



**UNIVERSIDAD NACIONAL
AUTÓNOMA DE MÉXICO**

**FACULTAD DE ESTUDIOS SUPERIORES
ACATLÁN**

**OPTIMIZACIÓN NUMÉRICA DEL MÉTODO
INDIRECTO DE ELEMENTOS DE FRONTERA CON
APLICACIONES A LA EXPLORACIÓN PETROLERA**

TESIS

**QUE PARA OBTENER EL TÍTULO DE
LIC. EN MATEMÁTICAS APLICADAS Y COMPUTACIÓN**

**PRESENTA
VERÓNICA HERNÁNDEZ FLORES**

Asesor: LIC. ANABEL MORENO BALTAZAR.

Noviembre 2004



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

A mis padres: Víctor y Eulalia

Y

a Salvador

Agradecimientos

Quiero agradecer al Instituto Mexicano de Petróleo el apoyo que me brindó, especialmente al Dr. Carlos Ortiz Alemán, por la información y conocimientos que me proporcionó para el desarrollo de esta tesis. A la Universidad Nacional Autónoma de México, sobre todo a la FES ACATLAN, por permitirme estudiar la carrera de Matemáticas Aplicadas y Computación. A mi asesora de tesis Lic. Anabel Moreno Baltazar por la gran ayuda que me brindó para realizar este trabajo.

A mis padres Víctor y Eulalia por el gran cariño con que siempre me han impulsado para seguir adelante y porque siempre han estado conmigo apoyándome en lo que he necesitado. A mis hermanos, Víctor Hugo, Violeta, Vianey y Vanessa, por su comprensión y paciencia, y a mis sobrinos Daniela y Erick, por llenar de alegría mi vida.

A Salvador por su amor y comprensión, por compartir conmigo los momentos buenos y malos y por estar a mi lado para darme ánimos.

A mis amigos Patricia y Ángel por sus consejos y ayuda incondicional.

ÍNDICE

AGRADECIMIENTOS.....	i
PRÓLOGO.....	ii
1) INTRODUCCIÓN.....	1
1.1) CONCEPTOS BÁSICOS	1
1.1.1) <i>Antecedentes históricos</i>	1
1.1.2) <i>Exploración sismológica</i>	2
1.2) TEORÍA DE LAS ONDAS SÍSMICAS	3
1.2.1) <i>Elasticidad</i>	3
1.3) LEY DE HOOKE	4
1.4) ECUACIÓN DE NAVIER.....	6
1.5) PROPAGACIÓN DE ONDA.....	6
1.5.1) <i>Soluciones de la ecuación de onda</i>	6
1.5.2) <i>Soluciones de onda plana</i>	7
1.6) PRINCIPIO DE HUYGENS.....	7
1.7) ONDAS P Y ONDAS S.....	8
1.8) CONDICIONES DE FRONTERA	11
1.9) ONDAS DE RAYLEIGH Y ONDAS DE LOVE.....	12
1.10) LEY DE SNELL	13
2) ANTECEDENTES	15
2.1) INTRODUCCIÓN A LOS EFECTOS LOCALES Y A LA DIFRACCIÓN DE ONDAS ELÁSTICAS	15
2.2) PROPAGACIÓN DE ONDAS EN UN MEDIO ELÁSTICO	16
2.3) DIFRACCIÓN DE ONDAS ELÁSTICAS POR IRREGULARIDADES SUPERFICIALES... 18	
2.4) SOLUCIONES NUMÉRICAS EN LA PROPAGACIÓN DE ONDAS ELÁSTICAS.....	19
3) MÉTODO INDIRECTO DE ELEMENTOS DE FRONTERA.....	21
3.1) REPRESENTACIÓN INTEGRAL USANDO FUENTES EN LAS FRONTERAS	22
3.2) DIFRACCIÓN DE ONDAS DEBIDO A ESTRUCTURAS TOPOGRÁFICAS SUPERFICIALES	23

