x$, donde Δx es el intervalo de muestreo, N representa el número total de puntos de

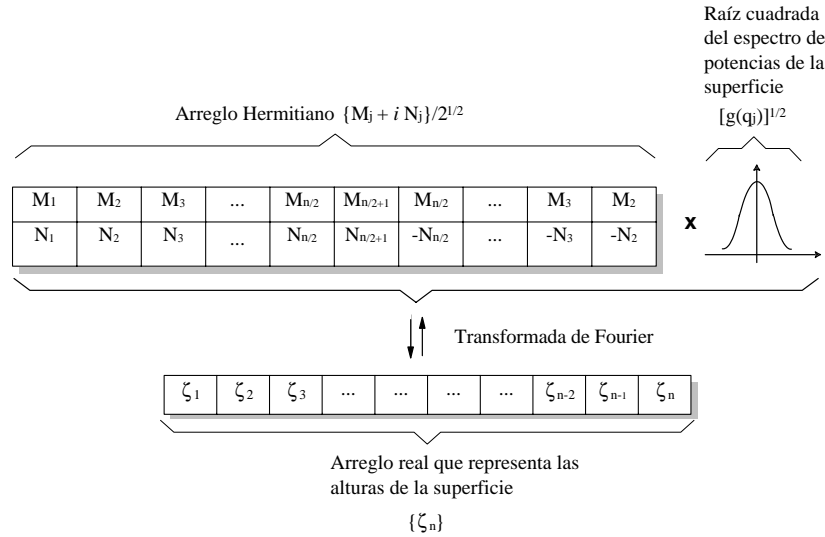


Figura 3: Representación numérica de una superficie aleatoria rugosa .

la superficie, L representa su longitud y $q_j = 2\pi j/L$. Los conjuntos aleatorios $\{M_j\}$ y $\{N_j\}$ contienen variables aleatorias gaussianas estadísticamente independientes entre sí, con media cero y desviación estándar uno y $g(|q_j|)$ representa la densidad espectral de potencia de las alturas de la superficie. En este trabajo se escogió la forma particular $g(|q_j|) = \sqrt{\pi}a \exp\left\{\frac{-a^2 q^2}{4}\right\}$. Para poder generar un conjunto de números aleatorios reales $\{\zeta_n\}$, es necesario que el arreglo complejo $[M_j + iN_j]$ sea Hermitiano. La primera y segunda derivadas del perfil, necesarias para resolver el problema de esparcimiento directo, pueden obtenerse al derivar la ecuación (68).

El procedimiento de generación numérica de una superficie utilizando la ecuación (68) se representa esquemáticamente en la figura 3. La transformada de Fourier del producto entre el arreglo complejo $[\{M_j\} + i\{N_j\}]/\sqrt{2}$ y la raíz cuadrada del espectro de potencias de la superficie $g(|q_j|)$ da como resultado al arreglo real ζ_n que representa las alturas de la superficie.

II.3.2 Aproximación de Shoenberg

En este apartado se presenta un esquema alternativo para la representación de superficies rugosas unidimensionales basado en conceptos de Teoría de la Aproximación y de una rama de la Geometría Computacional conocida como Diseño Geométrico Asistido por Computadora (CAGD por sus siglas en inglés). Inicialmente se describen las características principales de un tipo especial de funciones polinomiales conocidas como Spline Básicas. Se ilustra la generación de éstas a partir de un algoritmo recursivo y posteriormente se muestra su aplicación para generar una superficie rugosa unidimensional representada mediante una curva plana abierta.

Definición de las Funciones Spline Básicas (B-Splines)

Las funciones spline básicas (B-Splines) son funciones polinomiales por tramos definidas a partir de una sucesión de números $\mathbf{t} = \{t_j\}_1^{m=k+n}$ estrictamente creciente, conocida como *secuencia o vector de nodos*, cuyos elementos no necesariamente son distintos entre sí. El índice superior m es el número total de elementos en la sucesión, k es el orden de los polinomios que conforman a la función B-Spline y n es el número de funciones B-Spline de grado $p = k - 1$ distintas de cero en el intervalo semi-abierto $[t_j, t_{j+k})$.

La j -ésima función B-spline de orden $k = 1$ se define como

$$B_{j,1,t}(x_1) \triangleq \begin{cases} 1 & \text{para } t_j \leq x_1 < t_{j+1}, \\ 0 & \text{para cualquier otro caso.} \end{cases} \quad (69)$$

Para $k > 1$, las funciones B-spline se definen y evalúan en forma recursiva mediante el *Algoritmo de Cox-de Boor* [Cox, 1972; de Boor, 1972]

$$B_{j,k,t}(x_1) \triangleq \omega_{j,k,t}(x_1)B_{j,k-1,t}(x_1) + (1 - \omega_{j+1,k,t}(x_1))B_{j+1,k-1,t}(x_1) \quad (70)$$

donde

$$\omega_{j,k,t}(x_1) \triangleq \begin{cases} \frac{x_1 - t_j}{t_{j+k-1} - t_j} & \text{if } t_j \neq t_{j+k-1}, \\ 0 & \text{para cualquier otro caso.} \end{cases} \quad (71)$$

Particularmente, la función $B_{j,k,t}(x_1)$ es nula fuera del intervalo $[t_j, \dots, t_{j+k})$ y es positiva dentro de éste. Es decir,

$$B_{j,k,t}(x_1) > 0 \quad t_j < x_1 < t_{j+k},$$

si $t_j = t_{j+k} \implies B_{j,k,t}(x_1) = 0$

Las ecuaciones (69) y (71) ilustran la importancia del vector de nodos para la definición y la generación numérica de las funciones spline básicas (B-spline). Las secuencias de nodos pueden dividirse en las tres clases principales descritas continuación:

1. **Secuencia uniforme:** Aquella que cumple con la condición

$$t_{j+1} - t_j = \Delta t, \quad \forall j. \quad (72)$$

Donde $\Delta t > 0$ es una constante arbitraria. Algunos ejemplos de este tipo de

secuencias son

$$\{1, 2, 3, 4, 5, 6, 7, 8\},$$

$$\{0, 0.25, 0.5, 0.75, 1.0\},$$

$$\{-2.5, -1.4, -0.3, 0.8, 1.9, 3.0\}$$

2. **Secuencia uniforme abierta:** Aquella que cumple con las condiciones

$$t_1 = \dots = t_k = a, \quad t_{m+1}, \dots, t_{m+k} = b$$

y

$$t_{j+1+k} - t_{j+k} = \Delta t, \quad \text{para } j = 1, \dots, m. \quad (74)$$

Donde a y b son números reales que representan el nodo inicial y el nodo final, respectivamente. La constante Δt , al igual que en el inciso anterior, indica la separación entre los *nodos internos* comprendidos en el intervalo $[t_{k+1}, \dots, t_m]$.

Una propiedad de las funciones B-spline establece que en un nodo de multiplicidad ν la función $B_{j,k,t}(x_1)$ es $C^{k-\nu}$ continua [de Boor, 1978]. Es decir, $B_{j,k,t}(x_1)$ es derivable $k - \nu$ veces en el nodo respectivo. Este hecho permite que los polinomios que conforman a la función $B_{j,k,t}(x_1)$ coincidan perfectamente, satisfaciendo las condiciones de continuidad, en el nodo particular en el que se unen. Por otro lado, si $\nu = k$ la función será discontinua en el nodo de multiplicidad k . De esta manera, la multiplicidad de los nodos extremos en las secuencias *uniformes abiertas* implica que las funciones resultantes no sean derivables en los nodos extremos.

Algunos ejemplos para ilustrar la forma de las secuencias uniformes abiertas son

$$\{1, 1, 1, 2, 3, 4, 5, 6, 6, 6\} \quad \text{para } k = 3,$$

$$\{-6, -6, -6, -6, -4, -3, -2, -1, 0, 1, 2, 3, 4, 6, 6, 6, 6\} \quad \text{para } k = 4,$$

$$\{0.1, 0.1, 0.1, 0.1, 0.1, 0.3, 0.5, 0.7, 0.7, 0.7, 0.7, 0.7\} \quad \text{para } k = 5.$$

3. **Secuencia no uniforme:** Este es el caso general y la única condición que debe cumplir es ser estrictamente creciente. Algunos ejemplos son

$$\{1, 3, 7, 22, 23, 23, 49, 50, 50\},$$

$$\{1, 1, 1, 2, 2, 3, 4, 5, 6, 6, 6, 7, 7, 7\},$$

$$\{0.2, 0.7, 0.7, 0.7, 1.2, 1.2, 2.9, 3.6\}.$$

Debido a que las secuencias *uniformes abiertas* son las más utilizadas para la generación de curvas spline y, a su vez, estas últimas servirán para representar superficies rugosas unidimensionales, es conveniente ilustrar la generación de funciones B-spline por medio del siguiente ejemplo.

Considérese la secuencia de nodos uniforme abierta

$$\mathbf{t} = \{-6, -6, -6, -6, -4, -3, -2, -1, 0, 1, 2, 3, 4, 6, 6, 6, 6\}. \quad (77)$$

En la figura 4 se muestra la base de funciones spline básicas correspondientes a la secuencia (77) generadas empleando en forma sucesiva la relación de recurrencia (70).

La consecuencia de la multiplicidad de los nodos extremos se aprecia en la forma

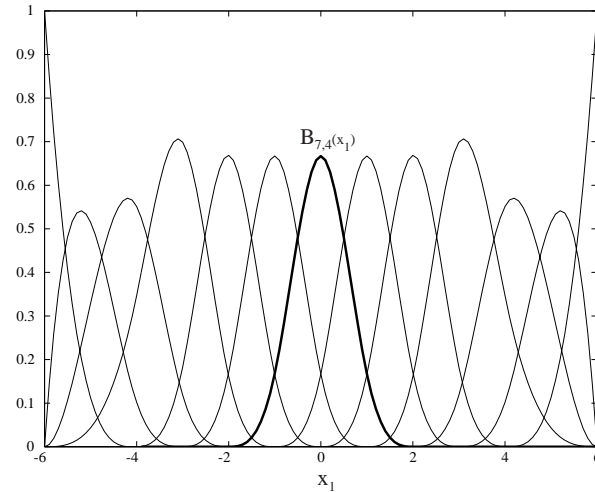


Figura 4: Base de funciones Spline generada, a partir de las ecuaciones (69) y (70) para la secuencia de nodos 77.

de las funciones $B_{1,4,t}(x)$ y $B_{13,4,t}(x)$, éstas presentan un pico localizado en la misma posición de los nodos múltiples. Posteriormente se verá que el efecto de la multiplicidad de nodos en la curva spline resultante es que ésta sea tangente a los puntos extremos del polígono de control. Por otro lado, el número de nodos múltiples indica el orden de las funciones spline básicas, para el ejemplo considerado es $k = 4$.

Una propiedad importante de las funciones B-spline puede apreciarse si, por un momento, se centra la atención en la función B-spline $B_{7,4,t}(x)$ de la figura 4. Mediante el uso del algoritmo recursivo de Cox-de Boor [ecuación (70)] se llega a que la función

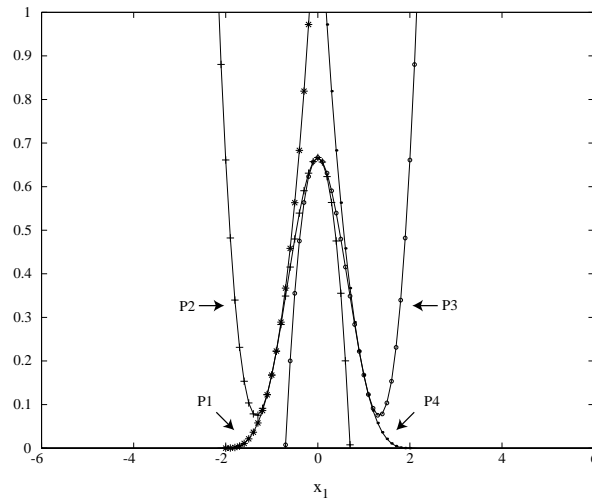


Figura 5: Función Spline Básica cúbica formada por la unión, en los nodos, de cuatro polinomios cúbicos.

$B_{7,4,t}(x)$ tiene la forma

$$B_{7,4,t}(x_1) = \begin{cases} \text{P1: } \left(\frac{1}{6}x_1^3 + x_1^2 + 2x_1 + \frac{8}{6}\right) & \text{para } -2 \leq x_1 < -1, \\ \text{P2: } \left(-\frac{1}{2}x_1^3 - x_1^2 + \frac{4}{6}\right) & \text{para } -1 \leq x_1 < 0, \\ \text{P3: } \left(\frac{1}{2}x_1^3 - x_1^2 + \frac{4}{6}\right) & \text{para } 0 \leq x_1 < 1, \\ \text{P4: } \left(-\frac{1}{6}x_1^3 + x_1^2 - 2x_1 + \frac{8}{6}\right) & \text{para } 1 \leq x_1 < 2. \end{cases} \quad (78)$$

A partir de las ecuaciones (78) es claro que la función $B_{7,4,t}(x_1)$ es polinomial y continúa por tramos, está compuesta por $k = 4$ polinomios cúbicos unidos en los nodos respectivos $t_7 = -2, t_8 = -1, t_9 = 0, t_{10} = 1, t_{11} = 2$ dentro del intervalo $[t_7, t_{11})$.

Curvas Spline

Es posible representar una superficie unidimensional $\zeta(x_1)$ utilizando un esquema basado en la Aproximación de Shoenberg [de Boor, 1978]. Si se supone que ζ_n es una

curva plana definida en el intervalo $[a, b]$, su representación paramétrica será una curva spline de la forma [de Boor, 2002]

$$\zeta_n = \sum_{j=1}^m P_j(\tau_j, \alpha_j) B_{j,k,t}(x_1^n) \quad \text{for } n = 1, \dots, N \quad (79)$$

donde N es el número total de puntos de muestreo $x_1^{(n)}$ sobre la superficie, $B_{j,k,t}(x_1)$ es la j -ésima *función spline básica* definida en la sección anterior, $\mathbf{t} = (t_j)_1^m$ es la *secuencia de nodos*, $P_j(\tau_j, \alpha_j)$ es el j -ésimo punto de control, α_j es un número real que se localiza en la j -ésima abscisa de Greville τ_j . Cabe señalar, que el origen de este nombre se remonta al planteamiento inicial de los esquemas de interpolación con splines, en los que determinación de las abscisas τ_j de los puntos a interpolar $P_j(\tau_j, \alpha_j)$ se llevaba a cabo en función del vector de nodos mediante la relación de Greville [de Boor, 1978; de Boor, 2002]

$$\tau_j = \frac{t_j + \dots + t_{j+k-1}}{k-1} \quad \forall j. \quad (80)$$

Sin embargo, como es el caso de la presente aplicación, si lo que se conoce (o puede determinarse) son las abscisas de Greville y no la secuencia de nodos, ésta puede suponerse uniforme abierta y los nodos internos podrán calcularse utilizando la ecuación (80) reescrita en la forma

$$t_{k+j} = \frac{\tau_j + \dots + \tau_{j+k-1}}{k-1} \quad j = 1, \dots, n. \quad (81)$$

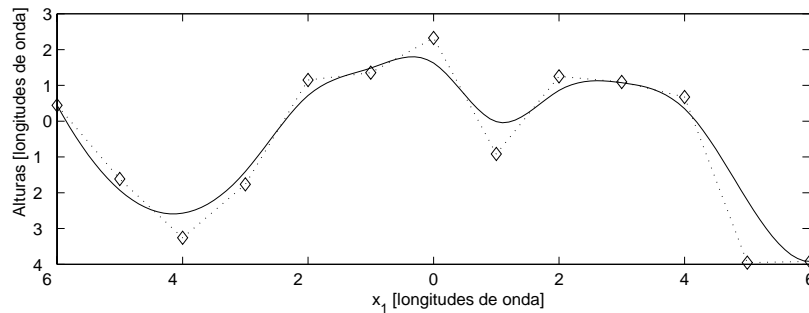


Figura 6: Superficie aleatoria representada como una curva spline (línea sólida). El polígono de control está dibujado con una línea punteada.

Debe mencionarse que el resultado que se obtiene al emplear la aproximación de Shoenberg es una curva suave y continua por tramos, la cual puede interpretarse como una versión suavizada del polígono de control que la contiene y cuyos vértices son los puntos $P_j(\tau_j, \alpha_j)$. Es decir, la curva generada por medio de la ecuación (79) *no pasa* por ninguno de los puntos de control, a diferencia de del esquema de interpolación con splines, en el que los puntos de control *necesariamente* están contenidos en la curva resultante.

En la figura 6 se muestra la representación de una superficie rugosa unidimensional mediante una curva spline. El número total de puntos de muestreo considerados fue $L = 121$. En este ejemplo, la posición de las abscisas de Greville fue determinada en forma arbitraria, el dominio de variación de la curva fue dividido en doce segmentos, cada uno de los cuales contiene 10 puntos de muestreo. La línea sólida es la superficie aleatoria generada. Se puede observar que está contenida dentro del polígono representado con la línea punteada. A su vez, los puntos de control $P_j(\tau_j, \alpha_j)$, corresponden a los diamantes en la figura. El orden de las funciones B-Spline fue $k = 4$. Puede observarse que el

efecto de la multiplicidad de nodos, descrita en la ecuación (74), es que el perfil $\zeta(x_1)$ es una curva abierta tangente al polígono de control únicamente en el primer y en el último puntos de control.

Capítulo III

El problema de esparcimiento inverso utilizando datos de amplitud y fase

En este capítulo se propone un algoritmo de inversión basado en principios de empacamiento de frentes de onda. Se consideran como datos de entrada la información de amplitud y fase de la onda esparcida en el campo lejano. En la primera sección se describen las características básicas del algoritmo propuesto y se ilustra su desempeño por medio de ejemplos específicos. Posteriormente, se analiza el efecto del muestreo en las reconstrucciones. A su vez, se establecen los límites de aplicación del algoritmo en presencia de ruido y de esparcimiento múltiple. La última sección del capítulo está dedicada al caso particular de la reconstrucción de superficies en medios estratificados. Se considera un sistema con tres capas, se muestran las ecuaciones básicas para la generación de los datos de esparcimiento y también las reconstrucciones obtenidas. Se propone un esquema para corregir la aberración esférica efectiva introducida por la

propagación de las ondas incidentes en un medio distinto al de incidencia. Finalmente, se presentan las conclusiones del capítulo.

III.1 Algoritmo de empatamiento de frentes de onda (AEFO)

La geometría del problema se muestra en la figura 1 del capítulo II. Sin embargo, en lugar de considerar que la superficie es iluminada por una onda plana [ecuación (32)], ahora se supone que el campo incidente está constituido por una superposición de ondas planas viajando en diferentes direcciones.

Se elige la forma particular

$$\begin{aligned} \psi_{inc}(x_1, x_3; \xi, \eta) = & \int_{-\omega/c}^{\omega/c} \frac{dk}{2\pi} P_{inc}(k) \exp [ik(x_1 - \xi)] \times \\ & \times \exp [-i\alpha_0(k)(x_3 - \eta)], \end{aligned} \quad (82)$$

donde $P_{inc}(k)$ representa una función de pupila, dada por

$$P_{inc}(k) = \text{rect} \left(\frac{k}{2k_{\max}} \right), \quad (83)$$

$\text{rect}(x)$ representa la función rectángulo [Goodman, 1968] y $k_{\max} = (\omega/c) \sin \theta_{0m}$. El ángulo θ_{0m} es el máximo valor que pueden tomar los ángulos de incidencia.

Debido a la linealidad del problema, el campo esparcido puede ser escrito como una

superposición de ondas planas, dada por

$$\begin{aligned} \psi_{sc}(x_1, x_3; \xi, \eta) &= \int_{-\omega/c}^{\omega/c} \frac{dk}{2\pi} P_{inc}(k) \exp\{-ik\xi + i\alpha_0(k)\eta\} \times \\ &\times \int_{-\infty}^{\infty} \frac{dq}{2\pi} R_{p,s}(q|k) \exp\{iqx_1 + i\alpha_0(q)x_3\}. \end{aligned} \quad (84)$$

Esta ecuación ilustra el hecho de que si $R_{p,s}(q|k)$ es conocido para todo q y k , la respuesta del sistema debida a un campo incidente arbitrario puede ser calculada.

Ahora se considera la función

$$V(\xi, \eta) = \int_{-\infty}^{\infty} |\psi_{sc}(x_1, x_3; \xi, \eta) + \psi_{ref}(x_1, x_3; \xi, \eta)|^2 dx_1, \quad (85)$$

donde $\psi_{ref}(x_1, x_3; \xi, \eta)$ representa una onda ficticia de referencia que emana de una zona centrada en el punto (ξ, η) . Se escoge esta onda con una forma idéntica a un patrón de difracción, de manera que

$$\begin{aligned} \psi_{ref}(x_1, x_3; \xi, \eta) &= \int_{-\omega/c}^{\omega/c} \frac{dq}{2\pi} P_{ref}(q) \exp[iq(x_1 - \xi)] \times \\ &\times \exp[i\alpha_0(q)(x_3 - \eta)]. \end{aligned} \quad (86)$$

A su vez, $P_{ref}(q)$ está dada por

$$P_{ref}(q) = \text{rect}\left(\frac{q}{2q_{\max}}\right). \quad (87)$$

Al igual que en la ecuación (83), $\text{rect}(x)$ representa una función rectángulo y $q_{\max} = (\omega/c) \sin \theta_{sm}$. Como se verá más claramente en expresiones posteriores, el ángulo θ_{sm} puede interpretarse como el máximo valor que pueden tomar los ángulos de esparcimiento.

Entonces, la función $V(\xi, \eta)$ representa la integral de la intensidad que resulta de la superposición de dos campos que se alejan de la superficie. Más adelante se demostrará que la función $V(\xi, \eta)$ no depende de x_3 , por lo que puede escribirse como

$$V(\xi, \eta) = 2I_0 + 2\Re\{U(\xi, \eta)\}, \quad (88)$$

donde $2I_0$ es una constante dada por

$$2I_0 = \int_{-\infty}^{\infty} [|\psi_{sc}(x_1, x_3; \xi, \eta)|^2 + |\psi_{ref}(x_1, x_3; \xi, \eta)|^2] dx_1, \quad (89)$$

y $\Re\{U(\xi, \eta)\}$ representa la parte real de la función $U(\xi, \eta)$, que ha sido definida como

$$U(\xi, \eta) = \int_{-\infty}^{\infty} \psi_{sc}(x_1, x_3; \xi, \eta) \psi_{ref}^*(x_1, x_3; \xi, \eta) dx_1. \quad (90)$$

Al sustituir las expresiones para los campos incidente y esparcido en la ecuación

(90) se tiene que

$$\begin{aligned}
 U(\xi, \eta) = & \int_{-\infty}^{\infty} \left[\int_{-\omega/c}^{\omega/c} \frac{dk}{2\pi} P_{inc}(k) \exp\{-ik\xi + i\alpha_0(k)\eta\} \right. \\
 & \times \int_{-\infty}^{\infty} \frac{dq}{2\pi} R_{p,s}(q|k) \exp\{iqx_1 + i\alpha_0(q)x_3\} \times \\
 & \left. \times \int_{-\omega/c}^{\omega/c} \frac{dp}{2\pi} P_{ref}^*(p) \exp\{ip(x_1 - \xi) - i\alpha_0(p)(x_3 - \eta)\} \right] dx_1. \quad (91)
 \end{aligned}$$

Ahora bien, como

$$\int_{-\infty}^{\infty} \exp\{i(q-p)x_1\} dx_1 = 2\pi\delta(q-p), \quad (92)$$

se llega a que

$$U(\xi, \eta) = \int_{-\omega/c}^{\omega/c} \frac{dk}{2\pi} \int_{-\omega/c}^{\omega/c} \frac{dq}{2\pi} P_{inc}(k) P_{ref}(q) R_{p,s}(q|k) \Phi(q, k|\xi, \eta), \quad (93)$$

donde se ha definido la función

$$\Phi(q, k|\xi, \eta) = \exp\{-i(k-q)\xi + i(\alpha_0(k) + \alpha_0(q))\eta\}. \quad (94)$$

El algoritmo propuesto está basado en la premisa de que la función $F_{real}^{p,s}(\xi, \eta) = 2\Re\{U(\xi, \eta)\}$ es maximizada o minimizada (dependiendo de la polarización) cuando $(\xi, \eta) = (x_1, \zeta(x_1))$; esto es, cuando (ξ, η) coincide con un punto de la superficie. Esta idea puede visualizarse haciendo uso de la figura 7. En la figura 7 (a) se tiene un espejo

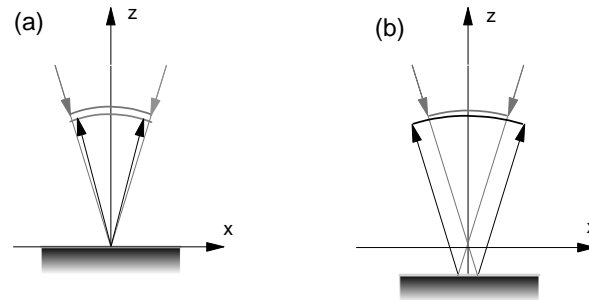


Figura 7: Geometría para ilustrar el principio de operación del algoritmo de empata-miento de frentes de onda.

plano sobre el que incide un haz convergente cuyo foco es el origen de coordenadas. Si se supone un haz de referencia divergente emanando del origen, es claro que el frente de onda de referencia coincidirá perfectamente con el del haz reflejado. Si los haces tienen la misma fase se tendrá interferencia constructiva, aunque esto puede cambiar debido a variaciones de fase en la reflexión de la luz en el espejo. Por otro lado, si existe un desenfoque como el mostrado en la figura 7(b), los frentes de onda del haz de referencia y del reflejado difieren, y este hecho resulta en una intensidad integrada intermedia. Para el caso de superficies rugosas la situación se complica. Más que intentar una visualización, se mostrará que el algoritmo funciona para diversos casos y posteriormente se presentará un breve estudio sobre sus limitaciones. Aunque el tratamiento parece ser muy distinto, puede demostrarse que existe una relación entre este algoritmo y el algoritmo basado en un microscopio confocal interferométrico propuesto por Quartel y Sheppard [Quartel y Sheppard, 1996b]. Por otro lado, puede mostrarse que el algoritmo basado en microscopía confocal [Quartel y Sheppard, 1996a]

maximiza la función

$$F_{\text{conf}}(\xi, \eta) = |U(\xi, \eta)|^2. \quad (95)$$

Como se ilustrará en la siguiente sección, la diferencia principal entre el algoritmo de empataamiento de frentes de onda y el algoritmo confocal reside en que el mapa correspondiente a la función $F_{\text{real}}^{p,s}(\xi, \eta) = 2\Re\{U(\xi, \eta)\}$ consiste en una estructura de franjas que, en principio, asemejan la forma del perfil que se quiere reconstruir. Por otro lado, el mapa de la función (95) es una sola franja brillante, proporcional a la intensidad, que sigue la forma del perfil a reconstruir.

Es importante hacer notar que aunque el algoritmo propuesto(AEFO) sintetiza un haz convergente similar al que se obtendría al utilizar una lente, no es necesario utilizar ningún sistema de formación de imágenes en la adquisición de los datos. El algoritmo representa una manera de procesar la información de esparcimiento para recuperar el perfil que dió origen a ésta.

III.1.1 Reconstrucción de una superficie rugosa utilizando el AEFO

En esta etapa de la exposición se considera necesario un ejemplo para ilustrar el desempeño del algoritmo de empataamiento de frentes de onda. Como primer paso se generó, mediante el método espectral descrito en la sección II.3.1, una superficie que constituye una realización de un proceso aleatorio gaussiano con función de correlación también gaussiana. El perfil resultante tiene una longitud de correlación $a = 2.0\lambda$ y

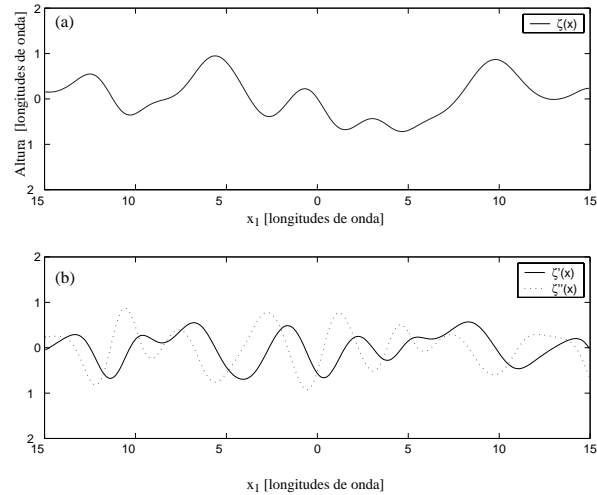


Figura 8: (a) Perfil utilizado para la generación de los datos de esparcimiento (b) Primera y segunda derivadas del perfil.

una desviación estándar de alturas $\delta = 0.5\lambda$. A su vez, contiene un número total de puntos de muestreo $N = 300$, separados entre sí por una distancia $\Delta x = \frac{\lambda}{10}$. Para este caso particular, la longitud total de la superficie fue $L = 30\lambda$. El perfil se muestra, junto con su primera y su segunda derivadas, en la figura 8.