3.3) DISCRETIZACIÓN	25
4) FORMULACIÓN IBEM/GRADIENTE CONJUGADO	28
4.1) DESCRIPCIÓN DEL ALGORITMO DEL GRADIENTE CONJUGADO	31
4.2) ALMACENAMIENTO DE MATRICES DISPERSAS	34
4.2.1) Almacenamiento de listas de enteros	36
4.2.2) Representación y almacenamiento de grafos (diagramas)	37
4.2.3) Almacenamiento diagonal de matrices de Banda	38
4.2.4) Almacenamiento "envolvente" de matrices simétricas	39
4.2.5) Esquemas de almacenamiento dispersos ligados	40
4.2.6) El formato row-wise disperso	42
4.2.7) Representaciones ordenadas y desordenadas	44
4.2.8) Compresión de Sherman	46
4.2.9) Almacenamiento de matrices de bloque-particiones	46
4.3) TECNOLOGÍA DE MATRICES RALAS	47
4.4) MÉTODOS DIRECTOS E ITERATIVOS	48
4.5) ESQUEMAS DE TRUNCAMIENTO MATRICIAL	49
4.5.1) Algunas definiciones y propiedades	52
4.5.2) Descomposición LU	52
4.5.3) Eliminación de Gauss	57
4.5.4) Factorización de Cholesky de una matriz definida positiva simétrica ..	59
4.5.5) Implementación práctica de la factorización de Cholesky	62
5) RESPUESTA SÍSMICA 3D DE LA ESTRUCTURA DEL SÓTANO PROFUNDA DE LA CUENCA DE GRANADA (SUDESTE DE ESPAÑA)	64
5.1) RESUMEN	64
5.2) LA CUENCA DE GRANADA	66
5.3) MÉTODO Y DISCRETIZACIÓN	67
6) RESULTADOS	72
6.1) ANÁLISIS DE PROPAGACIÓN DE ONDA	77
6.2) CONCLUSIONES	82

BIBLIOGRAFÍA	84
GLOSARIO.....	86

Prólogo

El Método Indirecto de Elementos de Frontera (IBEM, por sus siglas en inglés), se utiliza para resolver problemas de configuraciones complejas fracturadas naturalmente. Se discretiza utilizando las funciones de Green. Da origen a grandes sistemas de ecuaciones, por lo cual se propone el empleo de criterios de truncamiento matricial en combinación con métodos de solución para matrices ralas o porosas.

El objetivo de esta tesis es desarrollar herramientas numéricas para la caracterización de yacimientos naturalmente fracturados basadas en el empleo del Método Indirecto de Elementos de Frontera (IBEM), así como simular numéricamente la propagación de ondas sísmicas en medios fracturados y obtener modelos teóricos que faciliten la interpretación de registros sísmicos multicomponente.

La hipótesis es que *la optimización numérica del Método Indirecto de Elementos de Frontera es aplicable al estudio de la respuesta sísmica de estructuras geológicas complejas, bi y tridimensionales mediante la Tecnología de matrices ralas. Estos métodos están caracterizados por el truncamiento de la matriz de coeficientes asociada, lo cual reduce la complejidad del problema.*

Este trabajo consta de 6 capítulos, en los dos primeros se habla de conceptos básicos de Sismología, que es el área para la cual se aplica este método, también se explican métodos que se pueden utilizar para resolver este tipo de problemas, sus ventajas y desventajas; en el capítulo tercero se explica el método IBEM, y en los dos últimos capítulos, se muestra una aplicación real del método y los resultados obtenidos de este estudio.