Inicialmente la superficie se supuso perfectamente conductora y los datos de esparcimiento se obtuvieron al iluminarla con ondas planas a diferentes ángulos de incidencia. Se consideraron 151 ángulos de incidencia y de esparcimiento. El máximo valor que éstos podían tomar se fijó en 80° .

En la figura 9 se muestran la amplitud y la fase correspondientes a la amplitud compleja esparcida por la superficie en la figura 8. Los datos pueden visualizarse como una matriz de 151×151 elementos en la que cada columna corresponde a un patrón de esparcimiento generado al iluminar a la superficie con un ángulo de incidencia específico.

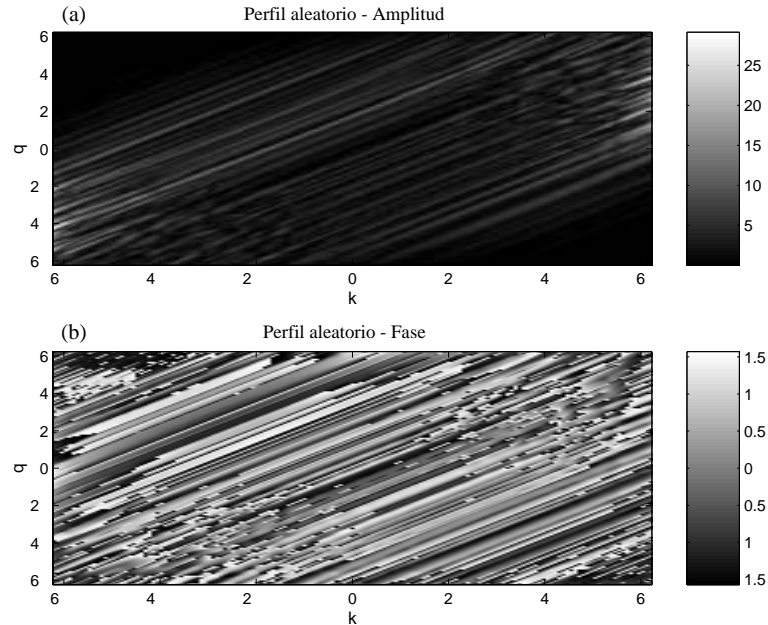


Figura 9: Datos de esparcimiento en función de los números de onda q y k . (a) Amplitud, (b) Fase.

En la figura 10 se presentan los resultados obtenidos al aplicar el algoritmo propuesto utilizando los datos de la figura 9. El mapa de la función $F_{real}^{p,s} = \Re\{U(\xi, \eta)\}$, para las polarizaciones p y s , se muestra en las figuras 10(a) y 10(b), respectivamente. Se observa un patrón de franjas que sigue la forma del perfil original, dibujado con una línea continua para facilitar la visualización. Cuando se considera polarización p , se tiene una franja brillante que sigue la forma del perfil y que coincide con su posición. El mismo comportamiento tiene lugar para la polarización s . Sin embargo, en ese caso, es una franja oscura la que describe al perfil. La diferencia existente entre las reconstrucciones obtenidas con ambas polarizaciones se debe a un cambio de π radianes en la fase, producido por la reflexión de las ondas con polarización s en el conductor perfecto.

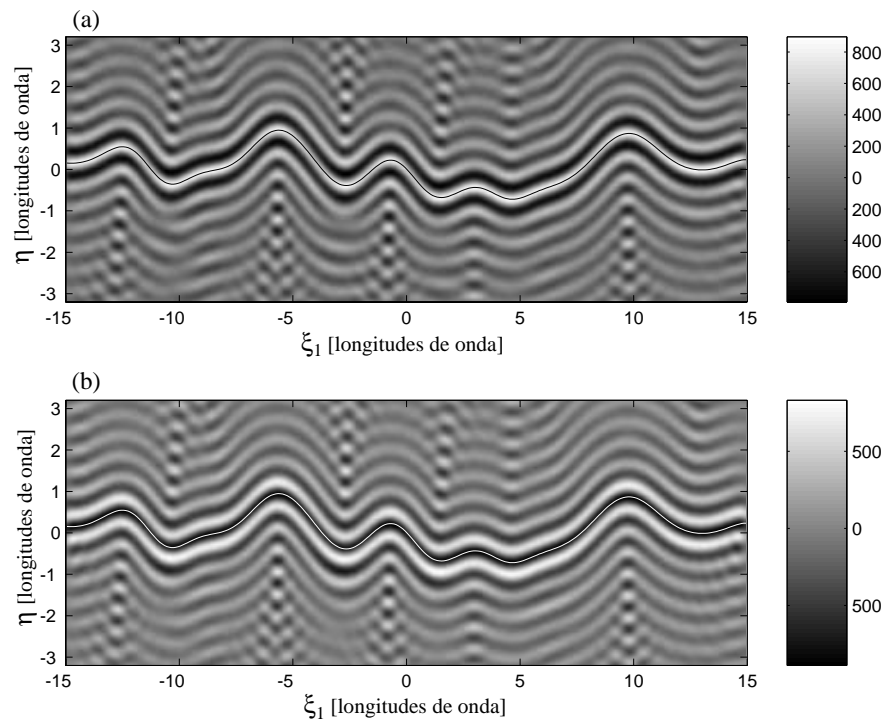


Figura 10: Mapas de la función $F_{real}^{p,s}(\xi, \eta)$ utilizando los datos de la figura 9:(a) Polarización p, (b) polarización s. El perfil original se muestra con una línea continua y contraste invertido.

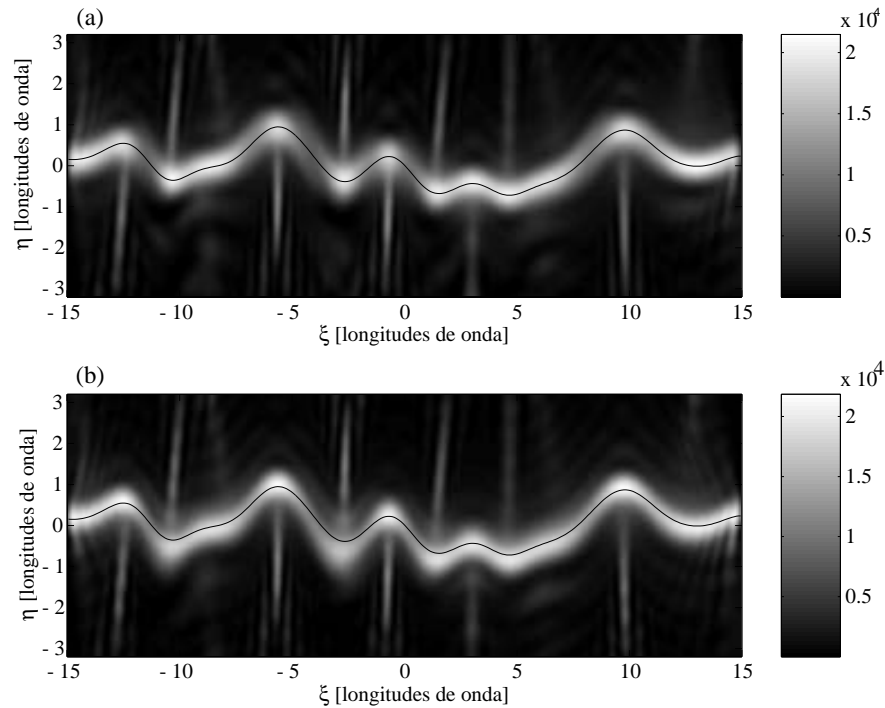


Figura 11: Mapa de la función $F_{conf}^{p,s}(\xi, \eta)$ utilizando los datos de la figura 9:(a) Polarización p, (b) polarización s. El perfil original se muestra con una línea continua.

Con el fin de comparar los resultados obtenidos con los reportados por Quartel y Sheppard en [Quartel y Sheppard, 1996a], se presentan, en la figura 11, los mapas de la ecuación (95) correspondientes a las polarizaciones p y s. Se observa una clara semejanza del perfil original, dibujado con una línea oscura, con la franja brillante en cada una de las figuras.

En esta tesis, se consideró la posibilidad de hacer más estable la reconstrucción mediante un esquema de disminución de múltiple esparcimiento. Se denota por $R_{p,s}^{(1)}(q|k)$ a la amplitud esparcida, calculada con la aproximación de Kirchhoff, y por $R_{p,s}^{(n)}(q|k)$ a la solución de orden n obtenida al resolver las ecuaciones de esparcimiento en forma iterativa (ver apéndice B). Esta solución aproximada puede interpretarse como la am-

plitud esparcida debida a n eventos de esparcimiento. Se ha demostrado que [Michel *et al.*, 1992],

$$R_s^{(n)}(q|k) = (-1)^n R_p^{(n)}(q|k) \quad n > 0, \quad (96)$$

Entonces, con el fin de reducir los efectos de esparcimiento múltiple, específicamente los debidos a eventos de esparcimiento de orden par, puede definirse la función tal que

$$F_{real}^{s-p}(\xi, \eta) = F_{real}^s(\xi, \eta) - F_{real}^p(\xi, \eta). \quad (97)$$

Para esto, se utilizó el hecho de que $U(\xi, \eta)$ constituye una transformación lineal de $R(q|k)$.

En la figura 12 se muestra el perfil estimado al encontrar los máximos de la función $F_{real}^{s-p}(\xi, \eta)$. Es posible utilizar algoritmos de detección más sofisticados involucrando, por ejemplo, esquemas de interpolación parabólica [Quartel y Sheppard, 1996b]. Sin embargo, en este trabajo solamente se detectaron los máximos absolutos del mapa de la ecuación (97) que, como se observa, coinciden totalmente con el perfil original utilizado para generar los datos de esparcimiento.

El algoritmo propuesto también se probó en la reconstrucción de superficies dieléctricas y metálicas, para lo cual se eligieron vidrio ($\varepsilon_d(\omega) = 2.25$) y plata ($\varepsilon_m(\omega) = -17.2 + i0.498$). En este caso, los datos de esparcimiento necesarios fueron calculados utilizando la ecuación (27) y resolviendo el sistema de ecuaciones integrales acopladas dado por las ecuaciones (17) y (18). En la figura 13 se muestran los resultados obtenidos para cuando se tiene polarización p. Se observa un comportamiento similar al que

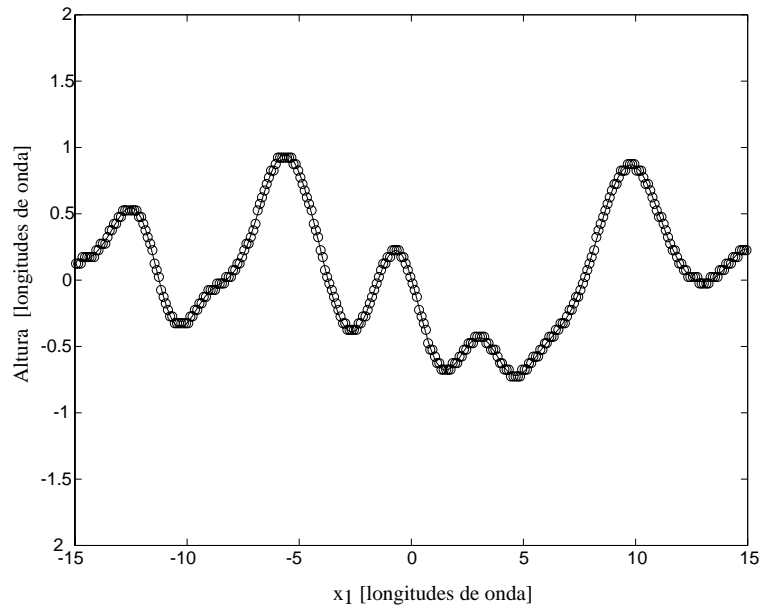


Figura 12: Perfil estimado a partir de los máximos de la función $F_{real}^{s-p}(\xi, \eta)$

se tiene en el caso de un conductor perfecto. Es decir, que aunque el contraste de las franjas es menor que en el caso anterior, el perfil original coincide con la franja brillante (máximo). Sin embargo, debe notarse que la estructura de franjas presenta una ligera asimetría con respecto a la posición del perfil. Este hecho, reportado como resultado del trabajo experimental realizado por [Cox *et al.*, 1982], sugiere la posibilidad de utilizar este método no únicamente para la reconstrucción del perfil de la muestra, sino también la determinación de sus propiedades ópticas.

El muestreo

Un aspecto importante en la determinación de la amplitud esparcida $R_{p,s}(q | k)$, ya sea en forma numérica o experimental, es que se lleva a cabo mediante el uso de un conjunto discreto de valores de q y de k . Entonces, es importante determinar el número

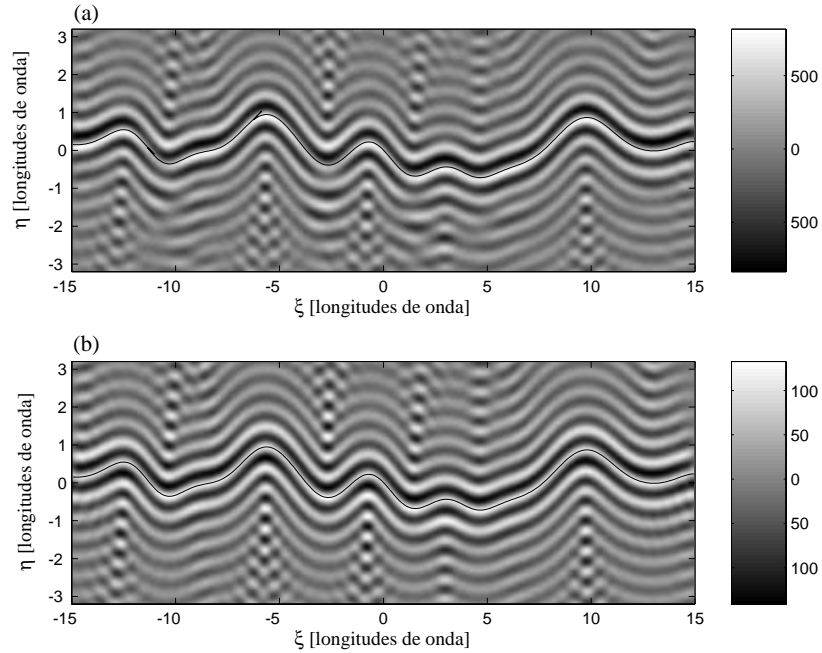


Figura 13: Reconstrucciones obtenidas al utilizar a) Plata, b) vidrio. En ambos casos se consideró polarización p.

mínimo de ángulos de incidencia y de esparcimiento requeridos por el algoritmo para evitar redundancia en la información y para muestrear adecuadamente los patrones de esparcimiento.

En la sección anterior se consideró el campo incidente dado por la ecuación (82). Para facilitar la visualización en la determinación del muestreo óptimo, se considera el caso $\xi = \eta = 0$. Si además $x_3 = 0$, la ecuación se reduce a

$$\psi(x_1, 0)_{inc} = \int_{-\infty}^{\infty} \frac{dk}{2\pi} P_{inc}(k) \exp(ikx_1), \quad (98)$$

que puede interpretarse como una transformada de Fourier inversa. Al invocar la

definición del par de transformadas de Fourier a través de las expresiones:

$$F(k) = \int_{-\infty}^{\infty} dx_1 f(x_1) \exp(-ikx_1), \quad (99)$$

$$f(x_1) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} F(k) \exp(ikx_1), \quad (100)$$

la ecuación del campo incidente en función de x_1 puede reescribirse como

$$\psi(x_1, 0)_{inc} = F^{-1} \{P_{inc}(k)\}, \quad (101)$$

donde $F^{-1} \{ \}$ es el operador transformada de Fourier inversa y $P_{inc}(k)$ es la función de pupila. Considerando un muestreo uniforme, la función de pupila se escribe en la forma

$$P_{inc}(k) = \text{rect} \left(\frac{k}{2k_{\max}} \right) \sum_{n=-\infty}^{\infty} \delta(k - n\Delta k), \quad (102)$$

donde Δk representa el intervalo de muestreo y k_{\max} el máximo de la componente paralela del vector de onda utilizado. Al sustituir la expresión anterior en la ecuación del campo incidente se obtiene

$$\psi(x_1)_{inc} = F^{-1} \left\{ \text{rect} \left(\frac{k}{2k_{\max}} \right) \sum_{n=-\infty}^{\infty} \delta(k - n\Delta k) \right\}. \quad (103)$$

Empleando el teorema de la convolución, se llega a que

$$\psi(x_1, 0)_{inc} = F^{-1} \left\{ \text{rect} \left(\frac{k}{2k_{\max}} \right) \right\} \otimes F^{-1} \left\{ \sum_{n=-\infty}^{\infty} \delta(k - n\Delta k) \right\}, \quad (104)$$

donde

$$F^{-1} \left\{ \text{rect} \left(\frac{k}{2k_{\max}} \right) \right\} = \frac{k_{\max}}{\pi} \text{sinc} (k_{\max} x_1), \quad (105)$$

$$F^{-1} \left\{ \sum_{n=-\infty}^{\infty} \delta (k - n\Delta k) \right\} = \frac{2\pi}{\Delta k} \sum_{n=-\infty}^{\infty} \delta \left(x_1 - \frac{2\pi}{\Delta k} n \right). \quad (106)$$

Finalmente, después de algunas sustituciones y de recordar que la convolución de una función con una delta de Dirac dá como resultado la función original centrada en la posición de la delta, se encuentra que

$$\psi(x_1, 0)_{inc} = \frac{2k_{\max}}{\Delta k} \sum_{n=-\infty}^{\infty} \text{sinc} \left(k_{\max} \left(x_1 - \frac{2\pi}{\Delta k} n \right) \right). \quad (107)$$

La expresión anterior representa la transformada de Fourier inversa de la función de pupila muestreada. Se tiene entonces una serie de réplicas separadas entre sí por una distancia $\frac{2\pi}{\Delta k}$.

Por otro lado, el teorema del muestreo establece que para poder recuperar la información contenida en una señal con ancho de banda limitado, la frecuencia de muestreo debe ser igual o mayor al doble de la frecuencia máxima de la señal original; esto permitiría que en el dominio de las frecuencias, la réplicas del espectro de la señal estuvieran lo suficientemente separadas entre sí para evitar traslapes, ya que éstos generarían un deterioro en la información.

En el problema que se analiza en este trabajo, el muestreo se realiza en el espacio de Fourier, y los traslapes deben evitarse en el espacio de coordenadas. Es claro

entonces, que la longitud de la superficie juega un papel importante. La superficie puede representarse por:

$$\zeta(x_1) = \text{rect}\left(\frac{x_1}{L}\right) \varsigma(x_1), \quad (108)$$

donde $\varsigma(x_1)$ es la función que describe al perfil superficial y L es la longitud de la superficie.

Si la idea es reconstruir el perfil original tomando el menor número de muestras posible, sin que esto vaya en detrimento de la calidad de la reconstrucción, la condición que debe satisfacerse para tener una sola réplica del haz incidente dentro de la superficie es que la distancia de separación entre réplicas sea mayor que la longitud física de la superficie. Es decir,

$$\Delta x > L. \quad (109)$$

De la expresión para el campo incidente y sus réplicas, se tiene que

$$\Delta x = \frac{2\pi}{\Delta k}. \quad (110)$$

Además, sabemos que

$$\Delta k = \frac{2k_{\max}}{N}, \quad (111)$$

donde N es el número de ángulos de incidencia y k_{\max} está determinado por el máximo valor que puede tomar el ángulo de incidencia.

Al combinar las ecuaciones anteriores, se obtiene una expresión que relaciona la longitud física de la superficie con el mínimo número de ángulos de incidencia que deben considerarse para obtener una buena reconstrucción del perfil. Se tiene que

$$N > \frac{Lk_{max}}{\pi}. \quad (112)$$

Si, por ejemplo, se considera la superficie de la figura 8, cuya longitud es $L = 30\lambda$ y además se supone que $\theta_{0m} = 80^\circ$, entonces $k_{max} = \frac{2\pi}{\lambda}\text{sen}(80^\circ)$ y por lo tanto

$$N \approx 60. \quad (113)$$

Con base en el ejemplo anterior, en la figura 14 se muestran los resultados obtenidos al considerar 151, 61 y 30 ángulos de incidencia. Para hacer más clara la exposición, se incluye en las gráficas el perfil de la superficie iluminada.

La figura 14(a) muestra el campo incidente, enfocado en el origen del sistema de coordenadas, cuando se tienen 151 ángulos de incidencia. En la figura 14(b) se redujo de 151 a 61 el número de ángulos de incidencia. Con respecto a la figura 14(a), se observa la aparición de réplicas del campo incidente sobre la superficie. Una disminución en el número de ángulos de incidencia genera una reducción del período de separación entre réplicas. Sin embargo, se observa que mientras este período sea mayor que la longitud de la superficie, ésta será iluminada por un solo haz. Por el contrario, si el número de ángulos de muestreo disminuye aún más [figura 14(c)] el período de separación puede llegar a ser menor que la longitud de la superficie, dando como resultado que ésta sea

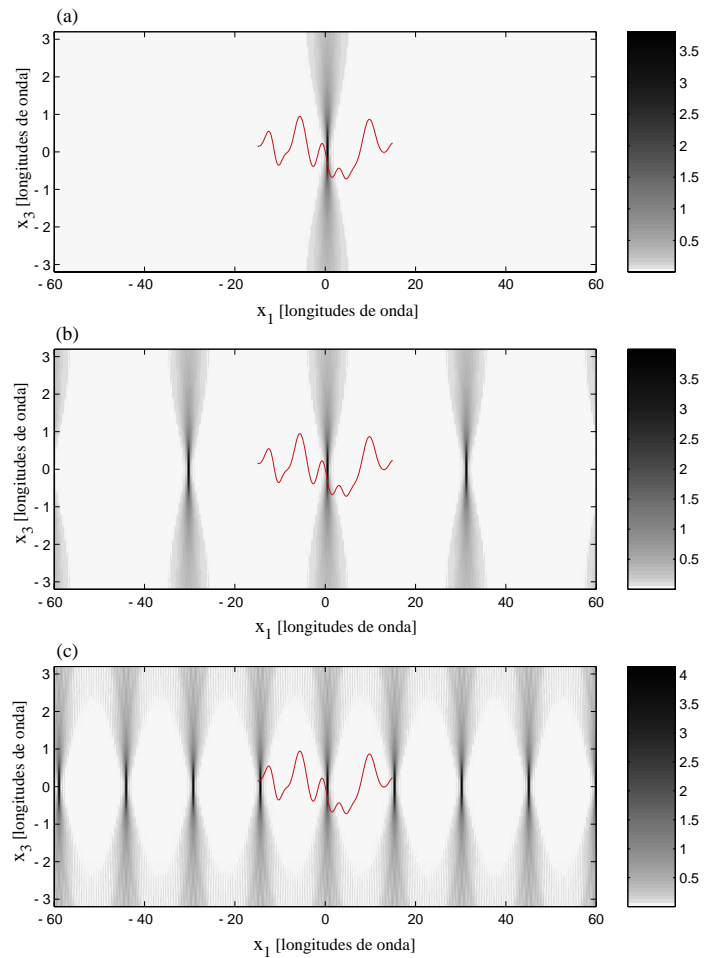


Figura 14: Ilustración del campo incidente cuando se tienen: (a) 151, (b) 61 y (c) 30 ángulos de incidencia, respectivamente.

iluminada por varios haces. Es claro que la consecuencia de un muestreo inadecuado será la interferencia entre los patrones de esparcimiento generados por diferentes zonas de la superficie y este hecho se verá reflejado en la pérdida de contraste en las franjas y en ambigüedades en ciertas zonas, como puede verse en la figura 15.

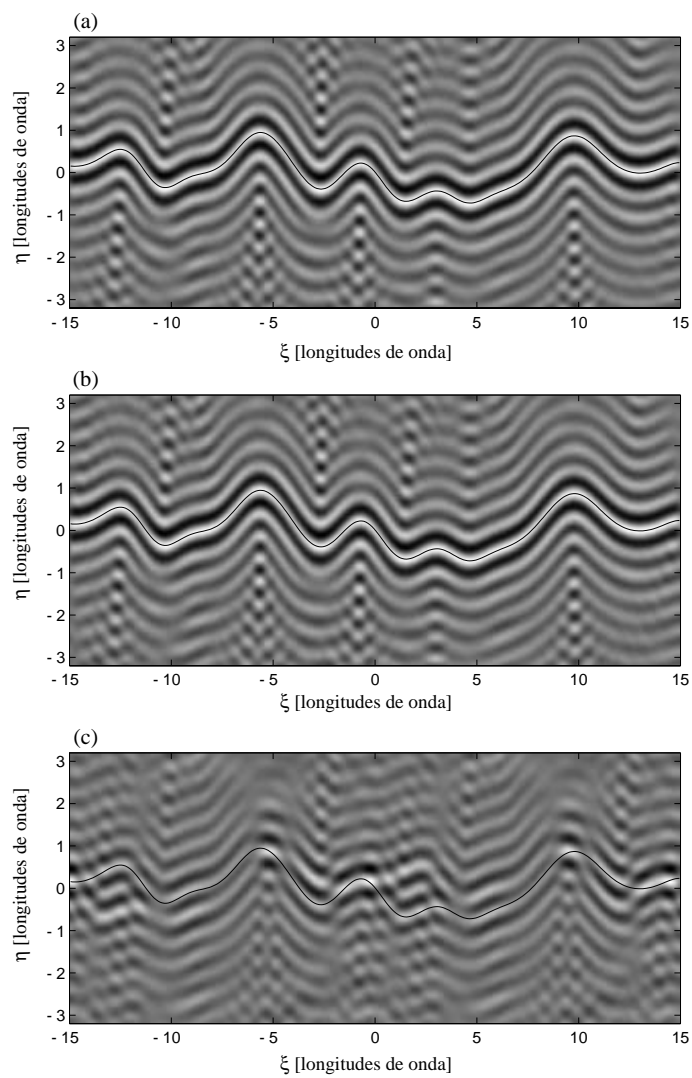


Figura 15: Mapa de $F_{real}^p(\xi, \eta)$ para la superficie perfectamente conductora mostrada en la figura 8, considerando (a) 151, (b) 61 y (c) 30 ángulos de incidencia, respectivamente.

El efecto del ruido

Un aspecto que debió explorarse fue la sensibilidad que presentaba el algoritmo de empatación de frentes de onda al ruido. Con el fin de lograr este objetivo se multiplicó la matriz de datos de esparcimiento $R_{p,s}(q|k)$ por una variable aleatoria compleja $A_N(q|k)$, para la cual se escogieron las siguientes propiedades:

$$A_N(q|k) = a_n(q|k) \exp \{ip_n(q|k)\}, \quad (114)$$

donde el módulo $a_n(q|k)$ se definió por la relación

$$a_n(q|k) = \frac{1}{\sqrt{2}} \sqrt{x^2 + y^2}, \quad (115)$$

siendo x e y variables aleatorias gaussianas con promedio cero y varianza uno. De igual forma, el término de fase p_n es una variable aleatoria gaussiana con promedio cero y varianza uno. Las variables x, y y p_n son estadísticamente independientes.

Al multiplicar los datos $R_{p,s}(q|k)$ por $A_N(q|k)$, se obtiene una nueva matriz

$$R_N(q|k) = R_{p,s}(q|k)A_N(q|k). \quad (116)$$

La intensidad tendrá entonces la forma:

$$\mathcal{I}_N(q|k) = |R_N(q|k)|^2 = I(q|k) |A_N(q|k)|^2, \quad (117)$$

donde se ha definido

$$I(q|k) = |R_{p,s}(q|k)|^2, \quad (118)$$

La intensidad ruidosa promedio está dada por

$$\langle \mathcal{I}_N(q|k) \rangle = I(q|k) \langle |A_N(q|k)|^2 \rangle, \quad (119)$$

donde

$$\langle |A_N(q|k)|^2 \rangle = 1 \quad (120)$$

y por lo tanto

$$\langle \mathcal{I}_N(q|k) \rangle = I(q|k). \quad (121)$$

Este último resultado implica que la intensidad "ruidosa" es, en promedio, igual a la intensidad sin ruido.

Por otro lado, siguiendo el proceso descrito en [Goodman, 1985], puede mostrarse que la función de densidad de probabilidad para la amplitud $a_n(q|k)$, tiene la forma

$$P_{a_n}(x) = \begin{cases} \frac{x}{\sigma_a^2} \exp \left\{ -\frac{x^2}{2\sigma_a^2} \right\} & \text{para } x > 0 \\ 0 & \text{para cualquier otro caso.} \end{cases}$$

Esta función de densidad de probabilidad se conoce como distribución de Rayleigh.

A partir de este resultado, mediante un procedimiento análogo al descrito en [Dainty, 1975], puede mostrarse que la distribución de probabilidad de la intensidad es una

exponencial decreciente de la forma:

$$P_{I_N} = \frac{1}{\langle I_N \rangle} \exp \left\{ -\frac{x}{\langle I_N \rangle} \right\} \quad (122)$$

La fase, por otro lado, posee una distribución gaussiana dada por

$$P_p(x) = \frac{x}{\sqrt{2\pi}\sigma_p} \exp \left\{ -\frac{x^2}{2\sigma_p^2} \right\} \quad (123)$$

donde $-\infty < x < \infty$ y $\sigma_p > 0$.

Como primer resultado, y con el fin de facilitar la comparación, en la figura 16 se muestran el módulo de la amplitud y la fase cuando no hay ruido, y también en presencia de éste. Debe señalarse que en las simulaciones realizadas en este trabajo, se consideró para la fase una $\sigma_p = \pi/4$. Se observa que la fase asemeja un patrón de moteado y sería posible pensar que toda la información se ha perdido. Sin embargo, en la figura 17 se muestran las reconstrucciones obtenidas antes y después de agregar el ruido. Puede verse que la reconstrucción se llevo a cabo sin gran deterioro. Por otro lado, al llevar a cabo las simulaciones variando σ_p , se llegó a que si ésta es mayor o igual $\pi/2$, entonces no es posible recuperar el perfil.

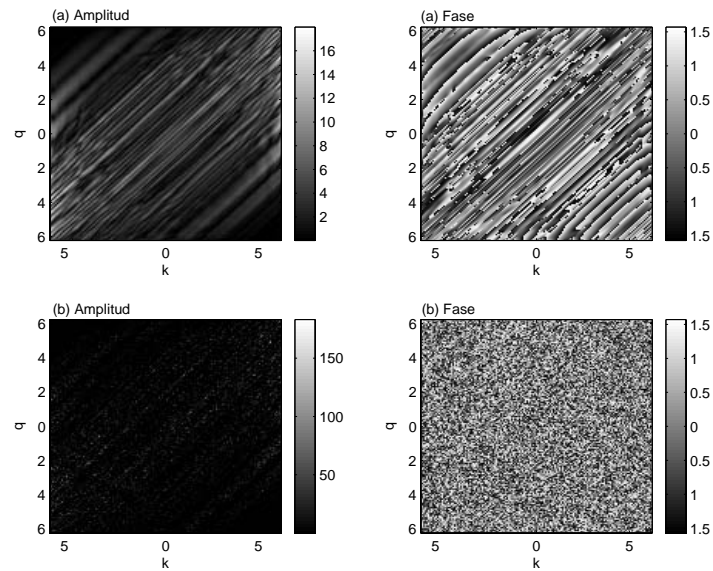


Figura 16: Amplitud y fase de los datos de esparcimiento, (a) sin ruido y (b) con ruido.

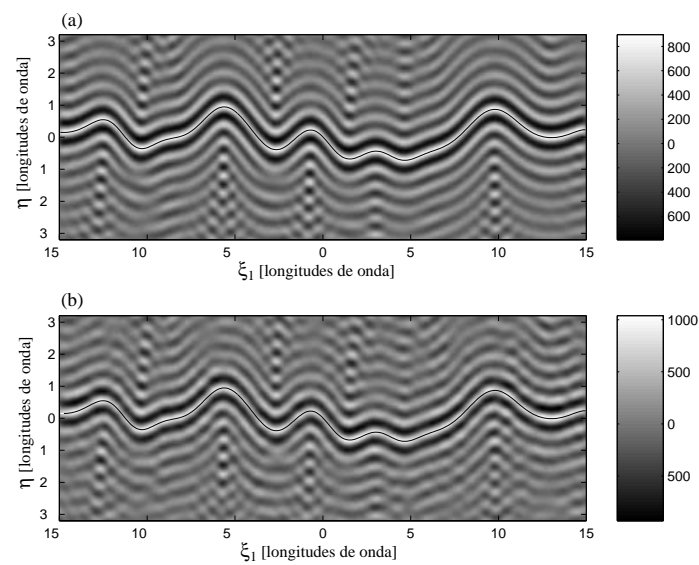


Figura 17: Mapa de $F_{real}^p(\xi, \eta)$ utilizando los datos de esparcimiento de la figura 16,(a) Datos sin ruido y (b) datos con ruido.

III.1.2 Aplicación del AEFO en sistema estratificado con tres capas

Una vez que la aplicabilidad del algoritmo propuesto había sido probada para el caso de superficies dieléctricas y superficies metálicas, se procedió a analizar el caso correspondiente a un medio estratificado. En este problema particular se consideran únicamente tres capas. Esto último no es restrictivo, ya que el algoritmo se puede aplicar a cualesquier datos de esparcimiento. Sin embargo, como se verá en breve, la consecuencia de tener más medios introduce una mayor complejidad en los cálculos de esparcimiento directo. En la figura 18 se muestra la geometría del problema. La región $x_3 > d/2 + \zeta(x_3)$ se consideró como el vacío y los medios restantes pueden ser dieléctricos o metales. A su vez, la forma de las superficies que conforman las fronteras se consideran como funciones arbitrarias separadas por una distancia d , que define el espesor promedio de la película (ver figura 18)

Al igual que en la sección 3, el campo esparcido $\psi_{sc}(x_1, x_3)$ puede calcularse con la ecuación (27), donde a su vez el espectro angular tiene la misma forma que la ecuación (9). Sin embargo, la determinación de las funciones fuente $\varphi(x_1|\omega)$ y $\chi(x_1|\omega)$ se torna complicada debido a que es necesario conocer no sólo el campo en la región $d/2 + \zeta_1(x_1) > x_3$ sino también dentro de la región $-d/2 + \zeta_2(x_1) > x_3$. Mediante la aplicación del teorema integral de Green y tomando en cuenta las condiciones de frontera respectivas a cada una de las regiones del medio, se ha mostrado [Lu *et al.*, 1991; Gu *et al.*, 1993] que las funciones fuente se obtienen resolviendo un sistema de

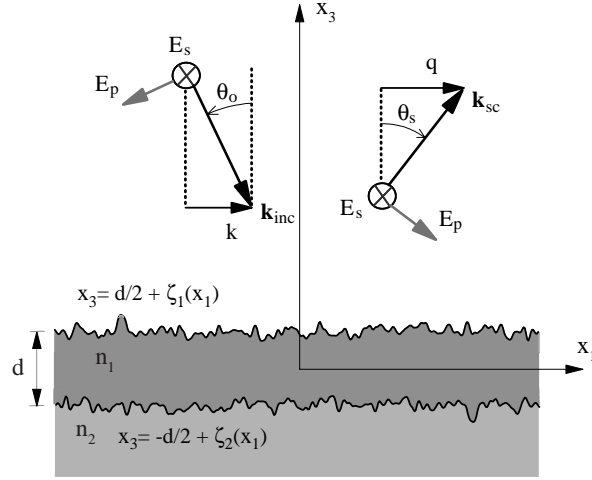


Figura 18: Geometría del problema de esparcimiento para el caso de superficies estratificadas.

cuatro ecuaciones integrales acopladas, que puede escribirse de la siguiente forma

$$\begin{aligned} \varphi^{(1)}(x_1|\omega)\theta(x_3 - (d/2 + \zeta_1(x_1))) &= \psi_{inc}(x_1|w) + \\ &+ \lim_{\epsilon \rightarrow 0} \int_{-\infty}^{\infty} dx'_1 [\varphi^{(1)}(x'_1|\omega)H^{(0)}(x_1|x'_1) - \chi^{(1)}(x'_1|\omega)L^{(0)}(x_1|x'_1)], \end{aligned} \quad (124)$$

$$\begin{aligned} 0 &= \lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 [\varphi^{(1)}(x'_1|\omega)H^{(11)}(x_1|x'_1) - \epsilon_d(\omega)\chi^{(1)}(x'_1|\omega)L^{(11)}(x_1|x'_1)] - \\ &- \lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 [\varphi^{(2)}(x'_1|\omega)H^{(12)}(x_1|x'_1) - \chi^{(2)}(x'_1|\omega)L^{(12)}(x_1|x'_1)], \end{aligned} \quad (125)$$

$$\begin{aligned}
\varphi^{(2)}(x_1|\omega) = & \\
& \lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 [-\varphi^{(1)}(x'_1|\omega)H^{(21)}(x_1|x'_1) - \epsilon_d(\omega)\chi^{(1)}(x'_1|\omega)L^{(21)}(x_1|x'_1)] - \\
& - \lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 [\varphi^{(2)}(x'_1|\omega)H^{(22)}(x_1|x'_1) - \chi^{(2)}(x'_1|\omega)L^{(22)}(x_1|x'_1)] \quad (126)
\end{aligned}$$

y

$$0 = \lim_{\epsilon \rightarrow 0} \frac{1}{4\pi} \int_{-\infty}^{\infty} dx'_1 \left[\varphi^{(2)}(x'_1|\omega)H^{(s)}(x_1|x'_1) - \frac{\epsilon_s(\omega)}{\epsilon_d(\omega)}\chi^{(2)}(x'_1|\omega)L^{(s)}(x_1|x'_1) \right], \quad (127)$$

donde las funciones fuente están definidas por

$$\varphi^{(1)}(x_1|\omega) = \psi^>(x_1, x_3|\omega) \Big|_{x_3=d/2+\zeta_1(x_1)+\epsilon}, \quad (128)$$

$$\chi^{(1)}(x_1|\omega) = \left[-\zeta'_1(x_1)\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] \psi^>(x_1, x_3|\omega) \Big|_{x_3=d/2+\zeta_1(x_1)+\epsilon}, \quad (129)$$

$$\varphi^{(2)}(x_1|\omega) = \psi^<(x_1, x_3|\omega) \Big|_{x_3=-d/2+\zeta_2(x_1)}, \quad (130)$$

$$\chi^{(2)}(x_1|\omega) = \left[-\zeta'_2(x_1)\frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] \psi^<(x_1, x_3|\omega) \Big|_{x_3=-d/2+\zeta_2(x_1)}. \quad (131)$$

A su vez, los kernels de las ecuaciones integrales están definidos por

$$H^{(0)}(x_1|x'_1) = \left[-\zeta'_1(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_0(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (132)$$

$$L^{(0)}(x_1|x'_1) = G_0(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (133)$$

$$H^{(11)}(x_1|x'_1) = \left[-\zeta'_1(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (134)$$

$$L^{(11)}(x_1|x'_1) = G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (135)$$

$$H^{(12)}(x_1|x'_1) = \left[-\zeta'_2(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (136)$$

$$L^{(12)}(x_1|x'_1) = G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=\zeta_1(x_1)+\epsilon}} \quad (137)$$

$$L^{(22)}(x_1|x'_1) = G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (138)$$

$$H^{(22)}(x_1|x'_1) = \left[-\zeta'_2(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (139)$$

$$H^{(21)}(x_1|x'_1) = \left[-\zeta'_1(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (140)$$

$$L^{(21)}(x_1|x'_1) = G_d(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=\zeta_1(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (141)$$

$$H^{(s)}(x_1|x'_1) = \left[-\zeta'_2(x_1) \frac{\partial}{\partial x_1} + \frac{\partial}{\partial x_3} \right] G_s(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (142)$$

$$L^{(s)}(x_1|x'_1) = G_s(x_1, x'_1|x_3, x'_3) \Bigg|_{\substack{x'_3=-d+\zeta_2(x'_1) \\ x_3=-d+\zeta_2(x_1)+\epsilon}} \quad (143)$$

donde $G_0(x_1, x'_1|x_3, x'_3)$, $G_d(x_1, x'_1|x_3, x'_3)$ y $G_s(x_1, x'_1|x_3, x'_3)$ son las funciones de Green del vacío, de la película y del , respectivamente.

Las cuatro ecuaciones integrales acopladas son convertidas en ecuaciones matriciales para su solución numérica [Lu *et al.*, 1991; Gu *et al.*, 1993].

Una vez que la amplitud compleja $R_{p,s}(q|k)$ ha sido determinada, es posible intentar la reconstrucción de los perfiles correspondientes a las superficies $\zeta_1(x_1)$ y $\zeta_2(x_1)$ mediante el cálculo de la función $U(\xi, \eta)$ utilizando la ecuación (93).

La figura 19 muestra los resultados obtenidos al aplicar el AEFO en la reconstrucción de las superficies de un sistema compuesto por las interfaces aire-vidrio-metal. Se

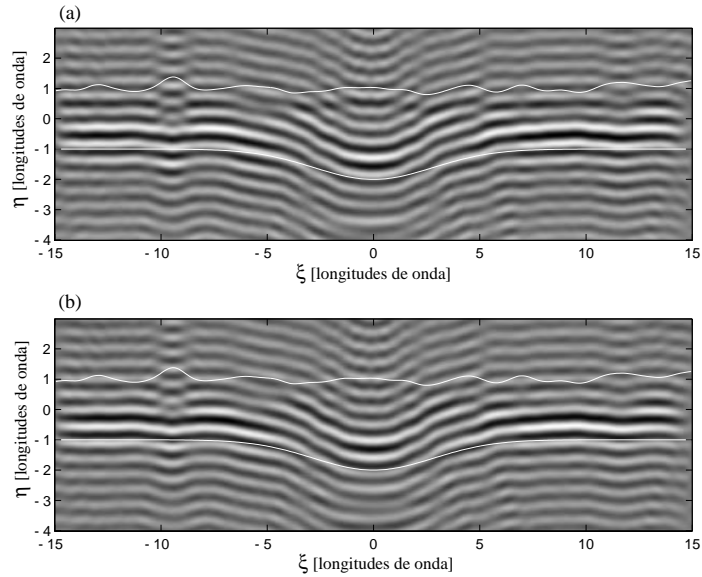


Figura 19: Mapa de $F_{real}^{p,s}(\xi, \eta)$ correspondiente a un medio estratificado constituido por una película de plata recubierta con vidrio. Los perfiles originales se muestran con líneas blancas. Para la superficie superior se consideró una desviación estándar de alturas $\delta = 0.1\lambda$ y una longitud de correlación $a = 1.0\lambda$. Para la superficie inferior se consideró un ancho $w = 4.0\lambda$ y una altura $h = -1.0\lambda$. (a) Polarización s, (b) polarización p.

consideró como sustrato una superficie de plata cuyo perfil era un surco gaussiano de ancho $w = 4.0\lambda$ y de altura $h = -1.0\lambda$. El dieléctrico fue una película de vidrio cuyo perfil era una realización particular de un proceso aleatorio gaussiano con función de correlación también gaussiana y con una desviación estándar de alturas $\delta = 0.1\lambda$ y longitud de correlación $a = 1.0\lambda$. aleatoria. Las constantes dieléctricas fueron $\epsilon_s(\omega) = -17.25 + i0.498$ y $\epsilon_p(\omega) = 2.25$ para el sustrato y para la película, respectivamente. La longitud física de ambas interfaces fue de $L = 30\lambda$. Para los cálculos, se utilizaron 61 ángulos de incidencia y de esparcimiento. El valor del ángulo de corte fue de 80° y el grosor de la película de vidrio fue de $d = 2.0\lambda$.

Puede observarse, para ambas polarizaciones, que la reconstrucción sigue de manera

aproximada la forma correspondiente al perfil de la superficie metálica. Sin embargo, a diferencia del caso en el cual se tiene una sola superficie, existe un deterioro en el patrón de franjas y la posición del perfil original no coincide con las franjas de mejor contraste. Visualizando este hecho en términos del empataamiento de frentes de onda, puede verse que se debe principalmente a la aberración esférica que se produce al enfocar un haz, sin aberraciones en el medio original, a través de una interface. Se puede apreciar que el perfil de la superficie superior no se alcanza a visualizar. Este hecho tiene su origen en el efecto dominante del campo reflejado por la superficie metálica sobre el campo reflejado por la interface vacío-vidrio.

En la figura 20 se presentan los resultados obtenidos al considerar una película dieléctrica sobre un substrato también dieléctrico. Para el material de la película se tomó una constante dieléctrica ($\epsilon_p(\omega) = 1.7689$) y para el substrato se tomó ($\epsilon_s(\omega) = 4.00$). Ambas superficies son realizaciones de un proceso aleatorio gaussiano con función de correlación también gaussiana. Las estadísticas del perfil superior son las mismas que las consideradas para la superficie rugosa mostrada en la figura 19. Para la superficie inferior se tomó una desviación estándar de alturas $\delta = 0.3\lambda$ y una longitud de correlación $a = 3.0\lambda$.

El número de ángulos de incidencia y de esparcimiento fue el mismo que en la simulación anterior, al igual que la longitud física de las superficies. Sin embargo, para esta simulación el grosor de la película se incrementó a $d = 4.0\lambda$.

A diferencia de los resultados mostrados en la figura 19, en el sistema de franjas de la figura 20 no es posible visualizar del todo ninguno de los perfiles originales

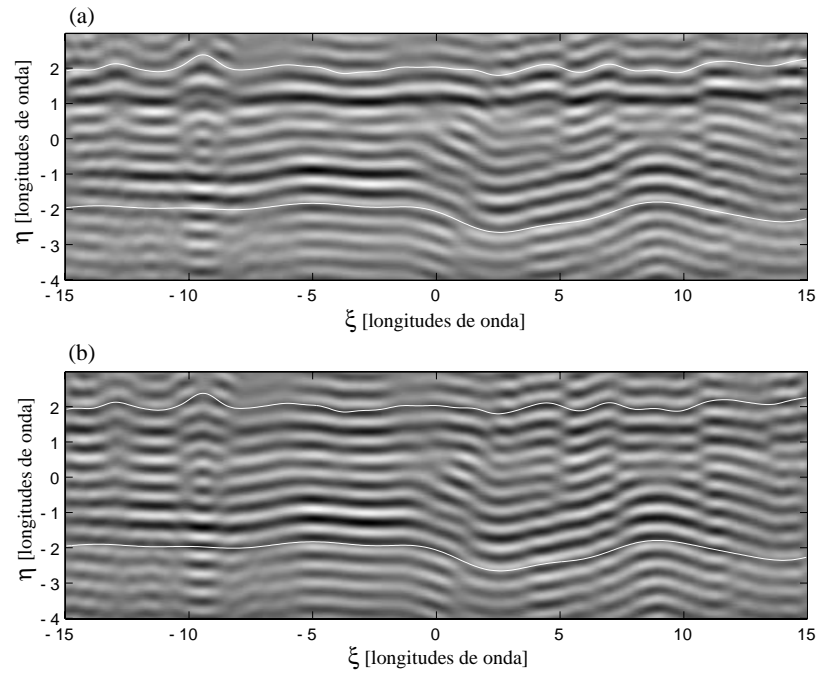


Figura 20: Mapa de $F_{real}^{p,s}(\xi, \eta)$ correspondiente a un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 1.7689$ y un sustrato dieléctrico con $\epsilon_s(\omega) = 4.00$. Ambas superficies son realizaciones de un proceso aleatorio gaussiano. Sus desviaciones estándar de alturas son $\delta_p = 0.1\lambda$ y $\delta_s = 0.3\lambda$ y sus longitudes de correlación son $a_p = 1.0\lambda$ y $a_s = 3.0\lambda$, respectivamente. (a) Polarización s, (b) polarización p.

(dibujados con una línea blanca). A su vez, se observa que las franjas de más alto contraste siguen en forma aproximada la forma de los perfiles. Sin embargo, existe un desplazamiento vertical con respecto a la posición de éstos. Este comportamiento se observa para ambas polarizaciones. Nuevamente es un efecto debido principalmente a la aberración esférica introducida al cambiar la luz de medio.

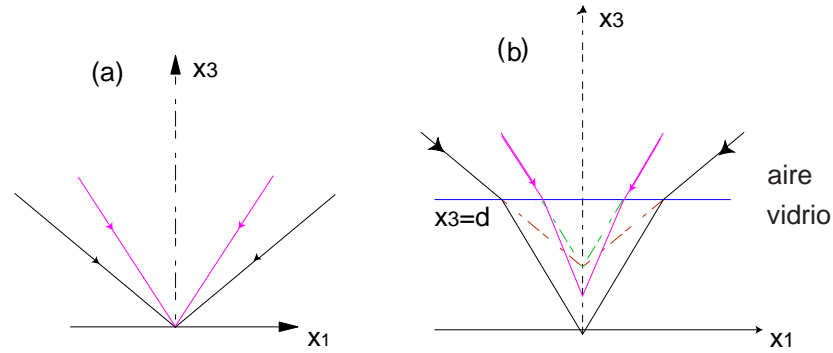


Figura 21: a) Haz sin aberraciones enfocado en el origen, b) Haz enfocado a través de una una interface plana.

Corrección de la aberración esférica introducida al enfocar un haz a través de una interface plana

En la figura 21(a), se muestra un haz convergente enfocado en el origen del sistema de coordenadas. Por otro lado, en la figura 21(b) se muestra la introducción de un material con una interface dieléctrica plana, anterior al foco, a una distancia d del plano $x_3 = 0$. Se ilustra el hecho de que, debido a la refracción en la interface, el foco de los rayos marginales no coincide con el de los rayos paraxiales. Esto representa aberración esférica y tiene como consecuencia un deterioro de la calidad del haz enfocado. La aberración se puede reducir disminuyendo el valor del ángulo de corte. Sin embargo, esto tiene como consecuencia un incremento en la profundidad de foco, que para la presente aplicación representa un deterioro en la resolución vertical. Es decir, que la reducción del ángulo de corte afecta negativamente la localización del perfil. Una manera más adecuada para disminuir la aberración consiste en la introducción de un factor corrector tanto en la pupila de entrada como en la de salida.

Para determinar la forma del factor de corrección se tomará como punto de partida la ecuación (82) y se supondrá por el momento, que el haz es enfocado en el origen, tal como se muestra en la figura 21(a).

$$\psi(x_1, x_3) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} P_{inc}(k) \exp \{ikx_1 - i\alpha_0(k)x_3\}. \quad (144)$$

Si ahora se coloca una interface dieléctrica en $x_3 = z$, es claro que el campo incidente en este plano tiene la forma

$$\psi^>(x_1, d) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} P_{inc}(k) \exp \{ikx_1 - i\alpha_0(k)z\}. \quad (145)$$

Cada onda plana incidente sobre la interface se transmite al otro medio de acuerdo con las ecuaciones de Fresnel. El campo incidente transmitido a través de la interface puede escribirse como

$$\psi^<(x_1, d) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} t(k) P_{inc}(k) \exp \{ikx_1 - i\alpha_0(k)z\}, \quad (146)$$

donde $t(k)$ es el coeficiente de transmisión de Fresnel.

Por otro lado, el campo en la región $x_3 < z$ puede escribirse de manera general como

$$\psi^<(x_1, x_3) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} P_t(k) \exp \{ikx_1 - i\alpha(k)x_3\}, \quad (147)$$

donde $\alpha(k) = \sqrt{(n_t \frac{\omega}{c})^2 - k^2}$. Vale la pena mencionar, que en la expresión anterior se

utilizó la componente del vector de onda del medio incidente, paralela a la interface, y no la del medio dieléctrico, debido a que éstas son iguales(conservación del momento a lo largo de x_1). La ecuación (146) puede considerarse como una condición de frontera, que a su vez debe ser satisfecha por la ecuación (147). Para que esto suceda debe cumplirse que

$$P_t(k) = t(k)P_{inc}(k) \exp \{-i [\alpha_0(k) - \alpha(k)] z\}. \quad (148)$$

De la sustitución de la ecuación (148) en la ecuación (147) se obtiene que el campo en la región $x_3 < z$ tiene la forma

$$\psi^<(x_1, x_3) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} t(k)P_{inc}(k) \exp \{-i [\alpha_0(k) - \alpha(k)] d\} \exp \{ikx_1 - i\alpha(k)x_3\}. \quad (149)$$

Las ecuaciones (147) y (149) expresan el campo en la región $x_3 < d$ en términos del espectro angular del campo en la frontera multiplicado por el propagador $\exp \{ikx_1 - i\alpha(k)x_3\}$.

Al enfocar un haz con una lente sin aberraciones en un medio limitado por una interface dieléctrica, éste introducirá una aberración representada por el factor de fase de la ecuación (148). La aberración puede corregirse incidiendo con un haz aberrado, de tal manera que se cancele la aberración introducida al pasar de un medio a otro. Es claro que la pupila requerida para el haz incidente está dada por el complejo conjugado de la aberración que introduce el medio. Representando por $P_d(k)$ la función de pupila

deseada en el segundo medio, se escoge a la pupila de incidencia de la forma

$$P_{inc}(k) = \frac{1}{t(k)} P_d(k) \exp \{i(\alpha_0(k) - \alpha(k))z\}. \quad (150)$$

Sustituyendo en la ecuación (148), se ve que la pupila en el segundo medio da como resultado la pupila deseada. En la figura 22 se muestran los resultados obtenidos al utilizar las pupilas de incidencia y de referencia con la forma de la ecuación (150). Al llevar a cabo la comparación directa con la figura 19, se hace evidente que la calidad de la reconstrucción mejora notablemente. Esto representa un resultado conocido en microscopía [Török *et al.*, 1995]. El problema de su implementación es que la pupila depende de z , la distancia a la que se desea enfocar en el segundo medio. Es decir, para su implementación en microscopía se requiere de una pupila adaptable. Para la presente aplicación, debido a que los haces incidente y de referencia se sintetizan con los datos de esparcimiento, no resulta práctico implementar estas pupilas.

También en la figura 22, se observa que patrón de franjas sigue la forma de los perfiles originales. Cabe resaltar que para introducir estas pupilas correctoras, se consideró que el perfil superior era conocido, de manera que se podía calcular la distancia z a la que se deseaba enfocar. En la figura 23, se presentan los resultados obtenidos al corregir la aberración para el caso de una película dieléctrica sobre un dieléctrico. Estos resultados deben compararse con la figura 20. Al igual que para cuando el substrato es un metal, se observa una mejora en la calidad de las franjas, permitiendo abordar el problema de reconstruir el perfil. Debido a que resulta poco práctico corregir la

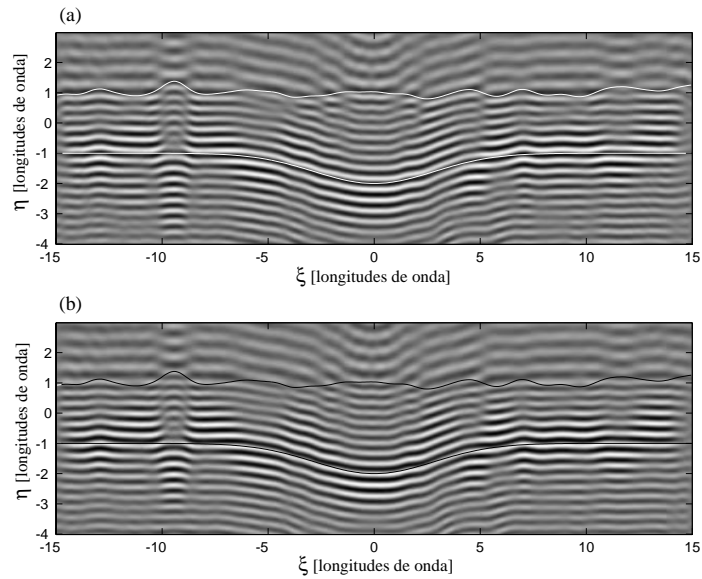


Figura 22: Mapa de $F_{real}^{p,s}(\xi, \eta)$ cuando se corrige la aberración esférica en un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 2.25$ y un sustrato metálico con $\epsilon_s(\omega) = -17.25 + i0.498$. (a) Polarización s, (b) polarización p

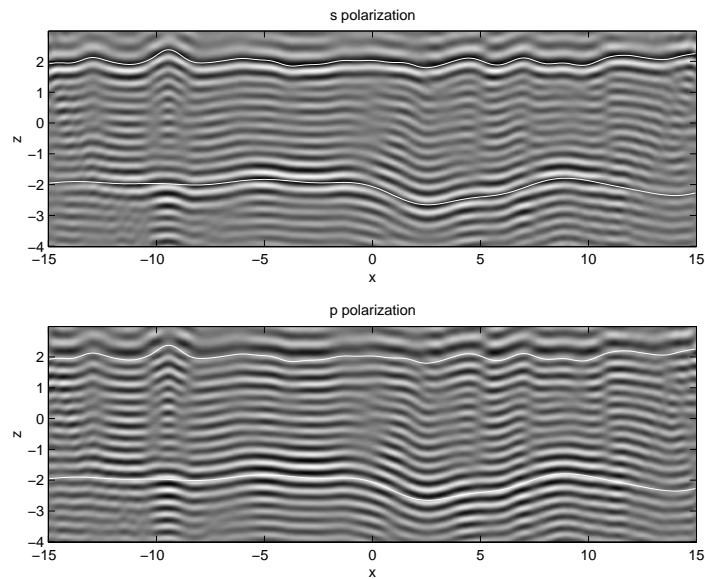


Figura 23: Mapa de $F_{real}^{p,s}(\xi, \eta)$ cuando se corrige la aberración esférica en un medio estratificado compuesto por una película dieléctrica con $\epsilon_p(\omega) = 1.7689$ y un sustrato dieléctrico con $\epsilon_s(\omega) = 4.00$. (a) Polarización s, (b) polarización p.