El capítulo 1, es una introducción a los conceptos básicos de Geología, teoría de ondas sísmicas, Propagación de onda, ondas P, S, de Rayleigh y de Love, así como antecedentes históricos para adentrarnos más en el tema de la Sismología. En el capítulo 2,

se hace una breve descripción de los diferentes métodos numéricos que se han aplicado al estudio de la propagación de ondas sísmicas en configuraciones irregulares tridimensionales. En el capítulo 3, se introduce brevemente el método indirecto de elementos de frontera. Se presenta la formulación integral y aspectos muy importantes como el planteamiento de las condiciones de frontera y su posterior discretización. Se discuten las principales ventajas y dificultades de este esquema numérico, el cual da como resultado un sistema de ecuaciones lineal, que dependiendo de las frecuencias, puede ser muy grande lo cual presenta una dificultad para resolverlo, por eso se propone el truncamiento de la matriz por medio de la llamada Tecnología de matrices ralas o porosas.. El capítulo 4 es el estudio de la aplicación de la Tecnología de matrices ralas para el almacenamiento y solución eficiente del sistema de ecuaciones lineales que surge de la formulación del método IBEM. Se presentan varios esquemas de almacenamiento así como esquemas de truncamiento matricial.

Capítulo 5, se introduce la aplicación del método optimizado al estudio de la propagación de ondas sísmicas, mediante un ejemplo real de la respuesta sísmica de la estructura de la Cuenca de Granada. Se intenta mejorar nuestra comprensión acerca de la propagación de ondas sísmicas en medios fracturados, que son de interés para el desarrollo de nuevas técnicas de procesamiento e interpretación de registros sísmicos multicomponente.

En el capítulo 6, se presentan los resultados conseguidos y se hace una revisión del impacto de éstos en el desarrollo de esta línea de investigación.

1) Introducción

1.1) Conceptos Básicos

1.1.1) Antecedentes históricos

La sismología es el estudio de la generación, propagación y registro de ondas elásticas en la Tierra (y otros cuerpos celestes) y de las fuentes que las producen. La sismología se ocupa del estudio de terremotos, sismos, temblores y de otras vibraciones producidas natural o artificialmente en la Tierra. El estudio de terremotos incluye su detección y la determinación de su localidad, de su magnitud, su energía y de los movimientos tectónicos causantes. Otros objetos de estudio son las vibraciones relacionadas con el volcanismo y aquellas generadas por los océanos, por el viento y ondas atmosféricas. Además la sismología se encarga del estudio de la estructura interna de la Tierra.

La exploración sismológica, es una disciplina relativamente joven, se inició por el año de 1923, aunque podría remontarse a la introducción de la ley de Hooke (1660), indicando una proporcionalidad entre presión y esfuerzo. Con el desarrollo de las ecuaciones para la teoría de la elasticidad por Navier y Cauchy en 1821-1822, la comprensión de la conducta de los materiales sólidos evolucionó rápidamente. En 1830, Poisson utilizó las ecuaciones de movimiento y las leyes constitutivas elásticas para mostrar que existen dos tipos fundamentales de ondas que se propagan a través del interior de los sólidos homogéneos: la onda P y la onda S. Rayleigh (1887) demostró la existencia de soluciones adicionales para las ecuaciones elásticas de movimiento de cuerpos con superficies libres. Estas son las ondas de Rayleigh, que involucran movimientos de onda reducidos y que se propagan a través de la superficie del cuerpo. En 1911, se descubrieron las ondas de Love, que igual que las ondas de Rayleigh, son ondas superficiales y resultan de la interacción de las ondas P y S con las condiciones frontera del cuerpo.

1.1.2) Exploración sismológica

Para la exploración sismológica, es importante utilizar ondas elásticas generadas artificialmente para localizar depósitos minerales (como hidrocarburos, yacimientos minerales, agua, depósitos geotérmicos, etc.), zonas arqueológicas y para obtener información geológica para la ingeniería. Con la exploración sismológica se obtienen datos que, cuando son utilizados junto con otros datos geofísicos, (como los datos de pozos exploratorios y los datos geológicos), y con conceptos de física y geología, pueden proporcionar información sobre la estructura y distribución de los diferentes tipos de rocas.

El trabajo sísmico es muy importante en la exploración petrolera, porque casi todas las empresas petroleras confían en la interpretación sísmica para seleccionar los sitios en donde perforarán los pozos petroleros.