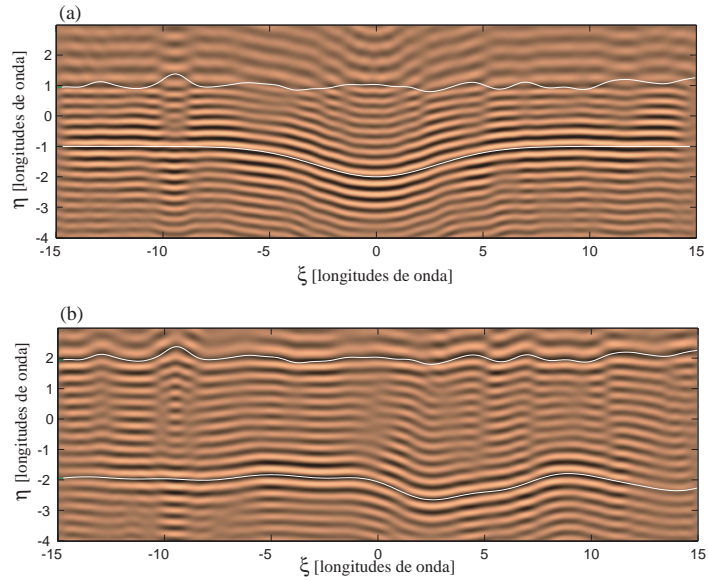


Figura 24: Mapas de la función $F_{real}^{s-p}(\xi, \eta)$ considerando el plano promedio de la interfaz superior y corrigiendo la aberración esférica. (a) Medio estratificado correspondiente a la figura (22), (b) medio estratificado correspondiente a la figura (23).

aberración siguiendo los detalles del perfil superior, se exploró la posibilidad de recuperar el perfil inferior utilizando una distancia z calculada con base en el plano promedio del perfil superior. Resulta interesante que, con esto, el sistema de franjas resultante adquiere un mejor contraste. A su vez, se buscó eliminar las contribuciones debidas a efectos de esparcimiento múltiple mediante uso de la ecuación (97). Los resultados obtenidos para los casos considerados se muestran en la figura 24. De los resultados anteriores surgieron dos posibilidades de aplicación. La primera consiste en la recuperación del perfil de una superficie altamente reflectora que se encuentre cubierta por una película dieléctrica. La otra consistiría en la determinación de los dos perfiles en películas dieléctricas.

III.2 Conclusiones parciales

En este capítulo se estudio un algoritmo de inversión basado en principios de empataamiento de frentes de onda. Su aplicabilidad fue probada para la reconstrucción de superficies unidimensionales perfectamente conductoras, dieléctricas y metálicas. A su vez, se propuso una estrategia de muestreo que relaciona el número de ángulos de incidencia y de esparcimiento con la longitud física de la superficie, evitando así una redundancia en la información utilizada y también el deterioro del mapa de reconstrucción. Con el fin de establecer los límites de validez del algoritmo, se estudió su funcionamiento en presencia de ruido. Los resultados obtenidos considerando ruido multiplicativo muestran que el algoritmo es poco sensible al ruido de amplitud, siendo el ruido de fase el que más afecta los resultados. Si la desviación estándar del ruido de fase es mayor o igual a $\sigma_p = \pi/2$, no se considera factible llevar a cabo la reconstrucción.

Un caso particular que se estudió fue posibilidad de reconstruir las interfaces de un medio estratificado con tres capas. Se mostró, que para mejorar los mapas de reconstrucción obtenidos, es necesario utilizar una pupila adaptable para corregir la aberración esférica que se genera cuando un haz convergente pasa a través de una interface plana.

Capítulo IV

El esparcimiento inverso como un problema de optimización

Con base en los resultados presentados en el capítulo anterior, puede afirmarse que el algoritmo de "empatamiento de frentes de onda" es robusto y posee una buena inmunidad al ruido. Sin embargo, para la aplicación de esta técnica se requiere tanto la amplitud como la fase del campo esparcido por la superficie que se está estudiando. Esto representa un problema, pues la adquisición experimental de estos datos es complicada.

En el capítulo II se mostró la posibilidad de abordar el problema inverso con datos de intensidad, planteándolo como un problema de optimización de parámetros que determinan, de alguna forma, las alturas de la superficie a reconstruir. La primera sección del presente capítulo está dedicada a la descripción de los aspectos principales de tres algoritmos de optimización que, por sus características propias, podrían ser

utilizados en la solución del problema inverso. En la segunda sección se presentan los resultados obtenidos al reconstruir superficies rugosas utilizando la representación espectral descrita en la sección II.3.1. A su vez, en la tercera parte se describen los algoritmos utilizados para la reconstrucción de superficies representadas por curvas spline, como se describió en la sección II.3.2. Se muestra también el efecto de la hibridación (mezcla de algoritmos) en la reconstrucción de superficies. Finalmente, se presentan las conclusiones del capítulo.

IV.1 Algoritmos de optimización

Existen tantos algoritmos de optimización como maneras de clasificarlos. Sin embargo, estos métodos poseen atributos comunes que permiten dividirlos en dos clases [Wehrens y Lutgarde, 2000]:

1. Métodos de optimización *clásicos*: Son aquellos cuya búsqueda del óptimo se basa en el conocimiento *a priori* de alguna característica de la función de costo, tal como su topología o su dirección de máxima variación. La convergencia de estos algoritmos hacia el óptimo está sujeta a que las suposiciones iniciales acerca de la influencia de los parámetros a optimizar sobre la función de costo sean acertadas. Algunos ejemplos de estos métodos para funciones unidimensionales y multidimensionales son la iteración regula falsi, la iteración Newton-Raphson, la interpolación lagrangiana, las estrategias de Newton, los métodos de búsqueda directa y las estrategias de gradiente, por mencionar unos cuantos.

Por otro lado, existen también métodos clásicos que no hacen uso de ningún conocimiento *a priori* de la topología de la función o de sus derivadas y que realizan la búsqueda del óptimo de manera estocástica. La aplicación repetida de los métodos estocásticos conduce a diferentes resultados, por lo que la repetibilidad del resultado final no está asegurada y usualmente es baja. El ejemplo más básico dentro de esta clase es la "búsqueda aleatoria".

2. Métodos de optimización *no clásicos* o *intermedios*: Son aquellos que para hallar el óptimo, hacen uso de las mejores características de los algoritmos descritos anteriormente. Es decir, estos métodos permiten la ocurrencia de procesos aleatorios durante la búsqueda y, al mismo tiempo, poseen la suficiente información acerca de la función objetivo, como para "dirigirse" hacia la dirección en la que se encuentra la solución. Los principios de operación de estos algoritmos se basan en la imitación de ciertos procesos naturales. Por ejemplo, la evolución de los individuos de una población o el enfriamiento térmico de un sistema.

Es importante mencionar que la elección del algoritmo de optimización depende del problema que se esté resolviendo. Propiedades de la función de costo, tales como: su diferenciabilidad, continuidad o número de dimensiones; por mencionar algunas, limitan el conjunto de algoritmos con los cuales es posible la determinación del óptimo. Dentro de este contexto, y a partir de la forma funcional de la ecuación (67), se observa que:

1. Debido a que las alturas de la superficie muestreada son números reales, su elec-

ción como parámetros a optimizar, hace que la función de costo sea multidimensional y además real. Es decir, está definida en \mathbf{R}^k .

2. En principio, no es posible determinar analíticamente las derivadas de la función de costo y hacerlo en forma numérica no es un proceso directo.
3. La utilización de datos de intensidad introduce una multimodalidad en la función. Es decir, es posible la existencia de varias soluciones (aproximadas) del problema [Walther, 1963; O'Neill y Walther, 1963].

Dentro del conjunto de los algoritmos de optimización multidimensionales se encuentran tres que podrían satisfacer las restricciones anteriores y que quizá podrían también utilizarse en la solución del problema inverso estudiado en este trabajo. El primero de ellos, conocido como el Método Simplex modificado [Nelder y Mead, 1965]. Este es un algoritmo clásico de búsqueda directa concebido originalmente para optimización experimental. Por su facilidad de implementación ha sido utilizado en una amplia gama de problemas de optimización. Por otro lado, los otros dos algoritmos son métodos intermedios, que basan sus principios de operación en la imitación de procesos naturales. Una de estas metodologías es el "Método de Recocido Simulado" y la otra cae dentro de los que se han llamado "Algoritmos Evolutivos".

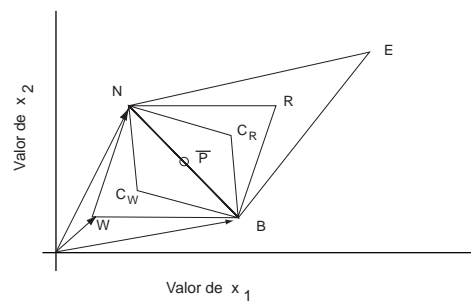
Antes de proceder a la aplicación de los los algoritmos mencionados en la reconstrucción de superficies rugosas, se considera conveniente describir brevemente las bases sobre las cuales cada uno de ellos fundamenta su proceso de búsqueda.

IV.1.1 Método Simplex

Un "simplex" es una figura geométrica que posee un número de vértices igual a la dimensión del espacio en el que se está trabajando más uno. Esto es, si la función cuyo mínimo quiere encontrarse pertenece al espacio R^k , el "simplex" correspondiente estará definido por $k + 1$ puntos en el espacio. Por ejemplo, para una función $f(x_1)$ el "simplex" correspondiente será una línea recta. Si la función es $f(x_1, x_2)$, el "simplex" será un triángulo (no necesariamente equilátero). Si la función fuese $f(x_1, x_2, x_3)$ el "simplex" sería un tetraedro, mientras que para una función con cuatro o más variables la visualización del "simplex" resulta más complicada.

Con el fin de ilustrar las operaciones que se llevan a cabo para encontrar el mínimo de una función mediante el uso del método simplex, sin por esto perder generalidad, se considera una función de dos variables. Con esto, se busca facilitar la exposición a través la interpretación geométrica del método. El "simplex" correspondiente a una función $f(x_1, x_2)$ se ilustra en la figura 25.

En este método el primer paso consiste en evaluar la función en tantos puntos como vértices tenga el "simplex", asociándole a cada uno de éstos el valor correspondiente. Es importante mencionar que los valores iniciales de las variables x_i son elegidos arbitrariamente. Una vez que se tiene el "simplex" inicial, se identifica entonces el mejor valor del vértice (f_B), el peor valor (f_W) y el más cercano al peor (f_N). Estos serán los que determinen tanto la dirección de la búsqueda como la operación a realizar para que el "simplex" se dirija hacia el óptimo.

OPERACIONES DEL SIMPLEX PARA UNA FUNCIÓN $f(x_1, x_2)$ 

- a) Reflexión $\bar{R} = (\bar{P} - \bar{W}) + \bar{P}$
- b) Expansión : Si $f(R) < f(B)$ entonces $\bar{E} = 2(\bar{P} - \bar{W}) + \bar{P}$
- c) Contracción de la Reflexión: Si $f(R) > f(B)$ y $f(R) < f(N)$ $\bar{C}_R = \frac{1}{2}(\bar{P} - \bar{W}) + \bar{P}$
- d) Contracción del peor valor: Si $f(R) > f(W)$ $\bar{C}_W = -\frac{1}{2}(\bar{P} - \bar{W}) + \bar{P}$

Donde P es el vector de posición del centroide del segmento de recta \overline{NB} y tiene forma:

$$\bar{P} = \frac{1}{2}(\bar{B} + \bar{N})$$

Figura 25: Ilustración de las operaciones del simplex para una función de dos variables.

La primera operación es la reflexión y tiene como objetivo eliminar al peor vértice del simplex mediante la determinación de un nuevo vértice localizado en la dirección opuesta al peor. Para un simplex cuya función de costo está dada por $f(x_1, x_2)$ se definen los vectores \mathbf{B} , \mathbf{N} y \mathbf{W} como los vectores de posición de los puntos correspondientes al mejor, al más cercano al peor y al peor valor de la función de costo, respectivamente (ver figura 25). Para una cuenca de tipo parabólico, es razonable suponer que la reflexión del peor valor con respecto al segmento de recta \overline{BN} representaría una mejora en la búsqueda del óptimo. La reflexión está representada por:

$$\mathbf{R} = 2(\mathbf{P} - \mathbf{W}) + \mathbf{W} = (\mathbf{P} - \mathbf{W}) + \mathbf{P}, \quad (151)$$

donde $\mathbf{P} = \frac{(\mathbf{B} + \mathbf{N})}{2}$ y es el vector de posición del centroide del segmento de recta \overline{BN} .

Una vez que se conoce el valor de la función en el punto \mathbf{R} , pueden presentarse las siguientes situaciones:

1. $f(\mathbf{R}) > f(\mathbf{B})$ y $f(\mathbf{R}) < f(\mathbf{N})$

En este caso, el nuevo simplex estará formado por los vértices \mathbf{B} , \mathbf{R} y \mathbf{N} y la operación a realizar sería nuevamente una reflexión, tal y como se describió anteriormente. Es importante destacar que el algoritmo original, propuesto en 1962 por [Spendley *et al.*, 1962], únicamente hacía uso de la reflexión para hallar el valor mínimo de la función. En 1965, [Nelder y Mead, 1965], propusieron un algoritmo similar al de Spendley, pero que incluía las modificaciones que se describen

a continuación.

2. $f(\mathbf{R}) < f(\mathbf{B})$

Si se obtiene un valor de la función mejor que el "óptimo inicial", es posible suponer que el simplex se está desplazando hacia la dirección en donde se encuentra localizado el "óptimo global". Entonces, para acelerar la búsqueda, puede llevarse a cabo una segunda operación conocida como "expansión", cuya interpretación geométrica puede describirse con la ayuda de la figura 25. Matemáticamente, la expansión se puede representar de la forma

$$\mathbf{E} = \mathbf{R} + (\mathbf{P} - \mathbf{W}) = 2(\mathbf{P} - \mathbf{W}) + \mathbf{P} \quad (152)$$

Si el resultado de la expansión es tal que $f(\mathbf{E}) < f(\mathbf{B})$, entonces se tendrá un nuevo simplex (mejor que el original), cuyos vértices serán \mathbf{B} , \mathbf{N} y \mathbf{E} .

3. $f(\mathbf{R}) > f(\mathbf{N})$ y $f(\mathbf{R}) < f(\mathbf{W})$

Si el resultado de la reflexión arroja un valor de la función objetivo muy cercano al peor, es claro que el simplex se está desplazando en una dirección no deseada. Es conveniente entonces llevar a cabo una "contracción de la reflexión" y, con base en el resultado, decidir cual será la siguiente operación a realizar. Haciendo otra vez referencia a la figura 25, la expresión para la "contracción de la reflexión" está dada por

$$\mathbf{C}_R = \frac{1}{2}(\mathbf{P} - \mathbf{W}) + \mathbf{P}, \quad (153)$$

donde el nuevo simplex estará formado por los vértices \mathbf{B} , \mathbf{N} y \mathbf{C}_R .

4. $f(\mathbf{R}) > f(\mathbf{W})$

También en este caso se tiene una reflexión fallida. El simplex no debe desplazarse en esa dirección, por lo que procede una "contracción del peor valor". La expresión correspondiente es:

$$\mathbf{C}_W = \mathbf{P} - \frac{1}{2}(\mathbf{P} - \mathbf{W}). \quad (154)$$

El nuevo simplex estará conformado por los vértices \mathbf{B} , \mathbf{N} y \mathbf{C}_W (Ver figura 25).

Finalmente, una vez que el simplex se encuentra dentro de una región cercana al óptimo, la operación necesaria para la determinación de éste es la contracción múltiple. Básicamente, y como su nombre lo indica, el simplex se contrae hasta convertirse en un punto. Cuando esto sucede, y una vez que se ha satisfecho un criterio, ya sea de convergencia o de terminación, la búsqueda cesa y todos los elementos de la población inicial han convergido a un óptimo.

El simplex se considera un método determinístico, ya que la aplicación en forma iterativa de las operaciones descritas anteriormente se lleva a cabo mediante reglas fijas, que no involucran procesos aleatorios, y cuyo fin es conducir al simplex a la determinación de un óptimo. A su vez, durante la búsqueda, el simplex se "adapta" a la topología de la función objetivo que se desea optimizar. Sin embargo, una clara desventaja de este método reside en el hecho de utilizar información local para la determinación del óptimo, ya que esto puede conducir a la convergencia prematura a un mínimo local.

IV.1.2 Método de Recocido Simulado

El recocido simulado (simulated annealing), al igual que el "simplex", es un método de optimización de funciones multidimensionales que durante la búsqueda usa únicamente valores de la función objetivo y no de las derivadas de ésta. Sin embargo, una característica de este método es la utilización de procesos aleatorios, los cuales tienen una dirección controlada por la lógica del algoritmo. Por este hecho, el recocido simulado es considerado un método estocástico.

El origen de este método se remonta a 1953 cuando [Metropolis *et al.*, 1953] propusieron un algoritmo para calcular los estados de un cristal que se enfriaba lentamente. Sin embargo, no fué sino hasta 1984 que [Kirkpatrick *et al.*, 1983] retomaron la idea de Metropolis y la aplicaron para la solución de un problema de optimización relacionado con el diseño de circuitos electrónicos.

La idea central del método de recocido simulado reside en la imitación del proceso con el que un material alcanza un estado de mínima energía al reducir su temperatura lentamente. En esencia, el recocido simulado es relativamente simple; la secuencia de pasos a seguir para la determinación del óptimo se enlista a continuación:

1. Se evalúa la función de costo $f_1(\mathbf{x}_1)$ en un punto \mathbf{x}_1 , que representa el vector de parámetros que caracterizan al problema.
2. Se introduce una variación aleatoria $\Delta\mathbf{x}$ tal que $\mathbf{x}_2 = \mathbf{x}_1 + \Delta\mathbf{x}$.
3. Se evalúa $f_1(\mathbf{x}_2)$

4. Se calcula $\Delta f = f_1(\mathbf{x}_2) - f_1(\mathbf{x}_1)$. Dos situaciones pueden tomar lugar a partir del cálculo de Δf . La primera sería que $\Delta f \leq 0$. En este caso, $f_1(\mathbf{x}_2) \leq f_1(\mathbf{x}_1)$ y la búsqueda del mínimo se está llevando a cabo en la dirección correcta. La nueva configuración \mathbf{x}_2 debe ser aceptada.

El segundo resultado posible es que $\Delta f > 0$. Si se acepta o no la nueva configuración es una decisión que se toma con base en una ley de probabilidades. Como se acostumbra en el método, se elige

$$P(\Delta f) = \exp \left\{ \frac{-\Delta f}{T} \right\}, \quad (155)$$

donde T es un parámetro de control (temperatura) cuyo valor decrece al aumentar el número de iteraciones. Si se cumple que

$$\epsilon_1 < P(\Delta f), \quad (156)$$

donde ϵ_1 es un número aleatorio uniformemente distribuido en el intervalo $[0,1]$, la nueva configuración \mathbf{x}_2 es aceptada y una nueva iteración del proceso se lleva a cabo. Por otro lado, si se tiene que

$$\epsilon_1 > P(\Delta f) \quad (157)$$

entonces, la configuración \mathbf{x}_2 será rechazada y la configuración anterior \mathbf{x}_1 será

utilizada para la siguiente iteración del proceso.

El decremento de la variable de control T simula el enfriamiento gradual de un sistema. El hecho de tener inicialmente valores grandes de T permite al algoritmo explorar una región grande del espacio de búsqueda, para posteriormente restringirse a una región pequeña. La inclusión de valores elevados de la función impide que el algoritmo se estanque en un óptimo local, lo cual resulta particularmente importante cuando se trabaja con funciones cuya topología es complicada.

IV.1.3 Algoritmos Evolutivos

El término *Algoritmos Evolutivos* es el nombre genérico utilizado para referirse a un grupo específico de métodos computacionales inspirados en los procesos de evolución que ocurren en la naturaleza y que han sido empleados para la solución de problemas complejos en los que la aplicación de métodos clásicos no es posible, o es limitada. Estas técnicas se subdividen en tres ramas principales. La primera, que surge en sus inicios como un modelo general de procesos adaptables, está formada por los "Algoritmos Genéticos" [Holland, 1992; de Jong, 1975], los cuales han sido aplicados extensivamente en las últimas décadas en la solución de problemas de optimización. La segunda rama se conoce como "Estrategias de Evolución" [Schwefel, 1995], inicialmente concebidas como un método de optimización experimental cuya aplicación se extendió, posteriormente, a la solución de problemas representados por funciones con parámetros reales. Finalmente, la tercera rama, la Programación Evolutiva [Fogel, 1962], surge como un

intento para la creación de inteligencia artificial, aunque también a sido empleada en la solución de problemas de optimización [Kim y Myung, 1997]. Es interesante mencionar que el surgimiento de las tres corrientes se da prácticamente al mismo tiempo, en forma independiente, a principios de los años 60. Sin embargo, no fué sino hacia fines de la década de los 70 cuando empezó la comunicación y el intercambio de ideas entre estos grupos de investigación.

El diseño de un algoritmo evolutivo se fundamenta en las características particulares de la función objetivo cuya solución quiere encontrarse. Sin embargo, a pesar de sus diferencias, todos los algoritmos evolutivos son métodos estocásticos basados en procesos heurísticos de búsqueda poblacional, que incorporan elementos aleatorios de variación y de selección [Salomon, 1997]. La premisa de la que parten es que la supervivencia de un individuo dentro de una población dependerá, en general, de su aptitud para adaptarse al medio.

La estructura de un algoritmo evolutivo canónico se muestra en la figura 26. El proceso de búsqueda empieza con la generación de una población inicial formada por elementos que a su vez son conjuntos de variables objeto. Con el fin de seleccionar únicamente a los elementos con el potencial para generar soluciones cada vez más cercanas a la óptima, se evalúa a cada uno de ellos y se le asocia un valor de aptitud que definirá, mediante el uso de un criterio que puede ser determinístico o estocástico, si son incluidos en el ciclo principal del algoritmo o si, por el contrario, deben ser desechados.

Los cambios en la población inicial tienen lugar mediante el uso de operadores de

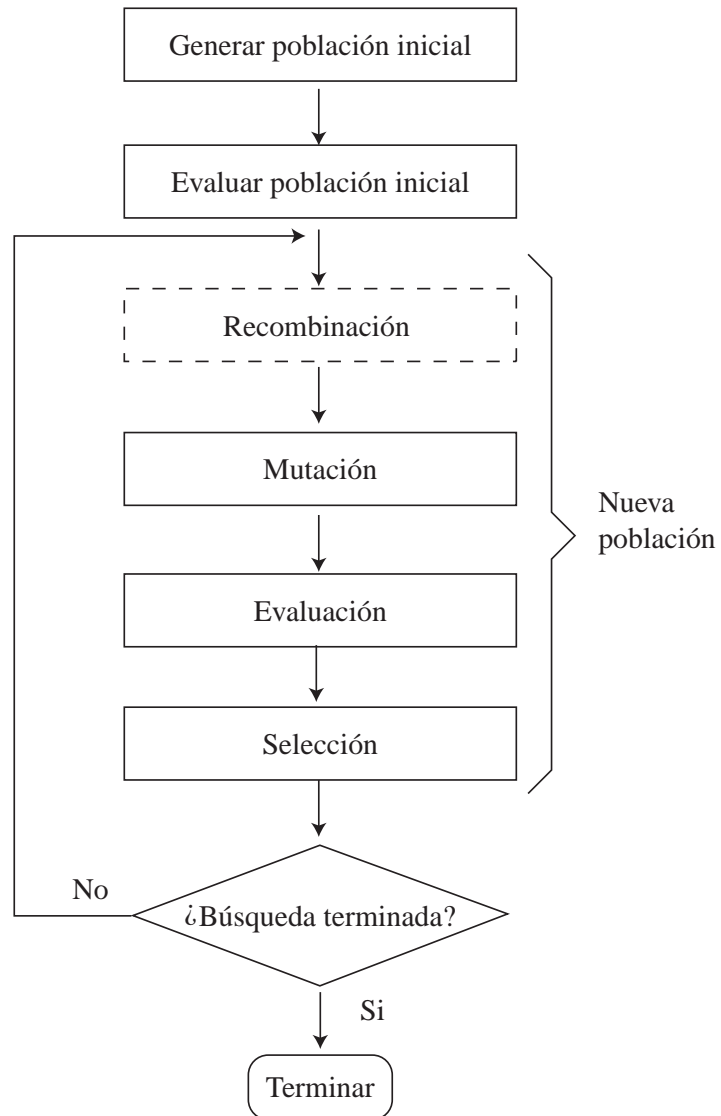


Figura 26: Diagrama de bloques de Algoritmo Evolutivo Canónico.

variación tales como la recombinación y la mutación. El primero tiene la función de *explotar* el espacio de búsqueda a través de la mezcla entre los elementos más aptos. La mutación, por otro lado, *explora* el espacio de búsqueda al generar, por medio de la introducción de variaciones aleatorias en los elementos de la población, nuevas posibles soluciones.

En cada iteración (generación) del algoritmo se genera una nueva población a partir de la cual se seleccionan los mejores elementos. Éstos, a su vez, formarán la población inicial para la siguiente generación. En principio, se espera que la repetición iterativa de estos pasos conduzca al óptimo de la función. Por último, el ciclo principal del algoritmo evolutivo finaliza cuando se ha alcanzado un criterio de terminación previamente establecido. Éste puede ser el número de iteraciones, la medida del error entre la solución encontrada y la óptima, la convergencia prematura a un óptimo local, etc.

En los inicios de la computación evolutiva existían diferencias importantes establecidas debido a las características propias de cada una de sus ramas. Por ejemplo, los Algoritmos Genéticos empleaban alfabetos binarios $\{0, 1\}$ para la representación de las variables objeto, mientras que las Estrategias Evolutivas utilizaban parámetros reales (aunque en [Beyer, 2001] se cita una referencia en la que se utilizan estrategias de evolución con alfabeto binario). Otra diferencia era el énfasis que cada uno de estos algoritmos ponía en los operadores genéticos. Es decir, los Algoritmos Genéticos daban prioridad al uso de la recombinación y de la selección, mientras que la mutación era únicamente utilizada para introducir cambios aleatorios en las variables objeto a partir del establecimiento de una probabilidad de ocurrencia baja. Por el contrario,

las Estrategias Evolutivas Canónicas solamente utilizaban la mutación y la selección. Sin embargo, en las últimas décadas, la brecha existente se ha ido angostando y las diferencias son cada vez menos significativas. Es posible encontrar variantes de Algoritmos Genéticos que utilizan una representación real y mutación, así como también es posible hallar Estrategias Evolutivas que utilizan recombinación, una codificación binaria, o ambas simultáneamente.