La sismología de exploración es una rama de la sismología de terremotos. Cuando ocurre un terremoto, la tierra se fractura, esta ruptura genera ondas sísmicas que se desplazan alejándose de la superficie de la fractura. Estas ondas se registran mediante el uso de sismógrafos. Los sismólogos utilizan los datos para deducir información sobre la naturaleza de las rocas a través de las cuales viajaron las ondas del terremoto.

Las ondas sísmicas son generadas sintéticamente con explosivos y otras fuentes de energía y son registradas por medio de geófonos o sismógrafos que detectan el movimiento de la tierra resultante.

Usualmente los datos se registran en forma digital de modo que se puedan utilizar procesamientos en computadora para extraer información o presentarla visualmente de forma que se efectúe rápidamente una interpretación geológica.

La información de las trayectorias corresponden a dos categorías principales: trayectorias de ondas precursoras o refractadas, en las cuales la parte principal de la trayectoria está a lo largo de la interfase entre dos capas de roca y, por lo tanto es aproximadamente horizontal, y trayectorias reflejadas en las que la onda se desplaza inicialmente hacia abajo y en algún punto se refleja, regresando a la superficie, siendo la trayectoria total prácticamente vertical.

El objetivo de la exploración sísmica consiste en deducir información acerca de las rocas, especialmente de las posiciones de las capas, a partir de los tiempos de llegada observados, de las variaciones en la amplitud, frecuencia y forma de onda.

El método sísmico tiene predominio sobre los demás métodos geofísicos por varios factores, el mas importante es su precisión, su alta resolución y la gran penetración que es posible alcanzar mediante el.

1.2) Teoría de las Ondas Sísmicas

1.2.1) Elasticidad

En el método sísmico se utiliza la propagación de ondas a través de la Tierra. Esta propagación depende de las propiedades elásticas de las rocas.

A diario, la Tierra nos demuestra su condición de medio "elástico". Los terremotos son un ejemplo de la capacidad de nuestro planeta para propagar ondas elásticas (ondas sísmicas).

La Teoría de la Elasticidad es fundamental en Sismología, ya que son las propiedades elásticas de las rocas las que permiten la propagación de ondas en La Tierra.

Si a un cuerpo sólido se le aplica una fuerza externa, este cambia su forma y tamaño, pero existen fuerzas internas que se resisten a este cambio y el cuerpo tiende a regresar a su condición original, cuando se eliminan las fuerzas externas.

A esta propiedad de resistir cambios de tamaño o forma y de regresar a la condición no deformada cuando se eliminan las fuerzas externas se llama *elasticidad*. Un cuerpo perfectamente elástico es aquel que se recupera completamente después de ser deformado.

En la teoría de la elasticidad, las relaciones entre fuerzas aplicadas y las deformaciones se expresan más convenientemente con los términos de esfuerzo y deformación.

La deformación es un cambio de forma o tamaño, generalmente es proporcional al esfuerzo que la produce, como se enuncia en la Ley de Hooke.

El esfuerzo se define como la fuerza por área unitaria. Cuando una fuerza se aplica a un cuerpo, el esfuerzo es la relación de la fuerza con el área que se aplica esa fuerza. Si la fuerza es perpendicular, se dice que el esfuerzo es un esfuerzo normal (o presión). Cuando la fuerza es tangencial al elemento del área, el esfuerzo es un esfuerzo cortante.

1.3) Ley de Hooke

Para calcular las deformaciones cuando se conocen los esfuerzos se debe conocer la relación entre esfuerzo y deformación. Cuando las deformaciones son pequeñas, esta relación está dada por la Ley de Hooke que establece que una deformación dada es directamente proporcional al esfuerzo que la produce.

Esta ley es para conocer el comportamiento de materiales idealmente elásticos. Esto significa que las deformaciones sufridas por estos materiales son siempre proporcionales a la magnitud de los esfuerzos que las provoquen.