No existe, como tal, una pauta a seguir en la elección del algoritmo más conveniente para la solución de un problema específico. Sin embargo, una regla empírica frecuentemente usada es que *el principio para el desarrollo de una heurística evolutiva se fundamenta en la representación de los parámetros a optimizar* [Bäck et al., 1996]. Con base en esta regla, y a partir de que la función de costo (67) se representa con parámetros reales, es posible pensar que las Estrategias de Evolución pueden ser la metodología más adecuada para la solución del problema estudiado en este trabajo. Por lo tanto, las secciones siguientes se enfocarán sobre los aspectos principales de esta clase de algoritmos evolutivos.

Estrategia de Evolución $(\mu/\rho_R, \lambda)$

Una característica propia de las estrategias de evolución es que operan sobre poblaciones de elementos que pertenecen al espacio de soluciones. Estos conjuntos poseen un tamaño definido que permanece constante a lo largo del proceso de búsqueda y que se denotan como μ para la población inicial y como λ para la población secundaria [Schwefel, 1995].

Cada elemento de la población tendrá la forma

$$\mathbf{a}(\mathbf{x}, \mathbf{s}, f(\mathbf{x})), \quad (158)$$

donde \mathbf{x} es un vector n -dimensional conocido como *vector de parámetros (o variables) objeto* cuyas componentes $\mathbf{x} = \{x_1, x_2, x_3, \dots, x_N\}^T$ pertenecen al *espacio de variables objeto*, que a su vez depende del problema particular que se esté resolviendo. El valor de aptitud se obtiene al evaluar la función de costo $f(\mathbf{x})$ en cada uno de los vectores objeto \mathbf{x} . El vector \mathbf{s} representa un conjunto de parámetros endógenos conocidos como *parámetros de estrategia*, cuya función se explicará en los siguientes párrafos.

La población inicial \mathfrak{B}_μ , para una iteración g dada, está dada por

$$\mathfrak{B}_\mu^{(g)} = \left\{ \mathbf{a}_1^{(g)}, \mathbf{a}_2^{(g)}, \dots, \mathbf{a}_\mu^{(g)} \right\}, \quad (159)$$

a su vez, la población secundaria puede representarse como

$$\tilde{\mathfrak{B}}_\lambda^{(g)} = \left\{ \tilde{\mathbf{a}}_1^{(g)}, \tilde{\mathbf{a}}_2^{(g)}, \dots, \tilde{\mathbf{a}}_\lambda^{(g)} \right\} \quad (160)$$

donde la tilde sobre cada uno de los elementos del conjunto indica que éstos han sido generados por medio de la aplicación de los operadores genéticos en los elementos del conjunto $\mathfrak{B}_\mu^{(g)}$ de la ecuación (159).

Una estrategia de evolución genérica, que a su vez engloba cualquier variante que

exista, puede escribirse en forma conceptual como

$$(\mu/\rho_R, ^+\lambda), \quad (161)$$

donde ρ es un parámetro exógeno, definido dentro del intervalo $1 \leq \rho \leq \mu$, que denota el número de elementos de la población inicial $\mathfrak{B}_\mu^{(g)}$ que se escogieran, por medio de un operador de *reproducción*, para recombinarse y generar un nuevo elemento de la población secundaria

$$\mathfrak{E} = \{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_\rho\}, \quad (162)$$

sobre la que posteriormente se aplicarán los operadores de variación. El subíndice R en la expresión (161) indica el esquema de recombinación que se está utilizando. Es decir, si la recombinación es dominante o intermedia. A su vez, existen dos esquemas de selección posibles. Al primero se le asocia el símbolo ”+” y se conoce como ”Estrategia Elitista,” mientras que a la ”Estrategia no-elitista” le corresponde el símbolo ”,”. Una explicación detallada sobre estos conceptos se da en los apartados correspondiente a los operadores genéticos mencionados.

La estrategia de evolución expresada por (161) puede escribirse como un pseudocódigo de la siguiente forma:

Algoritmo IV.1.1

```

1 begin
2    $g := 0;$ 

```

```

3  inicializa ( $\mathfrak{B}_\mu^{(0)} := \left\{ \left( \mathbf{x}_m^{(0)}, \mathbf{s}_m^{(0)}, f(\mathbf{x}_m^{(0)}) \right) \right\}$ );
4  while criterio de terminación no alcanzado
5      for  $l := 1$  to  $\lambda$  do
6           $\mathfrak{E}_\mu^{(0)} := \text{reproducción} \left( \mathfrak{B}_\mu^{(g)}, \rho \right)$ ;
7           $\mathbf{s}_l := s\_recombinación(\mathfrak{E}_l, \rho, R)$ ;
8           $\tilde{\mathbf{s}}_l := s\_mutación(\mathbf{s}_l)$ ;
9           $\mathbf{x}_l := x\_recombinación(\mathfrak{E}_l, \rho, R)$ ;
10          $\tilde{\mathbf{x}}_l := x\_mutación(\mathbf{x}_l, \tilde{\mathbf{s}}_l)$ ;
11          $\tilde{f}_l = f(\tilde{\mathbf{x}}_l)$ ;
12     end;
13      $\mathfrak{B}_\lambda^{(g)} := \left\{ \left( \tilde{\mathbf{x}}_l, \tilde{\mathbf{s}}_l, \tilde{f}_l \right) \right\}$ 
14     if esquema =  $(\mu, \lambda)$ 
15          $\mathfrak{B}_\mu^{(g+1)} := \text{selección}_{f_\mu} \left\{ \mathfrak{B}_\lambda^{(g)} \right\}$ ;
16     else esquema =  $(\mu + \lambda)$ 
17          $\mathfrak{B}_\mu^{(g+1)} := \text{selección}_{f_\mu} \left\{ \mathfrak{B}_\lambda^{(g)}, \mathfrak{B}_\mu^{(g)} \right\}$ ;
18     end
19      $g := g + 1$ ;
20 do
21 end

```

Selección

El operador selección $(+)$ genera, por medio de un proceso determinista, el conjunto \mathfrak{B}_μ que servirá como población inicial en la siguiente iteración $g + 1$ del algoritmo. Este operador, a diferencia del resto, actúa directamente sobre los valores de aptitud de los elementos de la población y no sobre los parámetros objeto. Es decir, únicamente aquellos elementos con los mejores valores de aptitud serán elegidos para la siguiente iteración.

La diferencia principal entre los esquemas $(\mu + \lambda)$ y (μ, λ) puede visualizarse a partir de la línea 15 del algoritmo (IV.1.1). Al emplear el esquema (μ, λ) únicamente serán seleccionados los elementos pertenecientes a la población secundaria $\tilde{\mathfrak{B}}_\lambda^{(g)}$ que resulta de la aplicación de los operadores de reproducción, de recombinación y de mutación. Una consecuencia importante del uso de este esquema es que el mejor elemento de la población $\mathfrak{B}_\mu^{(g+1)}$ puede ser peor que el mejor elemento de la población $\mathfrak{B}_\mu^{(g)}$. Aún a pesar de que este deterioro temporal en la aptitud podría hacer que el algoritmo diverja, esto permitiría evitar regiones de atracción hacia un óptimo local que conduciría al algoritmo a una convergencia prematura [Bäck *et al.*, 1997]. La característica anterior hace que el esquema (μ, λ) sea *no elitista* y de ahí el origen del nombre alternativo con el que también se le conoce.

La esencia del esquema $(\mu + \lambda)$ o *elitista* se muestra en la línea 17 del algoritmo (IV.1.1). La población $\mathfrak{B}_\mu^{(g+1)}$ estará conformada únicamente por los mejores elementos seleccionados de la unión entre la población inicial $\mathfrak{B}_\mu^{(g)}$ y la población secundaria

$\tilde{\mathfrak{B}}_\lambda^{(g)}$. A diferencia del esquema *no elitista*, en este caso no existe posibilidad de un deterioro temporal en la aptitud de los elementos. Este hecho garantiza un decremento monotónico (para el caso de minimización) de la función de costo. Sin embargo, este esquema no evita que el algoritmo escape de la vecindad de atracción de un óptimo local y, por lo tanto, es proclive a la convergencia prematura.

Mutación

En las Estrategias Evolutivas la mutación es el operador principal para llevar a cabo la exploración del espacio de búsqueda y, junto con la selección, forma la base sobre la cual se fundamentan las de estrategias de evolución.

La mutación consiste en la introducción de una variación aleatoria en las componentes del vector de variables objeto \mathbf{x} , al sumarle a cada uno de éstos números aleatorios normalmente distribuidos. Siguiendo la notación utilizada en la línea 10 del algoritmo (IV.1.1), se tiene

$$\tilde{\mathbf{x}}_l = \mathbf{x}_l + \mathbf{z} \quad (163)$$

donde \mathbf{z} representa el vector aleatorio. En su forma más general el vector de mutación está dado por

$$\mathbf{z} = \{N_1(0, \sigma_1), N_2(0, \sigma_2), N_3(0, \sigma_3), \dots, N_N(0, \sigma_N)\}^T, \quad (164)$$

donde, $N_i(0, \sigma_i)$ para $i = 1, \dots, N$ son números aleatorios gaussianos, estadísticamente independientes, con media cero y con una desviación estándar individual σ_i . En el

caso más simple se supone $\sigma = \sigma_i \forall i = 1, \dots, N$, con lo que la ecuación (163) puede reescribirse como

$$\tilde{\mathbf{x}}_l = \mathbf{x}_l + \sigma \mathbf{z}. \quad (165)$$

A partir de la expresión anterior puede verse que el parámetro σ influirá directamente en la distancia a la que se encuentre una solución mutada con respecto al óptimo que se busca.

Es posible controlar las variaciones que experimenta el parámetro σ durante la búsqueda por medio de la aplicación de los operadores genéticos. A este proceso se le conoce como *autoadaptación* y se encuentra incluido en las líneas (7) y (8), dentro del ciclo principal del algoritmo (IV.1.1).

La implementación práctica del principio de autoadaptación requiere de la definición de los parámetros exógenos τ' , τ y τ_0 , conocidos como índices de aprendizaje, que a su vez se relacionan con el número de variables objeto N por medio de las expresiones $\tau' \propto (\sqrt{2N})^{-1}$, $\tau \propto (\sqrt{2\sqrt{N}})^{-1}$ y $\tau_0 \propto (\sqrt{N})^{-1}$. Debe señalarse que el establecimiento de estos índices depende totalmente del problema que se esté resolviendo [Bäck y Schwefel, 1996; Bäck *et al.*, 1997].

Reproducción

La función del operador de reproducción es la generación del conjunto \mathfrak{E} , al que posteriormente se le aplicarán los operadores genéticos. El operador de reproducción se ilustra en la línea 6 del algoritmo (IV.1.1) y básicamente consiste de una selección

aleatoria, utilizando números uniformemente distribuidos en el intervalo $[1, \mu]$, de los ρ elementos que serán recombinados para formar la población intermedia \mathfrak{E} .

Recombinación

El objetivo principal de la recombinación es *explotar* el espacio de búsqueda mediante el intercambio de información entre aquellos elementos de la población cuya aptitud sea alta. A través de esta operación se espera, en principio, que las características de los elementos más aptos se transfieran de una generación a otra hasta alcanzar el criterio de terminación.

En la ecuación (161) y en las líneas (7) y (9) del algoritmo (IV.1.1), el operador de recombinación aparece denotado con la letra R . En las estrategias de evolución se utilizan dos esquemas principales, conocidos como recombinación intermedia y recombinación dominante [Beyer, 2001].

La recombinación intermedia se define como

$$\langle \mathbf{x}^{(g)} \rangle_{\rho} = \frac{1}{\rho} \sum_{i=1}^{\rho} \mathbf{x}_i^{(g)}, \quad (166)$$

donde ρ es el número de elementos pertenecientes al conjunto \mathfrak{E} que serán recombinados. Para el caso en el que $\rho > 2$ se tiene una recombinación múltiple. Cuando $\rho = \mu$ todos los miembros de la población inicial se recombinan entre sí y el vector resultante puede interpretarse como el centro de masa de la población.

Por otro lado, para describir el concepto de recombinación dominante el primer

paso es suponer que el vector de variables objeto \mathbf{x} puede expresarse en términos de una base ortonormal $\{\mathbf{e}_j\}$ tal que

$$\mathbf{x} = \sum_{j=1}^N x_j \mathbf{e}_j. \quad (167)$$

Entonces, la recombinación dominante puede definirse como

$$\mathbf{x}^{(g)} = \sum_{j=1}^N (\mathbf{e}_j^T \cdot \mathbf{x}_i) \mathbf{e}_j, \quad \forall \quad 1 \leq i \leq \rho. \quad (168)$$

Donde i es un número entero seleccionado aleatoriamente en el intervalo $[1, \rho]$ y ρ representa el número de elementos que se recombinarán. En la ecuación (168), el término dentro del paréntesis es la j -ésima componente del elemento recombinado en la dirección \mathbf{e}_j . Su módulo será igual al de la j -ésima componente en la misma dirección, pero de un elemento i del conjunto \mathfrak{E} seleccionado en forma aleatoria.

La recombinación dominante es también conocida como *recombinación discreta* [Beyer, 2001] y para el caso en el que $\rho = \mu$ se tiene el esquema de *recombinación discreta global* [Schwefel, 1995].

IV.2 Recuperación de perfiles utilizando la representación espectral

Debido a la facilidad de su implementación, el punto de partida para la resolución del problema con datos de intensidad se basó en la generación de la superficie con el método espectral, descrito en la sección II.3.1. La suposición inicial fue que la superficie a reconstruir pertenece a una clase estadística conocida, particularmente se supuso que el perfil era una realización de un proceso aleatorio gaussiano cuya longitud de correlación y desviación estándar de alturas también eran conocidas. Con este conocimiento *a priori*, el objeto era encontrar la realización particular que dió origen a los datos de esparcimiento.

En un primer intento de resolver el problema inverso se utilizaron los métodos simplex y recocido simulado, descritos en las secciones IV.1.1 y IV.1.2 respectivamente. Dentro del contexto del problema que se estudia, el tamaño del simplex estará determinado por el número de puntos de muestreo sobre la superficie a reconstruir más uno. Es decir, esta cantidad depende directamente de cuántos parámetros deban optimizarse. Aunque no se conocen restricciones en el número máximo de vértices que un simplex pueda tener, o por lo menos no se encontraron reportes al respecto en las referencias consultadas, inicialmente se propuso una superficie cuyo tamaño no provocara inestabilidades en el proceso de búsqueda del algoritmo y que fuera manejable con los recursos computacionales con los que se contaba.

El perfil a reconstruir se muestra en la figura 27. Consiste en una realización de

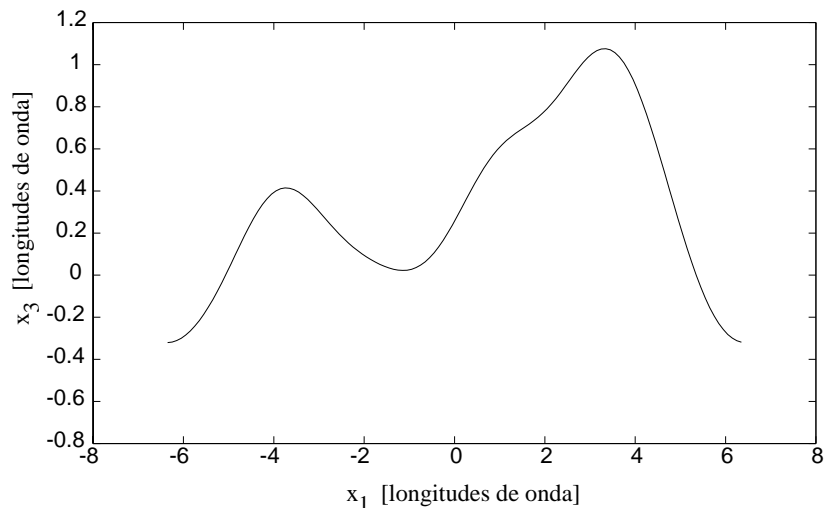


Figura 27: Perfil original a reconstruir.

un proceso aleatorio gaussiano, con función de correlación gaussiana, con desviación estándar de alturas $\delta = \frac{\lambda}{2}$ y longitud de correlación $a = 2\lambda$. Contiene un número total $N = 128$ puntos de muestreo, a intervalos de $\Delta x = \frac{\lambda}{10}$. Es necesario enfatizar que la letra λ ha venido utilizándose en forma indistinta para denotar a la longitud de onda del campo, notación usual en óptica, y también, siguiendo la notación propuesta por Schwefel [Schwefel, 1995], al tamaño de la población secundaria generada por medio de los operadores genéticos. Debido a que los contextos en los que dichas variables se emplean son distintos, no debe existir confusión alguna.

En principio, los datos que sirven de entrada para el algoritmo de inversión podrían obtenerse experimentalmente. Sin embargo, con el fin de estudiar y optimizar los algoritmos propuestos, en este trabajo se utilizan datos de esparcimiento generados a través de la solución numérica del problema de esparcimiento directo utilizando el método riguroso descrito por [Maradudin *et al.*, 1990]. Los datos de intensidad es-

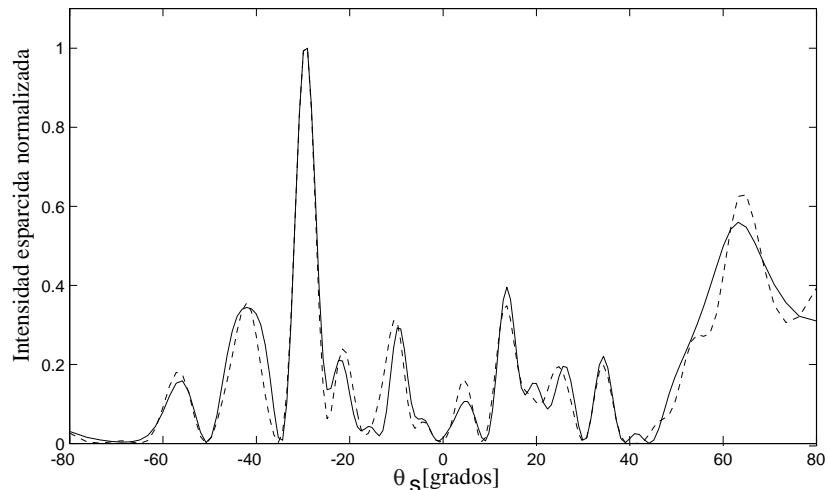


Figura 28: Patrón de esparcimiento generado por la superficie de la figura 27, al ser iluminada con una onda plana a incidencia normal. La línea punteada corresponde al patrón de esparcimiento generado, bajo las mismas condiciones de iluminación, por el perfil dibujado con círculos en la figura 34(a), cuya descripción se encuentra en la sección IV.2.3.

parcida se generaron iluminando el perfil de la figura 27, secuencialmente, con cuatro direcciones diferentes, definidas por los ángulos de incidencia $\theta_1 = -60^\circ$, $\theta_2 = -30^\circ$, $\theta_3 = 0^\circ$ y $\theta_4 = 40^\circ$. En la figura 28 se muestra, dibujado con una línea punteada, el patrón de esparcimiento obtenido al iluminar con una onda plana a incidencia normal, el perfil de la figura 27.

IV.2.1 Método Simplex

En los estudios iniciales con el método simplex, con el fin de entender su proceso de búsqueda y verificar que no hubiese errores de programación, se incluyó el perfil original dentro de la población inicial. El resto de los elementos de la población se generó aleatoriamente con el método espectral. Los resultados obtenidos se muestran en la figura 29. Se observa que el grado de similitud entre las reconstrucciones y el perfil

original depende del número de iteraciones del algoritmo. Esto, a su vez, tiene como consecuencia directa el incremento del tiempo de máquina requerido para completar la búsqueda, aunque esto último no representa una limitante significativa. Con los parámetros empleados, el algoritmo tarda aproximadamente 10 minutos en llevar a cabo 500 iteraciones en una computadora personal con procesador Pentium III a 640 MHz y con 64 Mb de memoria en RAM.

A pesar de que el método generaba los resultados esperados en las simulaciones que incluían el perfil original, cuando éste no se hallaba dentro de la población inicial el simplex experimentaba una convergencia prematura en un óptimo local que, en la mayoría de los experimentos numéricos realizados, no guarda ninguna semejanza con la superficie que se buscaba reconstruir. Este comportamiento es característico de los métodos de búsqueda local. A pesar de "adaptarse" a la topología de la función de costo, si existe un óptimo local, tenderán a estancarse y quedar atrapados en la vecindad de éste.

IV.2.2 Recocido Simulado

En la sección IV.1.2 se mencionó que el recocido simulado es un método intermedio que explora el espacio de búsqueda mediante una caminata aleatoria, regida a su vez por una distribución de probabilidad definida. Este comportamiento le permite al algoritmo evitar la convergencia prematura en el caso para el que la función de costo sea multimodal.

Por otro lado, es necesario mencionar que cuando este método fue concebido por sus

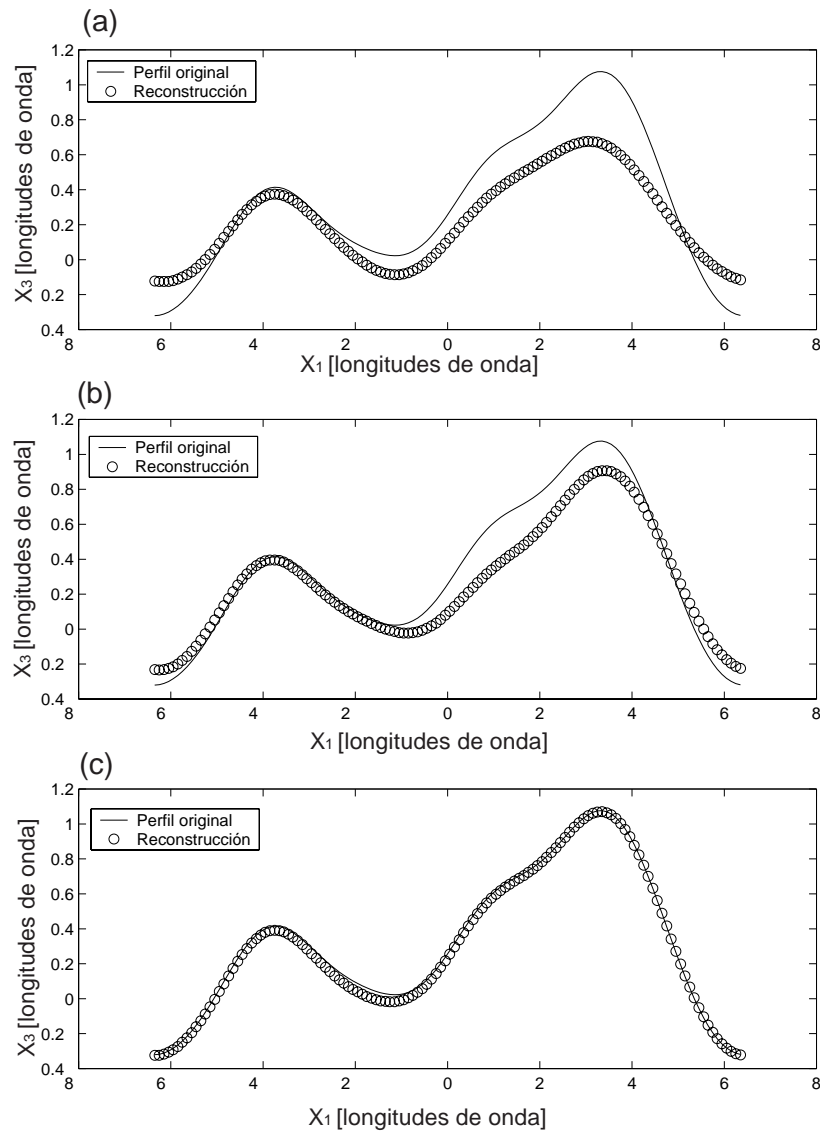


Figura 29: Reconstrucción de perfiles utilizando el método simplex incluyendo el perfil original en la población inicial. Reconstrucciones obtenidas después de (a) 300 iteraciones, (b) 500 iteraciones y (c) 1000 iteraciones.

creadores [Metropolis *et al.*, 1953; Kirkpatrick *et al.*, 1983] no hacía uso de poblaciones para llevar a cabo la búsqueda del óptimo. De hecho, puede interpretarse como una estrategia de evolución (1+1) con presión de selección variable en el tiempo y con tasa de mutación constante [Beyer, 2001]. Sin embargo, la implementación que se utiliza en este trabajo, descrita en [Press *et al.*, 1992], es un algoritmo híbrido que hace uso del simplex y del recocido simulado, y por lo tanto sí utiliza principios poblacionales para llevar a cabo la búsqueda del óptimo. Es decir, en cada iteración del recocido simulado un simplex, cuyos vértices están asociados con las variables objeto, explora el espacio de soluciones. De esta forma, el proceso de búsqueda global del recocido simulado se alterna con el de búsqueda local del simplex. La variación del parámetro de control permite inicializar al simplex en diferentes regiones del espacio de búsqueda y al mismo tiempo evita su estancamiento.

Las simulaciones con el método de recocido simulado se realizaron bajo las mismas condiciones que en el caso del método simplex, incluyendo el perfil original. La tasa de decremento del parámetro de control se fijó en 20% para cada iteración y las simulaciones se llevaron a cabo considerando tres diferentes temperaturas iniciales, definidas como $T_1 = 0$, $T_2 = 50$ y $T_3 = 100$.

Los resultados obtenidos con este algoritmo se muestran en la figura 30. La gráfica 30(a) muestra la gran similitud entre la reconstrucción y el perfil original. El desplazamiento vertical es entendible, ya que esto solo altera la fase del campo esparcido y además solamente se está utilizando información de intensidad. Para el caso $T = 0$, el algoritmo utilizado es equivalente al del simplex. La región del espacio objeto en donde

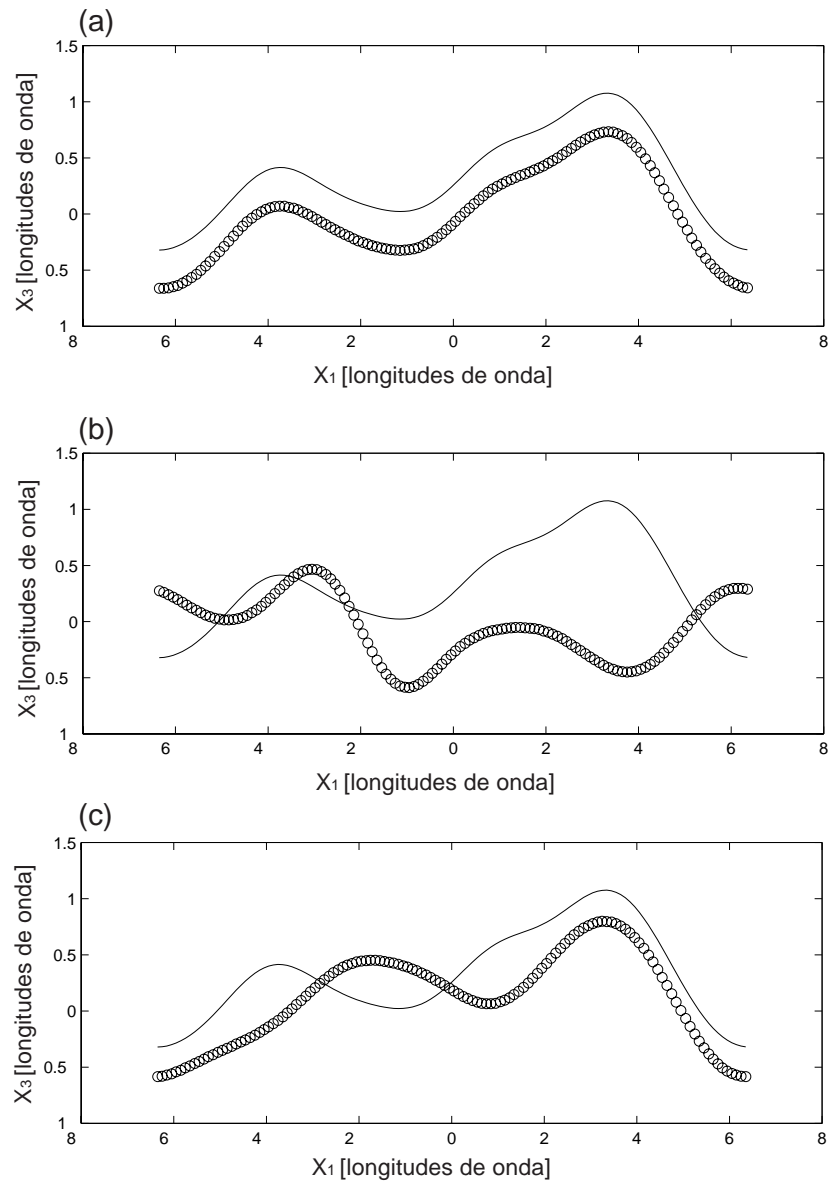


Figura 30: Reconstrucciones obtenidas con el recocido simulado para a) $T_1 = 0$, b) $T_2 = 50$ y c) $T_3 = 100$. La inicialización fue la misma para las tres corridas.

se inició la búsqueda favorece la convergencia del algoritmo hacia el óptimo local.

En la figura 30(b) se presentan los resultados que se obtienen cuando se incrementa el parámetro de control T . La inicialización fue la misma que para el caso anterior. Sin embargo, la reconstrucción difiere totalmente del perfil original. Aunque la probabilidad de aceptación de una solución depende del valor de T [ecuación (155)], la inclusión de variaciones aleatorias en la búsqueda puede alejar al algoritmo de la vecindad del óptimo. Para cuando la probabilidad de aceptación ha disminuido lo suficiente como para que se dé la convergencia, ésta no será en el óptimo global sino en un extremo local. Similarmente, en la figura 30(c) se observa que algunos segmentos de la reconstrucción asemejan partes del perfil original. Sin embargo, no puede afirmarse que la superficie reconstruida sea la que se buscaba.

A partir de los resultados anteriores puede decirse que el método de Recocido Simulado es menos proclive a la convergencia prematura, gracias al uso de elementos aleatorios durante la búsqueda. Sin embargo, este hecho no garantiza que al final se hallará la solución correcta. Por otro lado, la determinación del valor adecuado para el parámetro T no es directa y depende de las características del problema que se esté resolviendo. Por último, se observó que en las simulaciones realizadas sin incluir el perfil original, no hubo convergencia al óptimo buscado.

Un aspecto importante, que no se había tomado en cuenta al inicio de las simulaciones realizadas, es que las implementaciones del método simplex y del método de recocido simulado basan su proceso de búsqueda en un principio análogo al de recombinación intermedia. Es decir, los elementos de la población se mezclan entre sí, en cada

iteración del algoritmo, para generar un nuevo elemento. Esto introduce un problema, ya que después de la primera iteración, no es posible asegurar que los elementos de la población posean las mismas propiedades estadísticas del perfil buscado.

A partir de los resultados obtenidos con los métodos simplex y de recocido simulado, es clara la necesidad de un algoritmo de optimización global multidimensional que satisfaga, simultáneamente, las restricciones impuestas por las características de la función de costo y también por el esquema de representación.

IV.2.3 Algoritmo de inversión evolutivo sin recombinación

El diagrama de flujo correspondientes a las estrategias de evolución que se estudian en esta parte del trabajo se muestra en la figura 31. El paso inicial para las dos versiones del algoritmo es la generación y evaluación de la población inicial, sobre la que se aplicarán los operadores genéticos. La estrategia *elitista* está representada en la rama izquierda de la figura 31, mientras que la estrategia *no elitista* corresponde a la rama derecha. La secuencia de aplicación de los operadores genéticos es la misma para ambas estrategias. Debe señalarse que en las implementaciones exploradas en esta sección no se utilizó ningún esquema de recombinación y por eso el bloque correspondiente a este operador aparece dibujado con una línea punteada en la figura 31. Sin embargo, es necesario aclarar que sí hubiese sido posible emplear el esquema de recombinación dominante descrito en la sección IV.1.1, el cual satisface las restricciones impuestas por la representación espectral respecto a la pertenencia de los elementos de la población a una clase estadística específica. La diferencia fundamental entre las

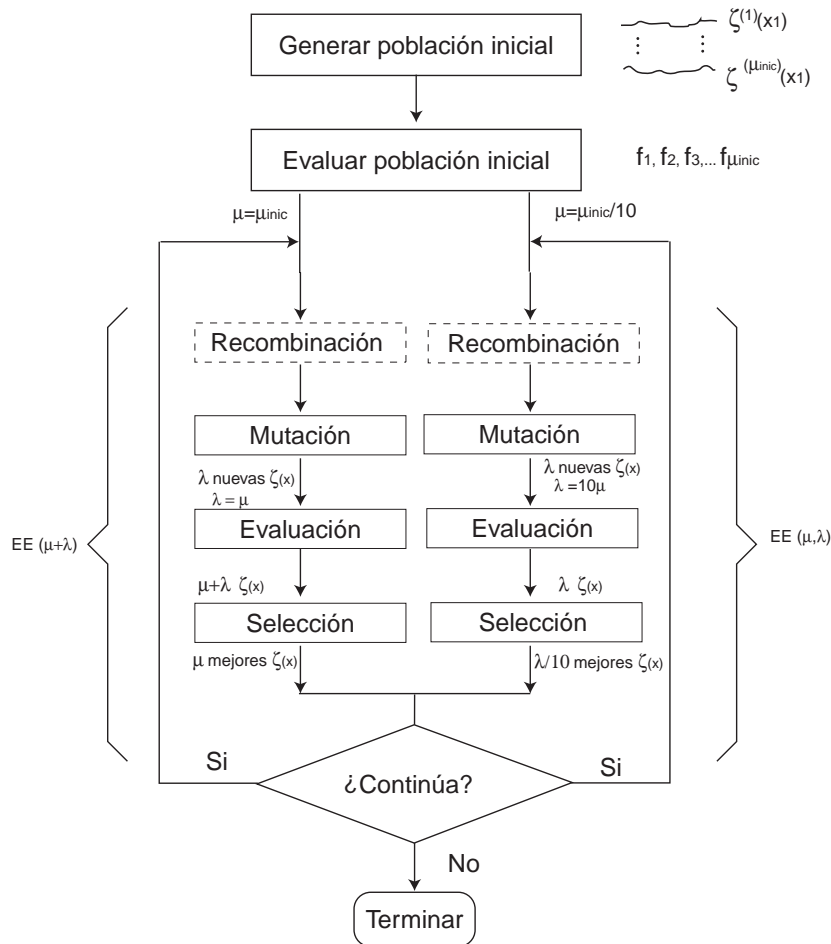


Figura 31: Diagrama de bloques de las estrategias evolución utilizadas.

estrategias propuestas radica en el esquema de selección. A su vez, aunque esto no necesariamente contribuye a la convergencia o la divergencia de los algoritmos, arbitrariamente se definieron las presiones de selección $\frac{\lambda}{\mu} = 1$ y $\frac{\lambda}{\mu} = 10$ para la estrategia *elitista* y para la *no elitista*, respectivamente. El concepto de presión de selección no había sido descrito anteriormente. Básicamente, expresa el número de elementos de la población secundaria, que serán generados por cada elemento de la población inicial.

Es importante mencionar que el esquema de mutación utilizado difiere del expuesto en la sección IV.1.1. La razón de esta diferencia obedece a la restricción impuesta por

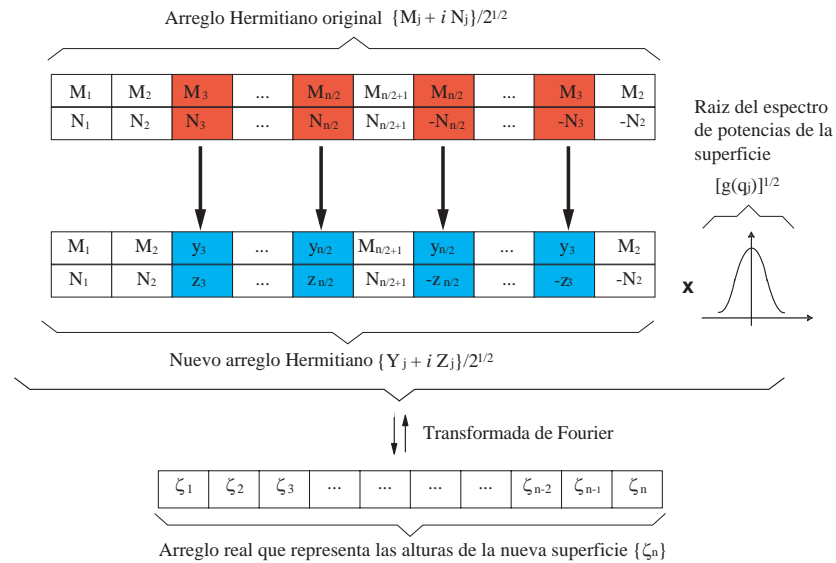


Figura 32: Esquema de mutación utilizado con la representación espectral.

las características propias del método espectral, ya que cambiar las variables objeto en forma independiente una de la otra podría conducir a una variación abrupta de las alturas de la superficie. La principal consecuencia de este hecho sería una situación fuera del interés físico del problema en cuestión.

Para llevar a cabo la mutación se propuso el esquema que se muestra en la figura 32. A partir de ésta y de la ecuación (68), se observa que se han producido cambios aleatorios en la superficie al introducir cambios en el arreglo Hermitiano $\{M_l + iN_l\}$. Siempre que los nuevos números posean las mismas propiedades estadísticas y que la hermiticidad del arreglo se conserve, la nueva superficie pertenecerá a la misma clase estadística definida anteriormente. Al igual que para el simplex y para el recocido simulado, el número total de puntos muestreados en la superficie fue $N = 128$. Después de un proceso de prueba y error se llegó a que el número de elementos a cambiar en el arreglo Hermitiano, para producir una mutación adecuada, debe ser aproximadamente

de 20. Una vez elegidas estas componentes, se sustituyen con nuevos números aleatorios que poseen las estadísticas definidas inicialmente.

Debe señalarse que el proceso para la obtención de las derivadas de las superficies mutadas, al utilizar este esquema de mutación, es directo. Aunque el esquema propuesto originalmente por Schwefel [Schwefel, 1995] puede también ser utilizado en el espacio de Fourier, su implementación no es tan simple y directa como el esquema de la figura 32, por lo que se usó éste en todas la simulaciones realizadas.

Las simulaciones realizadas con las estrategias evolutivas se llevaron a cabo considerando las mismas condiciones que con los métodos simplex y de recocido simulado.

Para las dos estrategias evolutivas estudiadas, cada elemento de la población inicial fue una realización de un proceso aleatorio estacionario gaussiano con media cero y función de correlación gaussiana. Para la estrategia $(\mu + \lambda)$ se eligieron $\mu = 100$ y $\lambda = 100$ mientras que para la estrategia (μ, λ) fueron $\mu = 10$ y $\lambda = 100$. El número de iteraciones fue $g = 300$, que además se consideró como el criterio de terminación.

En los experimentos numéricos realizados se consideraron cinco perfiles aleatorios a recuperar, empezando desde treinta posiciones iniciales diferentes escogidas aleatoriamente.

Debe mencionarse que de esos treinta intentos, no todos recuperaron la superficie original. Sin embargo, se estableció como criterio que aquella superficie cuyo valor de la función objetivo asociado fuese el más bajo sería la reconstrucción a elegir. Coincidentemente, en la mayoría de los casos en los que la función objetivo era baja el perfil recuperado era similar al original.

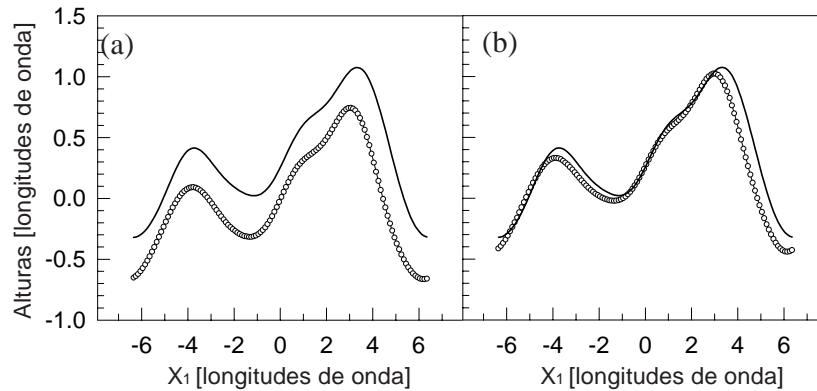


Figura 33: Perfiles reconstruidos con estrategias de evolución, (a) utilizando la estrategia (μ, λ) , b) utilizando la estrategia $(\mu + \lambda)$. El perfil original se muestra con la línea continua.

A continuación se presentan únicamente los resultados más significativos correspondientes a una sola de las cinco superficies estudiadas. El perfil utilizado para generar los datos de espacimiento se muestra en la figura 27. Las estadísticas de la superficie fueron las mismas que se habían definido en la sección IV.2.1. Con el fin de facilitar la comparación, en la figura 33 se muestran los resultados obtenidos al utilizar las dos estrategias evolutivas propuestas. En ambas gráficas el perfil original está dibujado con una línea continua.

Puede verse que en ambos casos el perfil ha sido recuperado en forma aceptable. El desplazamiento vertical del perfil en la figura 33(a) no se considera relevante en aplicaciones de perfilometría. Por otro lado, el desplazamiento no influye en el patrón de intensidad generado en el campo lejano, por lo cual el perfil recuperado es congruente con los datos suministrados al algoritmo.

Un resultado interesante, que ilustra la falta de unicidad en la solución obtenida al usar datos de intensidad se muestra en la figura 34. En la figura 34(a) se presentan el

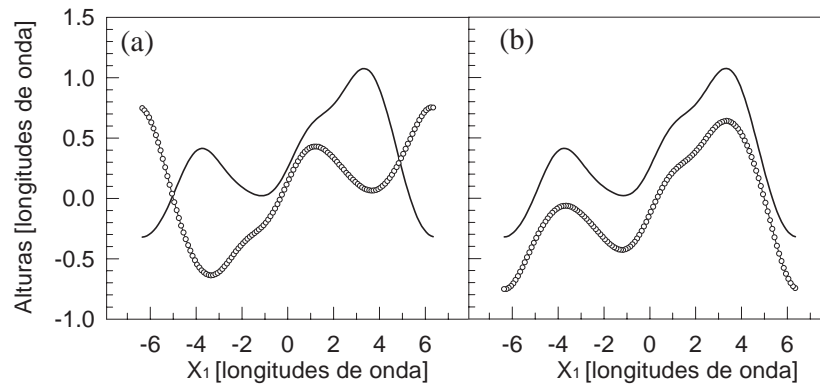


Figura 34: a) Perfil reconstruido, b) Perfil reconstruido reflejado con respecto a ambos ejes de simetría. El perfil original se muestra con una línea continua.

perfil original, con una línea continua, y el perfil recuperado usando la estrategia (μ, λ) , con una línea discontinua. La población inicial es distinta a la utilizada para obtener los resultados que se muestran en la figura 33. A simple vista, el perfil recuperado no tiene ninguna semejanza o relación con el original. Sin embargo, al llevar a cabo una reflexión respecto a los ejes de simetría x_1 y x_3 , se obtiene el perfil mostrado en la figura 34(b) que, salvo por el desplazamiento vertical, es muy parecido al perfil original. El patrón de esparcimiento obtenido con este perfil, considerando las mismas condiciones de iluminación que en el caso original, se muestra en la figura 28. Es claro que los patrones de esparcimiento son muy parecidos, lo cual ilustra la posibilidad de que dos perfiles diferentes generen el mismo patrón de esparcimiento. En el apéndice A se muestra que, bajo la Aproximación de Kirchhoff, el patrón de esparcimiento en intensidad es invariante ante una reflexión del perfil en x_1 y x_3 .

Influencia del ruido

Al igual que en el capítulo anterior, en esta sección se estudia el desempeño del algoritmo propuesto en presencia de ruido. Con este fin, se generó una matriz de datos ruidosos $\mathcal{I}_N^{(m)}(q | k)$, que se supuso igual a la suma de los patrones de moteado

$$\mathcal{I}_N^{(m)}(q | k) = I^{(m)}(q | k) + I_r(q | k), \quad (169)$$

donde $I^{(m)}(q | k)$ es la intensidad medida sin ruido e $I_r(q | k)$ es una variable aleatoria para la cual se escogieron las siguientes propiedades

$$P(I_r) = \exp \left\{ -\frac{I_r}{\langle I_r \rangle} \right\} \quad (170)$$

En la figura 35, se presentan las reconstrucciones obtenidas en presencia de ruido. En la figura 35a, se consideró que el ruido era igual al 5% del máximo pico de la intensidad esparcida $I^{(m)}(\zeta(x_1))$. Se observa que el perfil recuperado coincide con el original (línea continua). El mismo comportamiento se observa en la figura 35b, donde el nivel de ruido era igual al 10% del máximo pico de intensidad. Se presenta un desplazamiento vertical en la reconstrucción. Sin embargo, la forma de ésta coincide con la del perfil. En estas simulaciones se encontró que si el ruido era mayor al 10 % de la máxima intensidad medida, el algoritmo fallaba y la convergencia se daba en un óptimo local.

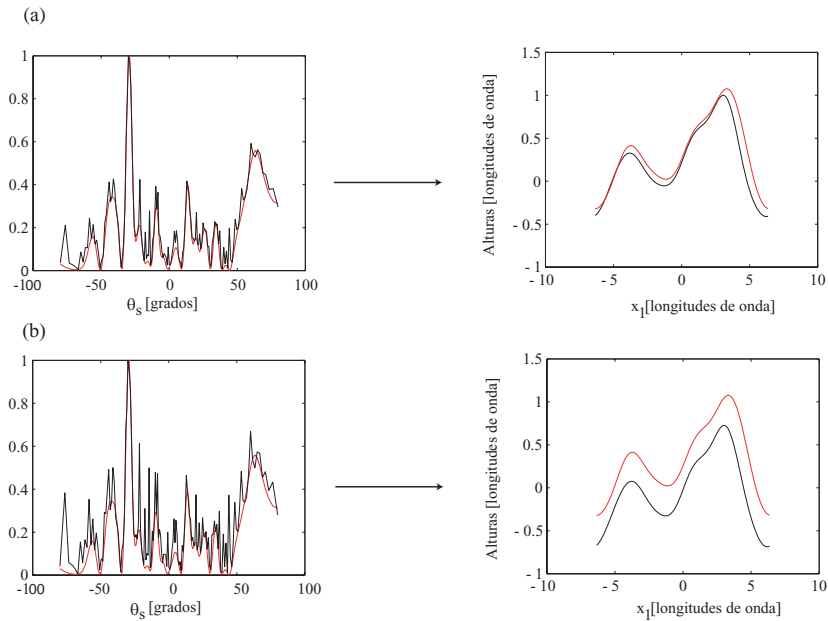


Figura 35: Reconstrucción de perfiles con datos ruidosos utilizando la estrategia de evolución (μ, λ) . (a) Reconstrucción con ruido al 5% del valor máximo de intensidad esparcida, b) Reconstrucción con ruido al 10% del valor máximo de intensidad esparcida. El perfil original se muestra con una línea continua.

IV.3 Recuperación de perfiles utilizando funciones B-spline

Con base en los resultados de la sección precedente pueden hacerse las siguientes observaciones:

- Las superficies cuyo perfil se buscaba recuperar estaban restringidas a una clase estadística específica.
- Una limitación importante del método espectral es la generación de superficies determinísticas. Debido a que se parte del supuesto que el espectro de potencias es conocido, se complica el tratamiento de superficies de tipo más general.

- El esquema de recombinación intermedia no puede utilizarse, ya que después de algunas iteraciones no se puede asegurar que las estadísticas de los elementos de la población secundaria serán iguales a las especificadas originalmente.

Las restricciones anteriores son consecuencia del esquema de representación utilizado. Una opción viable para salvarlas sería el uso de la Aproximación de Schoenberg expuesta en la sección II.3.2. Inicialmente, este esquema de representación alternativo permitiría la generación de superficies aleatorias o determinísticas continuas y derivables. Además, debido a que no se tienen restricciones en cuanto a que las superficies generadas pertenezcan a una clase estadística definida, sería posible utilizar la recombinación intermedia durante el proceso de búsqueda. A continuación se presentan los resultados obtenidos empleando este esquema alternativo.

IV.3.1 Algoritmo de inversión híbrido $(\mu/\rho, \lambda) + \text{Simplex}$

En particular, con este algoritmo híbrido el proceso de búsqueda se lleva a cabo en dos etapas. La primera consiste en determinar, por medio de una estrategia de evolución, una primera aproximación al óptimo. La segunda etapa comienza cuando la estrategia de evolución ha alcanzado el criterio de terminación. En ese momento, se utiliza el método simplex para encontrar el óptimo, incluyendo la solución encontrada por la estrategia de evolución en la población inicial del simplex.

La población inicial para las simulaciones realizadas es un conjunto de μ superficies aleatorias representadas mediante la ecuación (79). La estructura de la secuencia de

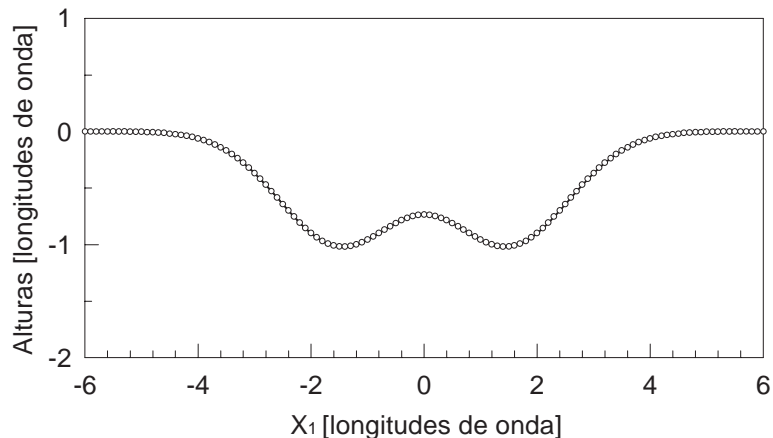


Figura 36: Perfil empleado para la generación de los datos de esparcimiento.

nodos utilizada corresponde a la descrita en la sección II.3.2 [ecuación (77)].

Por otro lado, las ordenadas α_j de los puntos de control $P_j(\tau_j, \alpha_j)$ son generadas, inicialmente, utilizando números aleatorios gaussianos con media cero y desviación estándar σ igual a la unidad. Por simplicidad se supone que ni las abscisas τ_j ni la secuencia de nodos $\{t_{ij}\}_1^{m+k}$ sufren cambios durante el proceso de búsqueda. Por otro lado, como consecuencia de la aplicación de los operadores genéticos, las ordenadas α_j cambian su posición en cada iteración del algoritmo al desplazarse únicamente en la dirección vertical (paralela al eje x_3). El valor de aptitud asociado a cada elemento de la población es calculado empleando la ecuación (67).

Para las simulaciones presentadas se buscó recuperar el perfil que se muestra en la figura 36. Debido a que el tiempo de cómputo requerido para hallar el óptimo aumenta cuando el número de puntos de muestreo se incrementa, la superficie elegida contiene $N = 121$ puntos muestreados a intervalos de $\lambda/10$.

Los datos de esparcimiento a partir de los cuales se recuperaría el perfil se generaron

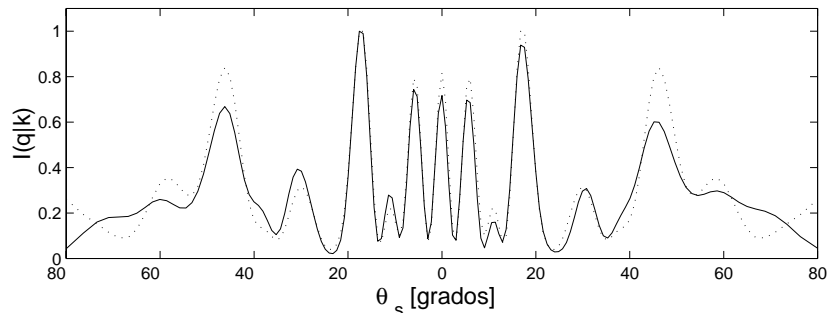


Figura 37: Patrón de intensidad esparcida por la superficie mostrada en la figura 36, para incidencia normal. La línea punteada muestra el patrón de esparcimiento producido por la superficie en la figura (39b)(línea continua) bajo las mismas condiciones de iluminación.

considerando las condiciones de iluminación descritas en la sección IV.2. El patrón de esparcimiento obtenido cuando se ilumina la superficie mostrada en la figura 36, a incidencia normal y con polarización s, se muestra(línea continua) en la figura 37.

En cada una de las estrategias estudiadas, la población inicial fué un conjunto de superficies aleatorias. Para las estrategias elitistas $(\mu + \lambda)$ y $(\mu/\rho + \lambda)$ se eligió $\mu = \lambda = 100$, mientras que para las no elitistas, (μ, λ) y $(\mu/\rho, \lambda)$, se tiene $\mu = 10$ y $\lambda = 100$. La mutación se llevó a cabo por medio de la ecuación (165), sin emplear técnicas de autoadaptación. El número de elementos, de la población inicial, que se seleccionaran para efectuar la recombinación fue $\rho = 2$ y el esquema utilizado fue la recombinación intermedia que se muestra en la ecuación (166). El criterio de terminación fue establecido con base en el máximo número de iteraciones, que para los experimentos numéricos realizados fue, $g = 300$.

Se utilizaron los algoritmos híbridos descritos para buscar el óptimo, partiendo de 30 posiciones distintas del espacio de búsqueda, elegidas aleatoriamente. Al igual que en los experimentos numéricos mostrados en la sección IV.2.3, no en todos los

intentos se recuperó el perfil original. Sin embargo, en la mayoría de los casos el valor de la función de costo más bajo encontrado, en las treinta simulaciones realizadas, nuevamente estaba asociado con una reconstrucción que asemejaba el perfil original. Por lo tanto, este fue el criterio establecido para determinar si una superficie había sido recuperada o no.

En la figura 38, se presentan las gráficas de convergencia para una corrida de cada uno de los algoritmos estudiados en esta parte del trabajo. El lado izquierdo corresponde a las estrategias de evolución utilizadas. La línea punteada muestra el comportamiento de la función de costo, en cada iteración, para las estrategias de evolución canónicas. A su vez, la línea continua corresponde a las estrategias de evolución que incluyen recombinación en su proceso de búsqueda. El lado derecho de la figura 38, muestra el desempeño del algoritmo cuando se utiliza el Simplex una vez que las estrategias de evolución han convergido a una solución. La discontinuidad que aparece en las gráficas que corresponden a las estrategias no elitistas [(b) y (c)] es consecuencia de la inclusión, en la población inicial del simplex, del mejor óptimo hallado por la estrategia de evolución en todo el proceso de búsqueda. Es decir, la mejor superficie reconstruida por la estrategia de evolución no necesariamente fue encontrada en la última iteración. Es importante hacer notar que los valores más altos de la función están asociados con las estrategias canónicas $(\mu^+\lambda)$ -ES [figuras 38(c) y (d)].

En la figura 39 se muestran los resultados obtenidos con los algoritmos híbridos $(\mu/\rho^+\lambda)$ -ES + Simplex. Para facilitar la comparación, el perfil original está dibujado con círculos. El óptimo hallado con la estrategia $(\mu/\rho + \lambda)$ es la línea punteada en la

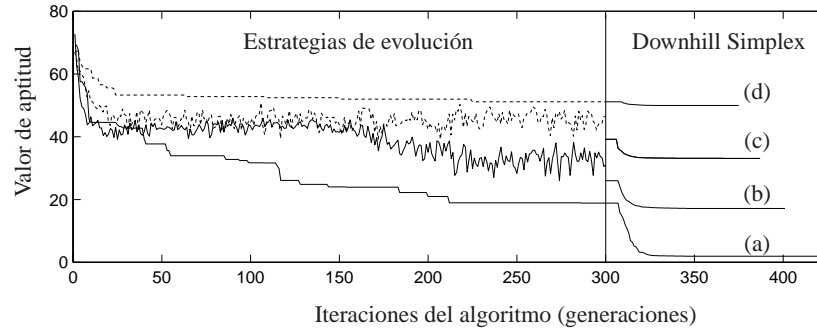


Figura 38: Gráfica de convergencia para la mejor solución encontrada por cada una de los algoritmos estudiados. Se tiene, respectivamente, (a) Estrategia $(\mu/\rho+\lambda)$ +Simplex, (b)Estrategia- $(\mu/\rho, \lambda)$ +Simplex, (c)Estrategia- (μ, λ) +Simplex y (d) Estrategia- $(\mu + \lambda)$ + Simplex.

figura 39(a). En este caso, la reconstrucción obtenida asemeja al perfil original salvo por algunas diferencias que se observan en los extremos de la superficie. La reconstrucción generada empleando el método simplex se muestra con una línea continua en la figura 39(a). Puede apreciarse que las diferencias en los extremos han desaparecido. A su vez, tal como se mencionó anteriormente, el desplazamiento vertical en la figura 33(a) resultante de utilizar datos de intensidad para llevar a cabo la reconstrucción, puede considerarse irrelevante en aplicaciones tales como perfilometría de superficies.