La ley de Hooke conduce a relaciones complicadas, pero cuando el medio es isotrópico, es decir, cuando las propiedades no dependen de la dirección, suele escribirse en términos de las dos constantes de Lamé (λ y μ), como se muestra en la siguiente ecuación:

$$\sigma_{ij} = \lambda \cdot e_{kk} \delta_{ij} + 2\mu \cdot e_{ij} \quad (1.1)$$

Donde σ_{ij} es el tensor de esfuerzos, e es el tensor de deformaciones, μ es una medida de la resistencia a la deformación cortante y se denomina como módulo de rigidez, λ no tiene una representación física definida, sin embargo, tiene relación con otros módulos elásticos. Aunque la ley tiene amplia aplicación no sirve para esfuerzos grandes, cuando el esfuerzo aumenta más allá de un límite elástico, la Ley de Hooke ya no se cumple y las deformaciones aumentan más rápidamente. Las deformaciones resultantes de esfuerzos que exceden este límite no desaparecen por completo cuando se eliminan los esfuerzos. Con mayor esfuerzo, se puede alcanzar un punto de deformación plástica en el que empieza el flujo plástico y la deformación plástica puede producir un decremento de la deformación. Algunos materiales no pasan por una fase de flujo plástico sino que se rompen antes (fig.1.1 a).

De igual manera, algunos materiales tienen también un comportamiento al esfuerzo que depende del tiempo (fig. 1.1 b). Cuando se sujetan a esfuerzo fijo, estos materiales ceden hasta que eventualmente se rompen. La deformación plástica no desaparece si no se elimina el esfuerzo.

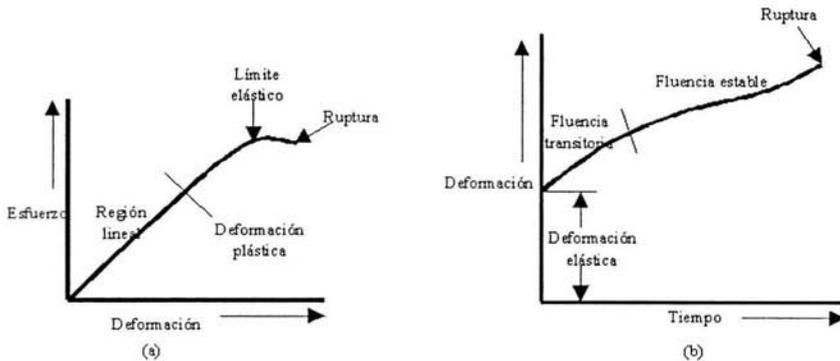


Figura 1.1 Relaciones de esfuerzo-deformación-tiempo. a) Esfuerzo contra deformación; b) Deformación contra tiempo.

1.4) Ecuación de Navier

Esta es la ecuación del movimiento particularizada para un medio sólido, infinito, perfectamente elástico, homogéneo e isótropo, donde los esfuerzos y deformaciones son infinitesimales. La ecuación se deriva de la segunda ley de Newton.

Se acostumbra escribir la ecuación empleando el operador nabla:

$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right)$, siendo esta la forma más conocida de la ecuación de Navier.

$$\mu^2 \bar{u} + (\lambda + \mu) \nabla(\nabla \cdot \bar{u}) + X = \rho \frac{\partial^2 \bar{u}}{\partial t^2} \quad (1.2)$$

Donde ∇ es el operador gradiente, ρ es la densidad de masa, u es el campo de desplazamientos y $\nabla^2 \bar{u} = \nabla(\nabla \cdot \bar{u}) - \nabla X(\nabla X \bar{u})$.

1.5) Propagación de Onda

1.5.1) Soluciones de la ecuación de onda

Manipulando la ecuación de Navier, se pueden obtener diferentes formas de la ecuación de onda, pero su forma mas general es la siguiente:

$$\frac{1}{v^2} \frac{\partial^2 \Psi}{\partial t^2} = \nabla^2 \Psi \quad (1.3)$$

donde v es una constante. Cualquier función $\Psi = f(x - vt)$ cuyas primeras dos derivadas sean continuas, es una solución para la ecuación 1.3. Esta solución Ψ es general para el caso unidimensional y proporciona una cantidad infinita de soluciones.