La reconstrucción de la figura 39(b) ilustra la multimodalidad de la función, cuando se utilizan datos de intensidad. Al igual que en la figura 34, el perfil obtenido asemeja al original pero reflejado con respecto a los ejes x_3 y x_1 . Un resultado aún más sorprendente puede obtenerse si se ilumina la reconstrucción de la figura 39b con una onda plana y se compara el patrón de esparcimiento generado, bajo las mismas condiciones de iluminación, con el correspondiente a la superficie original. La línea punteada en la figura 28 corresponde al patrón generado por la reconstrucción de la figura 39(b).

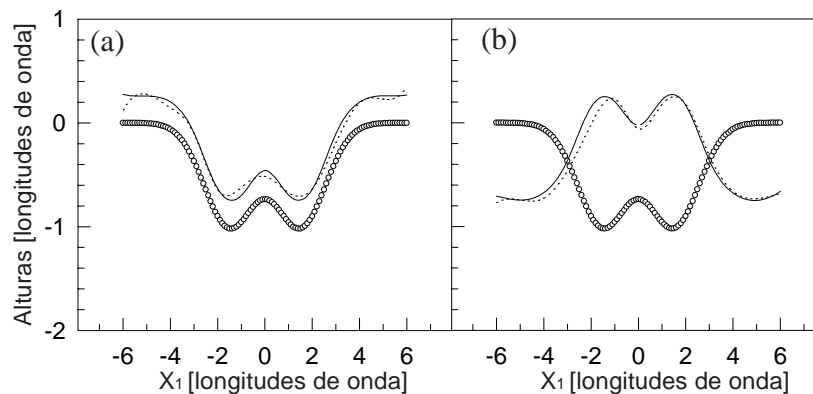


Figura 39: Reconstrucción del perfil utilizando: (a) La estrategia $(\mu/\rho+\lambda)$ -ES+ Downhill simplex, b) La estrategia $(\mu/\rho, \lambda)$ -ES+ Downhill simplex. El perfil original se muestra con círculos y las reconstrucciones están dibujadas con una línea punteada para las dos estrategias de evolución y con una línea continua para el método simplex.

Entonces, es claro que dos superficies rugosas con perfiles diferentes pueden generar el mismo patrón de esparcimiento. Los argumentos para visualizar o comprender este resultado son los mismos que se expusieron en la sección IV.2.3, y que aparecen en el apéndice A.

Debe mencionarse que para la obtención de los resultados presentados en la figura 39 no solamente se consideraron las mismas condiciones de iluminación, sino que la búsqueda del óptimo comenzó en el mismo punto del espacio de soluciones. La única diferencia entre ambas simulaciones fue el esquema de selección utilizado y este hecho ilustra su importancia en la exploración del espacio de búsqueda. Los resultados obtenidos con las estrategias híbridas $(\mu^+\lambda)$ -ES + Simplex se presentan en la figura 40. Los parámetros iniciales, así como las condiciones de iluminación, fueron las mismas que se utilizaron para las reconstrucciones mostradas en la figura (39). Sin embargo, para llevar a cabo la búsqueda del óptimo únicamente se utilizó la mutación. Puede

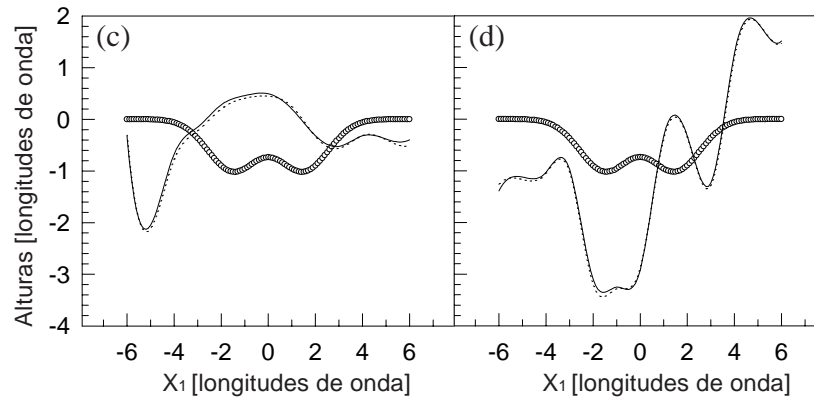


Figura 40: Reconstrucciones obtenidas utilizando las estrategias: (c) $(\mu, \lambda)+$ simplex y (d) $(\mu + \lambda)+$ simplex . El perfil original se muestra con círculos. Se utiliza una línea punteada para las reconstrucciones correspondientes a las estrategias de evolución y una línea continua para las reconstrucciones obtenidas con el Método Simplex.

observarse que las reconstrucciones obtenidas no se asemejan al perfil original. Este hecho, a su vez, corresponde a un valor alto de la función de costo, tal como se mostró en la figura 38. A su vez, la escasa diferencia entre las reconstrucciones dibujadas con líneas continuas y aquellas dibujadas con líneas punteadas muestra la tendencia del simplex a estancarse en un óptimo local, si inicia su búsqueda en una posición alejada de la vecindad del óptimo global.

IV.4 Conclusiones Parciales

En este capítulo se mostró la posibilidad de resolver el problema de esparcimiento inverso utilizando métodos estocásticos. A su vez, se mostró la importancia del esquema de representación de las variables objeto en la elección de un algoritmo de optimización adecuado al problema.

Por otro lado, de los experimentos numéricos realizados, se concluye que la represen-

tación con funciones spline básicas (B-splines) no solamente permite generar superficies deterministas o aleatorias de manera conveniente y sencilla, sino que además, abre la posibilidad de utilizar el operador de recombinación durante el proceso de búsqueda. En contraste, la representación espectral requiere del conocimiento *a priori* de las estadísticas de una superficie, para el caso de perfiles aleatorios, y no permite la recombinación durante la búsqueda del óptimo.

Finalmente, se mostró que la hibridación de las estrategias de evolución con un algoritmo de búsqueda local, dá como resultado una mejora en la solución encontrada, siempre y cuando la primera fase haya acercado a la solución a la vecindad del óptimo global.

Capítulo V

Discusión y Conclusiones

En los capítulos III y IV se presentaron dos técnicas de inversión para recuperar el perfil de superficies rugosas utilizando información de esparcimiento en el campo lejano. Se expusieron sus características principales y por medio de ejemplos específicos, se estudió su desempeño en diferentes situaciones.

En la primera parte de este capítulo se pretende dar una visión general del trabajo desarrollado en esta tesis mediante una comparación entre los métodos de inversión propuestos. Se consideran tres casos particulares que sirvan para ilustrar aspectos importantes de sus respectivos desempeños, así como también de la complejidad del problema estudiado. Posteriormente, con base en los resultados obtenidos, se presentará un breve resumen y se enunciarán las conclusiones generales del trabajo.

V.1 Discusión

Antes de proceder a la discusión de resultados, es conveniente describir las condiciones iniciales establecidas para llevar a cabo las simulaciones.

Los datos de entrada, ya sea de amplitud o de intensidad, fueron generados mediante el método riguroso descrito por [Maradudin *et al.*, 1990]. Se supuso que el campo incidente era una onda plana con polarización s . A su vez, se consideró una superficie perfectamente conductora con una longitud física de $L = 12\lambda$, muestreada a intervalos de $\Delta x = \frac{\lambda}{10}$.

Para los algoritmos evolutivos se utilizaron $N_{ang} = 4$ ángulos de incidencia, que fueron $\theta_0 = -60$, $\theta_0 = -30$, $\theta_0 = 0$ y $\theta_0 = 40$. Para el algoritmo de empatamiento de frentes de onda, con base en la longitud física elegida e invocando la relación (112), se determinó que el menor número de ángulos de incidencia y esparcimiento necesarios para reconstruir la superficie apropiadamente era $N_{ang} = 24$. A su vez, se consideró un ángulo de corte $\theta_{0m} = 80^\circ$ y el muestreo de q y de k se hizo de manera uniforme.

En cada uno de los ejemplos que se muestran, se consideraron treinta estados iniciales diferentes para las estrategias de evolución. Sin embargo, únicamente se presentarán los resultados más significativos.

En el primer caso que se presenta, la superficie a reconstruir fue generada con el método espectral descrito en la sección II.3.1. Los parámetros estadísticos que se consideraron fueron una longitud de correlación $a = 2.5\lambda$ y una desviación estándar de alturas $\delta = 1.0\lambda$. Los resultados obtenidos se muestran en la figura 41. En el mapa de

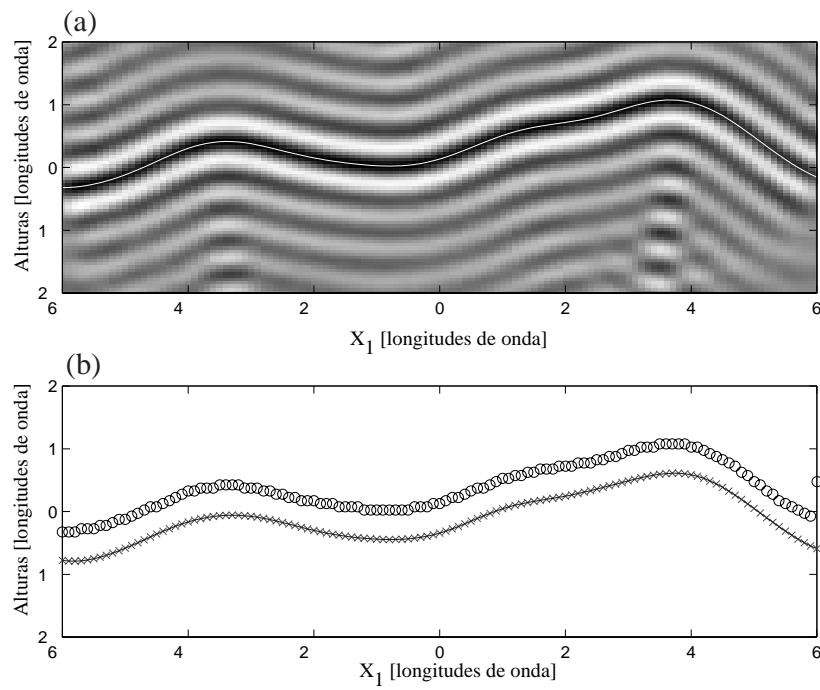


Figura 41: Reconstrucción de una superficie aleatoria perfectamente conductora : (a) Mapa de la función $F_{real}^{s-p}(\xi, \eta)$ empleando el algoritmo de empatamiento de frentes de onda, b) Perfiles estimados con los algoritmos de inversión propuestos.

la función $F_{real}^{s-p}(\xi, \eta)$ [figura 41(a)] se observa una franja central oscura que sigue la forma del perfil original, trazado con una línea blanca para facilitar su visualización. El superíndice $s - p$ indica que para la generación del mapa se empleó el esquema de disminución de múltiple esparcimiento, descrito en el capítulo III. En la figura 41(b) se presenta, dibujado con círculos, el perfil obtenido después de hallar los mínimos de la función $F_{real}^{s-p}(\xi, \eta)$. También se muestran, dibujados con una línea continua y con cruces, los perfiles obtenidos a la salida de las estrategias híbridas $(\mu/\rho + \lambda)$ y $(\mu/\rho, \lambda)$, respectivamente. La similitud entre las reconstrucciones es evidente, salvo por el desplazamiento vertical que se aprecia en la figura 41(b) y que es consecuencia de utilizar datos de intensidad y no de amplitud, como se discutió en el capítulo IV. En este primer ejemplo las tres familias de algoritmos considerados dan buenos resultados.

En el capítulo IV se vió que si la interacción de la luz con la superficie está dominada por esparcimiento sencillo (es decir, si la aproximación de Kirchhoff es válida) la reflexión del perfil con respecto a los ejes x_1 y x_3 da como resultado una superficie que genera el mismo patrón de esparcimiento en intensidad. Con respecto a la unicidad de la solución, esto parece indicar que es conveniente estar en situaciones en las que el esparcimiento múltiple juega un papel importante en la solución del problema directo. Resulta entonces interesante estudiar el desempeño de los algoritmos de inversión con superficies que dan origen a efectos de esparcimiento múltiple. Se considera ahora un surco triangular de ancho 2λ y profundidad λ , como se muestra en la figura 42. Vale la pena señalar que las pendientes del surco son de $\pm 45^\circ$, por lo que la ocurrencia de

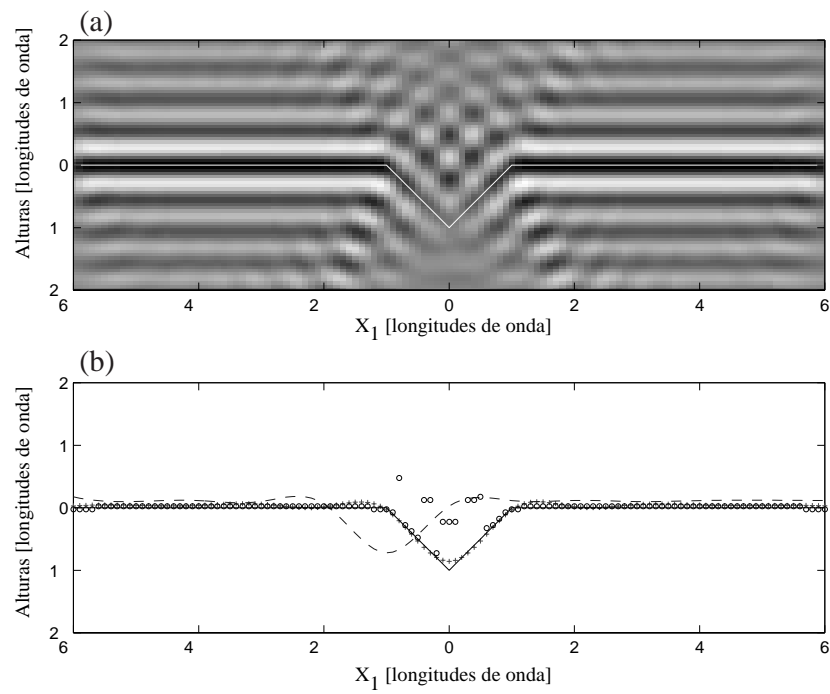


Figura 42: Reconstrucción de un surco triangular con pendientes de 45° utilizando: (a) Mapa de la función $F_{real}^{s-p}(\xi, \eta)$ empleando el algoritmo de empatamiento de frentes de onda, b) Perfiles estimados con los algoritmos de inversión propuestos.

esparcimiento múltiple es clara.

Los resultados obtenidos al aplicar los algoritmos de inversión propuestos se presentan en la figura 42. El mapa de la función $F_{real}^{s-p}(\xi, \eta)$ se muestra en la figura 42(a). A diferencia del caso anterior, se observa que el perfil original únicamente coincide con la franja oscura en los tramos horizontales. En la parte central, que corresponde al surco, se tiene un patrón de franjas más complicado y del que resulta difícil estimar el perfil. Este resultado era predecible, ya que el algoritmo de empatación de frentes de onda se basa en principios de esparcimiento sencillo. En la figura 42(b) se muestra que el perfil estimado localizando los mínimos del mapa de F_{real}^{s-p} presenta discontinuidades. Por otro lado, el perfil reconstruido con las estrategias de evolución propuestas es una versión suavizada del perfil original. La reconstrucción correspondiente a la estrategia no elitista está dibujada con cruces, mientras que la línea discontinua se utiliza para la estrategia elitista. Con esta última se ve que el surco aparece movido hacia la izquierda. La posible convergencia a una superficie corrida es comprensible, ya que la amplitud compleja del campo esparcido por un perfil desplazado difiere de la amplitud compleja del surco original, únicamente por un factor lineal de fase. Es decir, que las intensidades son iguales. A su vez, debido a que solamente se cuenta con información de esparcimiento en el campo lejano, la mejor resolución que se puede esperar en la reconstrucción es del orden de $\lambda/2$, de ahí el origen del suavizado del perfil recuperado.

Un resultado interesante, que a su vez ilustra una dificultad adicional para la resolución del problema, se obtiene cuando el ancho del surco se reduce en 0.5λ y la

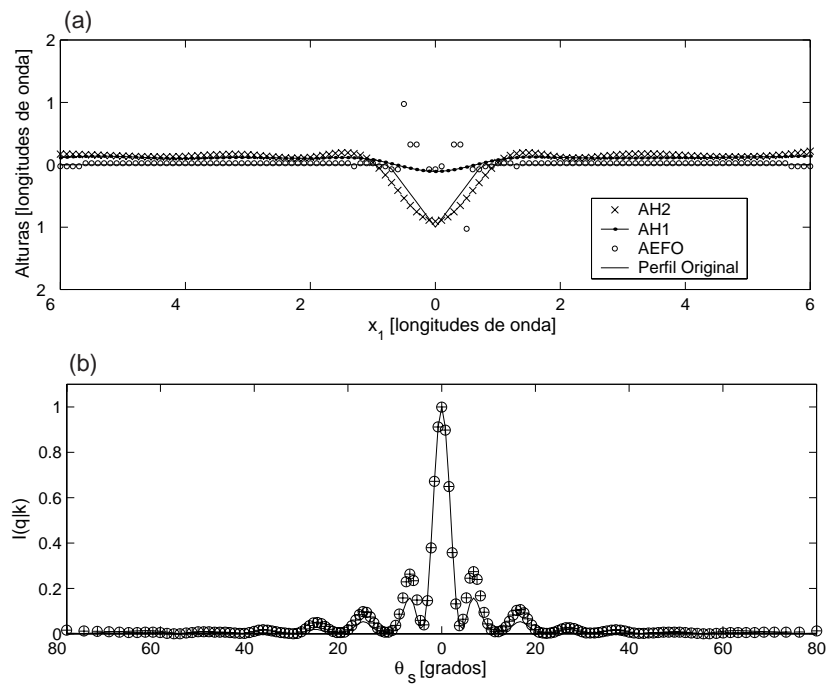


Figura 43: Reconstrucción de un surco triangular con pendientes de 55° . (a) Perfiles estimados con el AEFO y la estrategia híbrida $(\mu/\rho, \lambda)+$ simplex, considerando dos inicializaciones distintas. (b) Comparación entre el patrón de esparcimiento producido por el perfil original (línea continua) y los patrones generados por las reconstrucciones mostradas en la gráfica (a) de esta figura.

profundidad permanece constante. Es decir, que el ancho es ahora de 1.5λ y la profundidad de λ . En este caso, las pendientes del surco central se incrementan en 10° con respecto a las del ejemplo anterior. La línea continua de la figura 43(a) muestra el perfil original, mientras que el perfil estimado por el algoritmo de empatamiento de frentes de onda está dibujado con círculos. Las líneas restantes corresponden a los perfiles reconstruidos considerando dos estados iniciales distintos de la estrategia $(\mu/\rho, \lambda)$. Los saltos que aparecen en la estimación del perfil con el AEFO son una clara indicación de que éste falla ante situaciones de este tipo. A su vez, para una inicialización, el algoritmo híbrido no elitista genera una versión suavizada del surco, mostrada con cruces en la figura 43(a). Para la otra inicialización, se tiene una superficie suave con una pequeña depresión que coincide con la posición del surco original, pero que es menos profunda. El valor de aptitud es prácticamente el mismo en estos dos casos ($f \approx 17$ para ambas), de manera que con esta información, no es posible decidir cual sería la mejor solución.

En la figura 43(b) se muestran los patrones de intensidad esparcida correspondientes al perfil original (línea continua), . La línea con cruces muestra el patrón de esparcimiento generado por la reconstrucción AH2 de la figura 43(a) y la línea restante corresponde a la reconstrucción AH3. Se observa que los patrones de intensidad producidos por las tres superficies son similares, por lo que esta información no es suficiente para decidir cual de las dos superficies es la más cercana a la que generó los datos. En la figura 44 se muestran las amplitudes esparcidas, junto con sus respectivas fases, en función de los números de onda k y q . Mientras los patrones de amplitud guardan

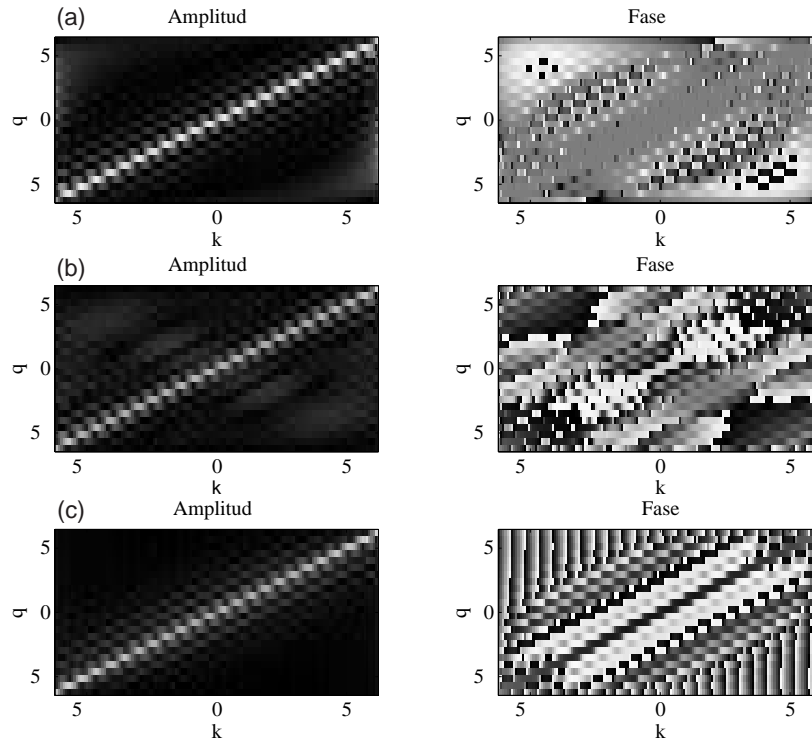


Figura 44: Información de amplitud y de fase esparcida por (a) el surco triangular dibujado con una línea continua en la figura 43a, (b) la reconstrucción dibujada con cruces en la figura 43a y (c) la reconstrucción dibujada con una línea punteada en la figura 43a.

cierta semejanza entre ellos, la fase difiere totalmente.

Este ejemplo particular ilustra dos situaciones en las que vale la pena ahondar. La primera es la existencia de dos soluciones que satisfacen a la función costo y de las cuales solamente una se asemeja al perfil original. La segunda situación es que a pesar de las limitaciones en la resolución, explicadas en párrafos anteriores, y de la falta de unicidad en la solución, el algoritmo híbrido arroja resultados promisorios como los mostrados en la figura 45 para las polarizaciones s y p , figuras 45(a) y 45(b), respectivamente. Las reconstrucciones están dibujadas con líneas continuas, mientras que al perfil original le corresponden las líneas punteadas. El hecho de que dos de

las reconstrucciones en la figura 45(b) sean más profundas que el perfil original puede deberse al esquema de representación y quizá sea necesario incrementar el número de puntos de control para tener una mayor precisión en los resultados.

Por otro lado, volviendo a la falta de unicidad en la solución del problema, si bien encontrar una manera de favorecerla no es un trabajo sencillo y requiere un estudio cuidadoso, los resultados anteriores sugieren como posible punto de partida la inclusión de los efectos de polarización en el proceso de búsqueda. Es decir, la función de costo inicial $f(\zeta(x_1), k_j)$ (ecuación (67)), podría redefinirse, por ejemplo, como

$$f(\zeta(x_1), k_j) = \sum_{j=1}^{N_{ang}} \int [|I_s^{(m)}(q|k_j) - I_s^{(c)}(q|k_j) + |I_p^{(m)}(q|k_j) - I_p^{(c)}(q|k_j)|] dq. \quad (171)$$

De esta forma, las posibles soluciones al problema estarán restringidas a solamente aquellas que satisfagan, simultáneamente, las condiciones impuestas al incluir las polarizaciones s y p.

Es necesario hacer notar que existen varias formas de reducir el espacio de búsqueda, además de la mostrada en la ecuación (171). Sin embargo, modificar la función de costo requiere de cierto cuidado, ya que este cambio repercutirá directamente en su topología y no solo se podría incrementar la complejidad del problema, sino incluso volverlo irresoluble.

Otra manera en la que se podría buscar la unicidad de la solución sería utilizando un esquema de iluminación alterno, en el que se incluyera en el proceso de búsqueda la información generada al iluminar con haces diferentes partes de la superficie. Cabe

señalar que opción podría ser especialmente útil para superficies con una longitud física mayor que la de las superficies consideraras en este trabajo.

No es posible afirmar *a priori* si la redefinición de la función de costo, el uso de esquemas de iluminación alternos o el empleo de un mayor número de puntos de control contribuirán a la unicidad de la solución o no, ya que es necesario realizar una gran cantidad de experimentos numéricos para poder aseverarlo. Sin embargo, con base en los resultados mostrados, puede pensarse que el potencial de los algoritmos de inversión propuestos no ha sido del todo explorado. Este hecho abre la posibilidad de, en el futuro, proponer algoritmos cada vez más robustos que consideren las alternativas anteriores en su proceso de búsqueda y cuyos resultados no solamente satisfagan la condición de unicidad en la solución, sino también permitan extender su aplicación a la reconstrucción de superficies reales.

V.2 Resumen y conclusiones

En esta tesis se estudió un problema inverso de esparcimiento que consiste en la reconstrucción de los perfiles de superficies rugosas unidimensionales a partir de datos de esparcimiento obtenidos en el campo lejano. Se propusieron dos métodos de inversión cuya aplicabilidad fue probada en diferentes situaciones. El primero de ellos hace uso de la información de amplitud y de fase, mientras que el segundo utiliza únicamente datos de intensidad.

La primer metodología se basa en principios de empatamiento de frentes de onda.

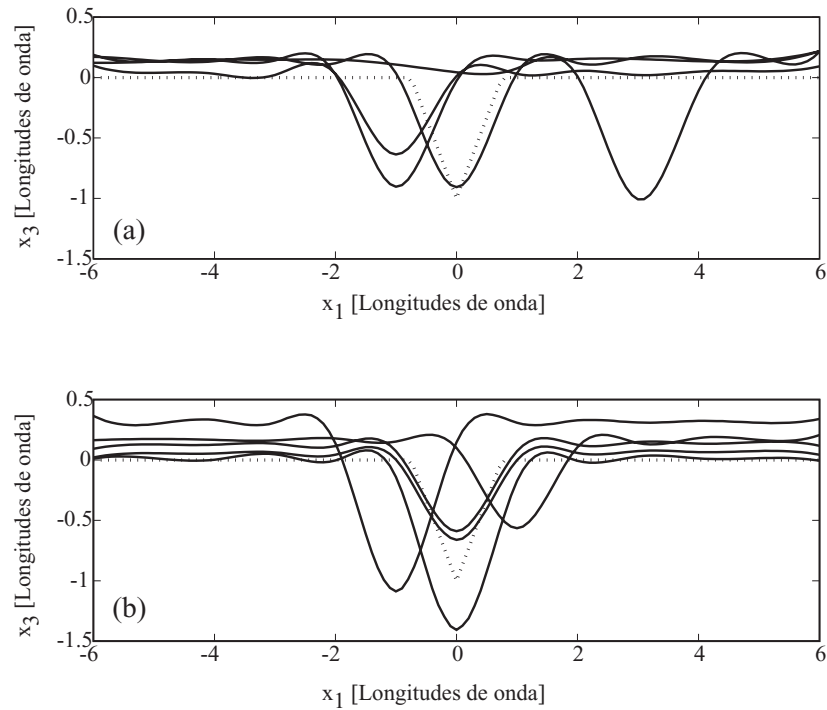


Figura 45: Reconstrucciones obtenidas con la estrategia $(\mu/\rho, \lambda)$ para (a) polarización s y (b) polarización p.

Se analizó su desempeño tanto para el caso de un conductor perfecto, como para medios dieléctricos y metálicos. A su vez, se estudió la influencia del muestreo en la calidad de las reconstrucciones obtenidas y se propuso una estrategia para evitar redundancia en la información utilizada. Se estudiaron las limitaciones del algoritmo en presencia de ruido. Se mostró que es posible mejorar la reconstrucción si se utiliza información de las dos polarizaciones elementales, a través de un esquema de disminución del espaciamiento doble.