1.5.2) Soluciones de onda plana

Una onda de cuerpo se define como una “perturbación” que viaja a través del medio. En esta notación, una perturbación Ψ es un cambio de volumen cuando $\Psi = \Delta$ y una rotación cuando $\Psi = 0$.

En la ecuación:

$$\Psi = f(x - vt) \quad (1.4)$$

se ve claramente que la perturbación viaja a lo largo del eje x . En la siguiente figura se muestra como viaja con una velocidad igual a la cantidad v .

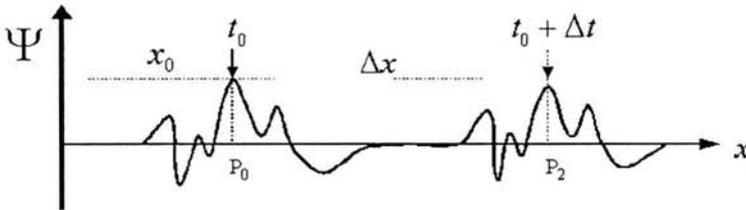


Figura 1.2 Ilustración de la velocidad de una onda.

Onda plana: cuando el valor de Ψ es independiente de y y z , la perturbación debe ser la misma en todas partes en un plano perpendicular al eje x .

La cantidad $x - vt$ se conoce como la fase. Las superficies sobre las cuales el movimiento de la onda es el mismo se conoce como frente de onda.

1.6) Principio de Huygens

Este principio es importante para comprender el viaje de la onda y con frecuencia es útil para dibujar posiciones sucesivas de frentes de onda. El principio de Huygens, establece que cada punto sobre un frente de onda se puede contemplar como una nueva fuente de onda.

Este principio ayuda a explicar como se transmiten las perturbaciones sísmicas en la Tierra a partir de la respectiva información.

En la difracción se utiliza este principio para construir frentes de onda de difracción, aunque este estudio implica el uso de matemáticas complejas, el planteamiento mediante el principio de Huygens no es matemático.

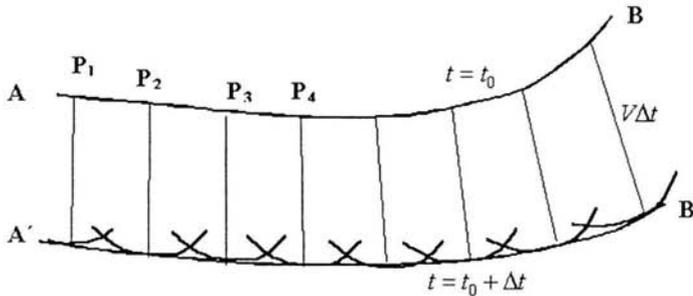


Figura 1.3 Uso del principio de Huygens para localizar nuevos frentes de onda

En figura 1.3, AB es el frente de onda en el tiempo t_0 y se desea encontrar el frente de onda en un tiempo posterior $t_0 + \Delta t$. Durante el intervalo Δt , la onda avanzará una distancia $V\Delta t$, siendo V la velocidad (que puede variar de un punto a otro). Se seleccionan puntos sobre el frente de onda, P_1 , P_2 , P_3 , y así sucesivamente, desde los cuales se dibujan arcos de radio $V\Delta t$.

1.7) Ondas P y ondas S

Al romperse la roca se generan ondas que se propagan a través de la Tierra, tanto en su interior como por su superficie. Como se mencionó anteriormente, básicamente existen dos tipos de ondas llamadas también Ondas de volumen, que se propagan en un medio isotrópico homogéneo, uno correspondiente a cambios en la dilatación (Δ), el otro a cambios en uno o más componentes de una rotación dada.