También se estudió la aplicabilidad del algoritmo de empatamiento de frentes de onda a medios estratificados. Al igual que en el caso de una interface simple, los perfiles a reconstruir pueden ser aleatorios o deterministas. Por simplicidad, se restringieron

los cálculos a muestras de tres capas. Para reconstruir perfiles internos fue necesario utilizar una pupila correctora para compensar por las aberraciones introducidas cuando un haz convergente pasa de un medio a otro.

Inicialmente, sería posible aplicar el AEFO para reconstruir los perfiles de superficies bidimensionales. Sin embargo, hasta este momento, no se cuenta con datos de esparcimiento directo confiables con los cuales puedan llevarse a cabo las simulaciones.

En el segundo método de inversión estudiado se aborda el problema inverso como un problema de optimización no lineal. Para aligerar los cálculos numéricos, se centró la atención sobre la reconstrucción de superficies rugosas unidimensionales perfectamente conductoras con polarización s .

Se mostró la posibilidad de recuperar el perfil de una superficie rugosa, partiendo de datos de intensidad en el campo lejano, sin necesidad de un paso intermedio para recuperar la fase del campo esparcido.

Las estrategias de evolución basadas en los dos esquemas de representación utilizados arrojaron resultados promisorios, comparables y en algunos casos mejores que los obtenidos con el AEFO. Sin embargo, cabe señalar que la representación espectral tiene una limitación importante, ya que es necesario conocer *a priori* la clase estadística a la que pertenece la superficie y no permite operaciones de recombinación. Es decir, que un número ρ de elementos pertenecientes a la población inicial se mezclen entre sí. En contraste, la representación con curvas spline sí permite operaciones de recombinación en la búsqueda del óptimo. Se obtuvieron mejores resultados con $\rho = 2$. Se observó que el uso de recombinación múltiple contribuía a la divergencia de los algoritmos de

inversión.

En los esquemas de selección utilizados se observó que la estrategia elitista, si bien garantiza un decremento monotónico del valor de aptitud, es proclive a la convergencia prematura a un óptimo local. En contraste, la estrategia no elitista evita el estancamiento al explorar el espacio de búsqueda en forma estocástica, permitiendo un deterioro temporal del valor de aptitud. A partir de las simulaciones realizadas, puede verse que las reconstrucciones obtenidas por ambos esquemas asemejan los perfiles originales. El establecimiento de una regla definida para decidir si una estrategia es mejor que la otra no es una tarea fácil, ya que para hacerlo en forma sistemática (por medio de técnicas estadísticas) es necesario realizar un gran número de experimentos numéricos, de manera que se tenga una muestra representativa del comportamiento de cada uno de los esquemas.

No se prevé ningún obstáculo para la aplicación de los métodos evolutivos propuestos a la reconstrucción de superficies metálicas o dieléctricas. La única consecuencia sería un incremento en el tiempo de cómputo necesario para resolver el problema directo en cada iteración. Sin embargo, dada la rapidez de los equipos de cómputo actuales, este hecho no representa una restricción significativa y, por otro lado, permitiría explorar la posibilidad de llevar a cabo reconstrucciones a partir de datos experimentales.

Una implicación importante al utilizar de datos de esparcimiento reales es el incremento en la longitud de las superficies a reconstruir. Este hecho repercutiría en la necesidad de ajustar parámetros tales como el número de elementos de las poblaciones inicial e intermedia, μ y λ respectivamente. A su vez, sería necesario incluir los

parámetros de aprendizaje τ , τ_0 y τ' para modificar las mutaciones en cada iteración. También, el número de elementos combinar ρ tendría que ser modificado. Sin embargo, debido a la naturaleza heurística de los algoritmos evolutivos, estos ajustes son, en sí mismos, parte del proceso de solución del problema particular que se está estudiando.

La aplicación de los algoritmos de inversión evolutivos a superficies bidimensionales no se ve factible en esta etapa. Ya que no solamente se incrementaría el grado de complejidad del problema, sino además la solución del problema directo en cada iteración haría al método muy costoso en términos computacionales.

En conclusión, se han presentado dos metodologías para la solución de un problema inverso en óptica, en particular, la reconstrucción de perfiles de superficies rugosas a partir de información de campo lejano.

Se mostró que la información de la fase del campo esparcido juega un papel fundamental en el proceso de reconstrucción del algoritmo de empatamiento de frentes de onda (AEFO). Los resultados obtenidos con este método indican que puede ser aplicado a una gran variedad de problemas que incluyen, bajo ciertas restricciones, medios estratificados. El AEFO presenta una buena inmunidad al ruido y su desempeño es estable aún en situaciones en donde hay esparcimiento múltiple. Sin embargo, el costo de su robustez, y de hacer uso de la información de fase, es que la adquisición de datos experimentales sea complicada y por lo tanto su aplicación en la reconstrucción de superficies reales no sea una proposición práctica.

Por otro lado, se ha abierto la posibilidad de recuperar el perfil de una superficie rugosa utilizando solamente datos de intensidad. El problema puede resolverse formu-

lándolo como un problema de optimización no lineal con restricciones. Se ha mostrado el potencial de las estrategias evolutivas para la solución de problemas inversos en óptica. Los resultados obtenidos son promisorios abren una amplia gama de aplicaciones tales como en perfilometría de superficies y ,por otro lado, también podrían utilizarse en problemas inversos de microscopía de campo cercano.

Apéndice A

El problema de la unicidad al utilizar datos de intensidad

El resultado presentado en las gráficas 34a y 39b correspondientes a las secciones IV.2.3 y IV.3, respectivamente, puede explicarse utilizando la aproximación de Kirchhoff. En la sección II.1.1 se demostró que la amplitud esparcida por una superficie $\zeta(x_1)$ perfectamente conductora tenía la forma

$$R^{AK}(q|k) = F(\theta_i, \theta_s) \int_{-\infty}^{\infty} dx'_1 \exp \{iv_x x'_1 - iv_z \zeta(x'_1)\} \quad (\text{A1})$$

donde $F(\theta_0, \theta_s)$ es un factor angular dado por:

$$F(\theta_0, \theta_s) = \pm \frac{1 + \cos(\theta_0) \cos(\theta_s) - \sin(\theta_0) \sin(\theta_s)}{\cos(\theta_0) + \cos(\theta_s)} \quad (\text{A2})$$

Si ahora se considera un nuevo perfil $z(x_1) = -\zeta(-x_1)$, la ecuación A1 se reescribe

como

$$R^{(z)}(q|k) = F(\theta_i, \theta_s) \int_{-\infty}^{\infty} dx'_1 \exp \{i v_x x'_1 + i v_z \zeta(-x'_1)\} \quad (\text{A3})$$

Haciendo el cambio de variable $u = -x_1$ y tomando el complejo conjugado de la expresión resultante se encuentra que:

$$R(q | k) = [R^{(z)}(q | k)]^*. \quad (\text{A4})$$

A partir de la ecuación (A4) es claro que ambas superficies generarán el mismo patrón de esparcimiento.

Apéndice B

Esquema de reducción de esparcimiento múltiple

En la primera parte de este apéndice se presenta un método aproximado para calcular en forma iterativa, considerando el caso de un conductor perfecto, las funciones fuente $\varphi(x_1 | \omega)$ y $\chi(x_1 | \omega)$. Posteriormente, con base en este resultado, se fundamenta el esquema de disminución de esparcimiento múltiple, representado por la ecuación (96).

Con base en la geometría mostrada en la figura 1, el campo total en la región $x_3 > \zeta(x_1)$ estará dado por la suma del campo incidente y el esparcido (ver las ecuaciones (8) y (9) de la sección II.1)

$$\psi(\mathbf{r}) = \psi(\mathbf{r})_{inc} + \psi(\mathbf{r})_{sc}, \quad (\text{B1})$$

donde

$$\psi(\mathbf{r})_{sc} = \frac{1}{4\pi} \int_{-\infty}^{\infty} \left(\varphi(\mathbf{r}' | \omega) \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) - G_0(\mathbf{r}' | \mathbf{r}) \chi(\mathbf{r}' | \omega) \right) dS', \quad (\text{B2})$$

Para el caso de un conductor perfecto y considerando la condición de frontera de Dirichlet (polarización s), la ecuación (B1) puede escribirse de la forma

$$\psi(\mathbf{r}) = \psi_{inc}(\mathbf{r}) - \frac{1}{4\pi} \int_{-\infty}^{\infty} dS' G_0(\mathbf{r}' | \mathbf{r}) \chi(\mathbf{r}' | \omega). \quad (\text{B3})$$

Por otro lado, para la condición de frontera de Neumann (polarización p) se tiene que

$$\psi(\mathbf{r}) = \psi_{inc}(\mathbf{r}) + \frac{1}{4\pi} \int_{-\infty}^{\infty} dS' \varphi(\mathbf{r}' | \omega) \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}). \quad (\text{B4})$$

Las funciones fuente en las ecuaciones B3 y B4 corresponden a las expresiones (20) y (19), respectivamente.

Siguiendo el procedimiento descrito por [Liszka y McCoy, 1982] (ver también [Maradudin *et al.*, 1990]), se puede mostrar que las respectivas funciones fuente (el campo y su derivada normal en la frontera $x_3 = \zeta(x_1)$) satisfacen las siguientes ecuaciones

$$\chi(\mathbf{r}' | \omega) = 2 \frac{\partial}{\partial n'} \psi_{inc}(\mathbf{r}') - 2 \frac{P}{4\pi} \int_{-\infty}^{\infty} dS' G_0(\mathbf{r}' | \mathbf{r}) \chi(\mathbf{r}' | \omega) \quad (\text{B5})$$

y

$$\varphi(\mathbf{r}' | \omega) = 2 \psi_{inc}(\mathbf{r}') + 2 \frac{P}{4\pi} \int_{-\infty}^{\infty} dS' \varphi(\mathbf{r}' | \omega) \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}), \quad (\text{B6})$$

donde la P indica el valor principal de Cauchy.

Las ecuaciones (B5) y (B6) son ecuaciones integrales tipo Fredholm de segunda especie, cuya solución puede determinarse mediante un desarrollo en serie de Neumann-Liouville, de la forma

$$\chi(\mathbf{r}' | \omega) = D^{(1)}(\mathbf{r}' | \omega) + D^{(2)}(\mathbf{r}' | \omega) + D^{(3)}(\mathbf{r}' | \omega) + \dots, \quad (\text{B7})$$

donde

$$D^{(1)}(\mathbf{r}' | \omega) = 2 \frac{\partial}{\partial n'} \psi_{inc}(\mathbf{r}'), \quad (\text{B8})$$

$$D^{(m)}(\mathbf{r}' | \omega) = -2 \frac{P}{4\pi} \int_{-\infty}^{\infty} dS' \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) D^{(m-1)}(\mathbf{r}' | \omega), \quad \text{para } m > 1. \quad (\text{B9})$$

A su vez,

$$\varphi(\mathbf{r}' | \omega) = N^{(1)}(\mathbf{r}' | \omega) + N^{(2)}(\mathbf{r}' | \omega) + N^{(3)}(\mathbf{r}' | \omega) + \dots, \quad (\text{B10})$$

donde

$$N^{(1)}(\mathbf{r}' | \omega) = 2\psi_{inc}(\mathbf{r}'), \quad (\text{B11})$$

$$N^{(m)}(\mathbf{r}' | \omega) = 2 \frac{P}{4\pi} \int_{-\infty}^{\infty} dS' \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) N^{(m-1)}(\mathbf{r}' | \omega), \quad \text{para } m > 1. \quad (\text{B12})$$

La sustitución de las ecuaciones (B7) y (B10) en los términos del campo esparcido de las ecuaciones (B3) y (B4), permite escribir los campos esparcidos en términos de

sumas infinitas, dadas por

$$\psi_s(\mathbf{r} | \omega)_{sc} = \psi_s^{(1)}(\mathbf{r} | \omega)_{sc} + \psi_s^{(2)}(\mathbf{r} | \omega)_{sc} + \psi_s^{(3)}(\mathbf{r} | \omega)_{sc} + \dots \quad (\text{B13})$$

$$\psi_p(\mathbf{r} | \omega)_{sc} = \psi_p^{(1)}(\mathbf{r} | \omega)_{sc} + \psi_p^{(2)}(\mathbf{r} | \omega)_{sc} + \psi_p^{(3)}(\mathbf{r} | \omega)_{sc} + \dots \quad (\text{B14})$$

donde

$$\psi_s^{(m)}(\mathbf{r} | \omega)_{sc} = -2\frac{P}{4\pi} \int_{-\infty}^{\infty} dS' G_0(\mathbf{r}' | \mathbf{r}) D^{(m)}(\mathbf{r}' | \omega), \quad \text{para } m > 0. \quad (\text{B15})$$

y

$$\psi_p^{(m)}(\mathbf{r} | \omega)_{sc} = 2\frac{P}{4\pi} \int_{-\infty}^{\infty} dS' \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) N^{(m)}(\mathbf{r}' | \omega), \quad \text{para } m > 0. \quad (\text{B16})$$

Puede observarse que el primer término de las ecuaciones (B13) y (B14), representa el campo esparcido en la aproximación de Kirchhoff. Es decir, se trata de una aproximación de esparcimiento sencillo. Por otro lado, el término $\psi^{(m)}(\mathbf{r} | \omega)_{sc}$ la amplitud de esparcimiento debida a m interacciones sucesivas entre el campo y la superficie [Liszka y McCoy, 1982].

Utilizando la aproximación de Kirchhoff, en la sección II.1.1 [ecuación (37)], se demostró que

$$R_s^{(1)}(q | k)_{sc} = -R_p^{(1)}(q | k)_{sc}. \quad (\text{B17})$$

Ahora lo que se quiere es generalizar este resultado para el caso de $m > 1$. El

resultado buscado es de la forma [Michel *et al.*, 1992]

$$\psi_s^{(m)}(\mathbf{r} | \omega)_{sc} = (-1)^{(m)} \psi_p^{(m)}(\mathbf{r} | \omega)_{sc}. \quad (\text{B18})$$

Para este fin, es necesario reescribir las ecuaciones (B9) y (B12), siguiendo el procedimiento mostrado en [Michel *et al.*, 1992], como

$$D^{(m)}(\mathbf{r}' | \omega) = 2\chi^{(m-1)}(\mathbf{r}' | \omega) - D^{(m-1)}(\mathbf{r}' | \omega) \quad (\text{B19})$$

y

$$N^{(m)}(\mathbf{r}' | \omega) = 2\varphi^{(m-1)}(\mathbf{r}' | \omega) - N^{(m-1)}(\mathbf{r}' | \omega), \quad (\text{B20})$$

respectivamente.

Sustituyendo en las ecuaciones (B15) y (B16) se obtienen las expresiones correspondientes a la m -ésima contribución de los campos esparcidos

$$\psi_s^{(m)}(\mathbf{r} | \omega)_{sc} = -\frac{2}{4\pi} \int_{-\infty}^{\infty} dS' G_0(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'} \psi_s^{(m-1)}(\mathbf{r}' | \omega)_{sc} - \psi_s^{(m-1)}(\mathbf{r} | \omega)_{sc} \quad (\text{B21})$$

y

$$\psi_p^{(m)}(\mathbf{r} | \omega)_{sc} = \frac{2}{4\pi} \int_{-\infty}^{\infty} dS' \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) \psi_p^{(m-1)}(\mathbf{r}' | \omega)_{sc} - \psi_p^{(m-1)}(\mathbf{r} | \omega)_{sc} \quad (\text{B22})$$

Suponiendo que la relación (B18) es válida para $m - 1$, con $m > 1$, mediante manipu-

laciones algebraicas, se llega a

$$\begin{aligned}
& \psi_s^{(m)}(\mathbf{r} | \omega)_{sc} - (-1)^{(m)} \psi_p^{(m)}(\mathbf{r} | \omega)_{sc} = \frac{2}{4\pi} \int_{-\infty}^{\infty} dS' \\
& \times \left[G_0(\mathbf{r}' | \mathbf{r}) \frac{\partial}{\partial n'} \psi_s^{(m-1)}(\mathbf{r}' | \omega)_{sc} - \frac{\partial}{\partial n'} G_0(\mathbf{r}' | \mathbf{r}) \psi_s^{(m-1)}(\mathbf{r}' | \omega)_{sc} \right] \\
& - 2\psi_s^{(m-1)}(\mathbf{r}' | \omega)_{sc}. \quad (\text{B23})
\end{aligned}$$

Si se aplican el teorema de Green y la condición de radiación de ondas salientes a las funciones $2\psi_s^{(m-1)}(\mathbf{r}' | \omega)_{sc}$ y $G_0(\mathbf{r}' | \mathbf{r})$, el lado derecho de la ecuación (B23) se hace nulo, por lo que el resultado final es la identidad (B18).

Apéndice C

Publicaciones y presentaciones en congresos

A continuación se enlistan las publicaciones generadas como producto del trabajo de investigación desarrollado en esta tesis, así como también los diferentes foros internacionales en los que se han reportado los resultados obtenidos.

Publicaciones

1. D. Macías, G. Olague and E. R. Méndez, "Hybrid Evolution Strategy-Downhill Simplex Algorithm for Inverse Light Scattering Problems," in *Applications of Evolutionary Computing*, S. Cagnoni *et al.*, eds., LNCS **2611**, Springer-Verlag, Berlin (2003), 399-409.
2. D. Macías, E. R. Méndez and V. Ruíz-Cortés. "Inverse Scattering with a wavefront-matching algorithm.", *J. Opt. Soc. Am. A* **19**, 2064, 2002.

3. D. Macías, G. Olague and E. R. Méndez, "Surface Profile Reconstruction from Scattered Intensity Data Using Evolutionary Strategies," in *Applications of Evolutionary Computing*, S. Cagnoni *et al.*, eds., LNCS **2279**, Springer-Verlag, Berlin (2002), 233-244.
4. D. Macías and E. R. Méndez, "Numerical study of an inverse scattering algorithm for metallic and dielectric surfaces", in *Surface Scattering and Diffraction for Advanced Metrology*, A. A. Maradudin and Z. H. Gu, eds., Proc. SPIE. **4447**, 34-41 (2001).
5. E. R. Méndez, D. Macías and V. Ruiz-Cortés, "Inverse scattering of rough surfaces and layered objects", in *Proceedings of the Workshop on Rough Surface Scattering and Related Phenomena*, Yountville, Calif., USA, October 15-18, (2000).
6. D. Macías, E. R. Méndez and V. Ruiz-Cortés, "Numerical study of an inverse scattering algorithm for perfectly conducting one-dimensional surfaces", in *Scattering and Surface Roughness*, A. A. Maradudin and Z. H. Gu, eds., Proc. SPIE. **4100**, 57-64 (2000).

Resúmenes

1. D. Macías, G. Olague and E. R. Méndez, "Surface profile reconstruction using far-field intensity data", Proc. PIERS, Cambridge, Massachusetts, (2002).
2. D. Macías and E. R. Méndez, *Wave-front matching algorithm for inverse scattering*

problems, Proc. IV RIAO/VII OPTILAS , Tandil, Provincia de Buenos Aires, República Argentina, (2001).

Presentaciones en congresos

- EVO-WORKSHOPS2003 , Colchester, Great Britain, 2003.
- PIERS2002 , Boston, Massachusetts, USA. 2002.
- EVO-WORKSHOPS2002 , Kinsale, Republic of Ireland, 2002.
- IV RIAO/VII OPTILAS, Tandil, Provincia de Buenos Aires, República Argentina, 2001.
- SPIE International Symposium on Optical Science and Technology, San Diego, California, 2001.
- SPIE International Symposium on Optical Science and Technology, San Diego, California, 2000.

Bibliografía

Bäck, T., Hammel, U., y Schwefel, H.-P. (1997). Evolutionary computation: Comments on the history and current state. *IEEE Transactions on Evolutionary Computation*, 1(1):3–17.

Bäck, T., Schültz, M., y Khuri, S. (1996). Evolution strategies: An alternative evolution computation method. In Alliot, J. M., Lutton, E., Ronald, E., Schoenhauer, M., y Snyers, D., editors, *Artificial Evolution*, pages 3–20, Berlin. Springer-Verlag.

Bäck, T. y Schwefel, H.-P. (1996). Evolutionary computation: An overview. In Fukuda, T., Furuhashi, T., y Fogel, D. B., editors, *Proc. 1996 IEEE Int'l Conf. Evolutionary Computation (ICEC'96), Nagoya*, pages 20–29, Piscataway NJ. IEEE Press.

Beckmann, P. y Spizzichino, A. (1963). *The Scattering of Electromagnetic Waves from Rough Surfaces*. Pergamon.

Beyer, H.-G. (2001). *The Theory of Evolution Strategies*. Springer-Verlag.

Born, M. y Wolf, E. (1980). *Principles of Optics*. Pergammon, sixth edition.

Chandley, P. (1976). Determination of the autocorrelation function of height on a rough surface from coherent light scattering. *Opt. Quant. Elec.*, 8:329–333.

Chandley, P. y Welford, W. (1975). A re-formulation of some results of p. beckmann for scattering from rough surfaces. *Opt. Quant. Elec.*, 7:393–397.

Cox, I. J., Hamilton, D. K., y Sheppard, C. J. R. (1982). Observation of optical signatures of materials. *Appl. Phys. Lett.*, 41:604–606.

Cox, M. G. (1972). The numerical evaluation of *b*-splines. *J. Inst. Math. Applics.*, 10:134–149.

Dainty, J. (1984). *Laser Speckle and Related Phenomena*. Springer Verlag, Germany.

Dainty, J. C. (1975). Stellar speckle interferometry. In Dainty, J. C., editor, *Laser Speckle and Related Phenomena*. Springer-Verlag.

de Boor, C. (1972). On calculating with *b*-splines. *J. Approx. Theory*, 6:50–62.

de Boor, C. (1978). *A Practical Guide to Splines*. Springer-Verlag.

de Boor, C. (2002). Spline basics. In Farin, G. E., Hoschek, J., y Kim, M., editors, *Handbook of Computer Aided Geometric Design*, pages 141–164. Elsevier.

de Jong, K. A. (1975). *An analysis of the behaviour of a class of genetic adaptive systems*. PhD thesis, University of Michigan.

DeSanto, J. A. (1979). Theoretical methods in ocean acoustics. In DeSanto, J. A., editor, *Ocean Acoustics*. Springer-Verlag.

Destouches, N., Guérin, C., Lequime, M., y H., G. (2001). Determination of the phase of the diffracted field in the optical domain. application to the reconstruction of surface profiles. *Opt. Comm.*, 198:233–239.

Elson, J. M. y Bennet, J. M. (1979). Relation between the angular dependence of scattering and the statistical properties of optical surfaces. *J. Opt. Soc. Am.*, 69:31–47.

Fienup, J. (1978). Reconstruction of an object from the modulus of its fourier transform. *Opt. Lett.*, 3:27–29.

Fogel, L. (1962). Autonomous automata. *Ind. Research*, 4:14–19.

Gerchberg, R. W. y Saxton, W. O. (1971). Phase determination from image and diffraction plane pictures in the electron microscope. *Optik*, 34:275–284.

Gerchberg, R. W. y Saxton, W. O. (1972). A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik*, 35:237–246.

Goodman, J. W. (1968). *Introduction to Fourier Optics*. McGraw-Hill.

Goodman, J. W. (1985). *Statistical Optics*. John Wiley and Sons Inc.

Gu, Z.-H., Lu, J.-Q., y Maradudin, A. A. (1993). Enhanced backscattering from a rough dielectric film on a glass substrate. *J. Opt. Soc. Am. A*, 10(8):1753–1764.

Holland, J. H. (1992). *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence*. MIT Press/Bradford Books, second edition.

- Jackson, J. D. (1965). *Classical Electrodynamics*. J. Wiley.
- Kim, J. H. y Myung, H. (1997). Evolutionary programming techniques for constrained optimization problems. *Transactions on Evolutionary Computation*, 1(2):129–140.
- Kirkpatrick, S., Gelatt, C. D., y Vecchi, M. P. (1983). Optimization by simulated annealing. *Science, Number 4598, 13 May 1983*, 220, 4598:671–680.
- Liszka, E. y McCoy, J. (1982). Scattering at a rough boundary—extensions of the Kirchhoff Approximation. *J. Acoust. Soc. Am*, 71:1093–1100.
- Lu, J.-Q., Maradudin, A. A., y Michel, T. (1991). Enhanced backscattering from a rough dielectric film on a reflecting substrate. *J. Opt. Soc. Am. B*, 8(2):311–318.
- Malyshkin, V., Simeonov, S., McGurn, A. R., y Maradudin, A. A. (1997). Determination of surface profile statistics from electromagnetic scattering data. *Opt. Lett.*, 22:58–60.
- Mandel, L. y Wolf, E. (1995). *Optical Coherence Quantum Optics*. Cambridge University Press.
- Maradudin, A. A., Michel, T., McGurn, A. R., y R., M. E. (1990). Enhanced backscattering of light from a random grating. *Ann. Phys*, 203:255–307.
- McGurn, A. R., Maradudin, A. A., y Celli, V. (1985). Localization effects in the scattering of light from a randomly rough surface. *Phys. Rev. B*, 31:4866–4871.

Méndez, E. R. y O'Donnell, K. A. (1987). Observation of depolarization and backscattering enhancement in light scattering from Gaussian random surfaces. *Opt. Commun.*, 61:91–95.

Metropolis, N., Rosenbluth, A., Rosenbluth, R., Teller, A., y Teller (1953). Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–1092.

Michel, T. R., Knotts, M. E., y O'Donnell, K. A. (1992). Stokes matrix of a one-dimensional perfectly conducting rough surface. *J. Opt. Soc. Am. A*, 9:585–.

Nelder, J. y Mead, R. (1965). A simplex method for function minimization. *Computer Journal*, 7:308–313.

Ogilvy, J. A. (1991). *Theory of Wave Scattering from Random Rough Surfaces*. Hilger.

O'Neill, E. L. y Walther, A. (1963). The question of phase in image formation. *Optica Acta*, 10:33–40.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., y Flannery, B. P. (1992). *Numerical Recipes in Fortran 77*. Cambridge University Press.

Quartel, J. C. y Sheppard, C. J. R. (1996a). Surface reconstruction with an algorithm based on confocal imaging. *J. Modern Optics*, 43:469–486.

Quartel, J. C. y Sheppard, C. J. R. (1996b). Surface reconstruction with an algorithm based on confocal interferometric profiling. *J. Modern Optics*, 43:591–605.

- Roger, A. y Maystre, D. (1980). Inverse scattering method in electromagnetic optics: Application to diffraction gratings. *J. Opt. Soc. Am.*, 70:1483–1495.
- Salomon, R. (1997). Evolutionary algorithms and gradient search: Similarities and differences. *IEEE Transactions on Evolutionary Computation*, 2:45–55.
- Schwefel, H.-P. (1995). *Evolution and Optimum Seeking*. Wiley Interscience.
- Sheppard, C. J. R. (1986). The spatial frequency cut-off in three dimensional imaging. *Optik*, 72(9):131–133.
- Spendley, W., B., H. G., y Himsworth, F. R. (1962). Sequential application of simplex design in optimisation and evolutionising operation. *Technometrics*, 4:441–446.
- Stover, J. C. y Serati, S. A. (1984). Calculation of surface statistics from light scatter. *Opt. Eng.*, 23:406–412.
- Thorsos, E. I. (1988). The validity of Kirchhoff Approximation for rough surface scattering using a Gaussian roughness spectrum. *J. Acoust. Soc. Am.*, 83:78–92.
- Thorsos, E. I. y Jackson, D. R. (1989). The validity of the perturbation approximation for rough surface scattering using a Gaussian roughness spectrum. *J. Acoust. Soc. Am.*, 86(1):261–277.
- Török, P., Varga, P., Laczik, Z., y Booker, G. R. (1995). Electromagnetic diffraction of light focused through a planar interface between materials of mismatched refractive indices: an integral representation. *J. Opt. Soc. Am. A*, 12:325–332.

Walther, A. (1963). The question of phase retrieval in optics. *Optica Acta*, 10:41–49.

Wehrens, R. y Lutgarde, M. B. (2000). Classical and nonclassical optimization methods. In Meyers, R., editor, *Encyclopedia of Analytical Chemistry*. John Wiley and Sons.

Welford, W. T. (1977). Optical estimation of statistics of surface roughness from light scattering measurements. *Opt. Quant. Elec.*, 9:269–287.

Wombell, R. J. y DeSanto, J. (1991a). Reconstruction of rough-surfaces with the kirchhoff approximation. *J. Opt. Soc. Am. A*, 8:1892–1897.

Wombell, R. J. y DeSanto, J. (1991b). The reconstruction of shallow rough-surface profiles from scattered field data. *Inverse Problems*, 7:L7–L12.