El primer tipo se conoce como onda P (primaria), este nombre se debe al hecho de que este tipo es el primer evento en el registro de un terremoto. Consiste en la transmisión de compresiones y refracciones de la roca, de forma similar a la propagación del sonido (fig.1.4 a).

El segundo tipo se denomina ondas S (secundaria), puesto que usualmente es el segundo evento observado en registros de terremotos. Consiste en la propagación de ondas de cizalla, donde las partículas se mueven en dirección perpendicular a la dirección de propagación de la perturbación (fig.1.4 b).

Estos dos tipos de ondas se pueden propagar por el interior de la Tierra.

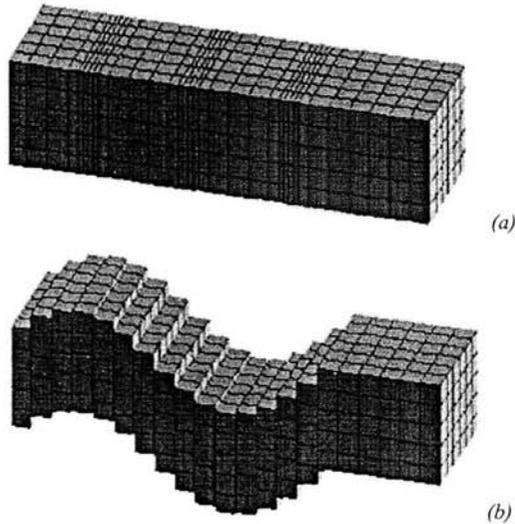


Figura 1. 4, a) Onda P; b) Onda S

La onda P tiene una velocidad α y la onda S una velocidad β en donde

$$\alpha = \{(\lambda + 2\mu) / \rho\}^{\frac{1}{2}} \quad \text{onda P} \quad (1.5)$$

$$\beta = (\mu / \rho)^{\frac{1}{2}} \quad \text{onda S} \quad (1.6)$$

Las ondas P son ampliamente utilizadas en la exploración sísmica, ya que se propagan con mayor velocidad (de ahí su nombre, *primarias*), presentando además la característica de poder propagarse por cualquier tipo de material, sea sólido o líquido. Las ondas S viajan a una velocidad algo menor (*secundarias*) y no se propagan por masas líquidas.

Debido a la diferencia en la velocidad de cada tipo de onda, cuando sentimos un terremoto las primeras sacudidas son debidas a las ondas P, siendo las siguientes, las ondas S y por último las ondas superficiales. La diferente velocidad de cada tipo de onda es, además, la propiedad que se utiliza para determinar la localización del foco del terremoto.

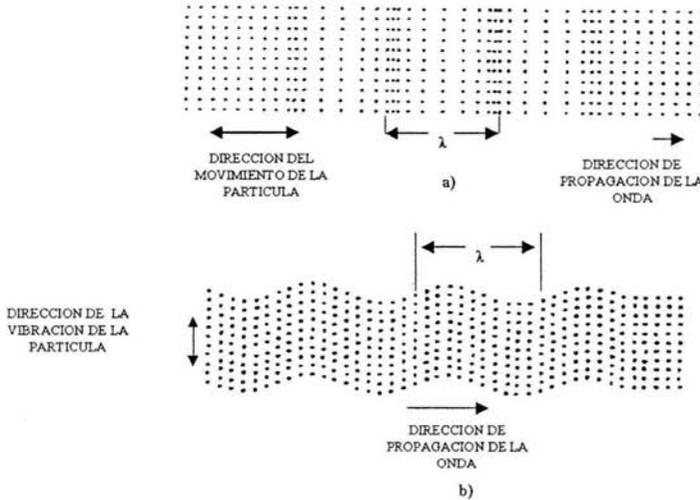


Figura 1. 5 Movimiento durante el paso de una onda plana, a) onda P, b) onda S.

Las ondas S tienen dos grados de libertad, a diferencia de las ondas P que tienen solo uno. El movimiento de las ondas S se descompone usualmente en una componente paralela llamada SH y otra perpendicular a la superficie del suelo conocida como SV.

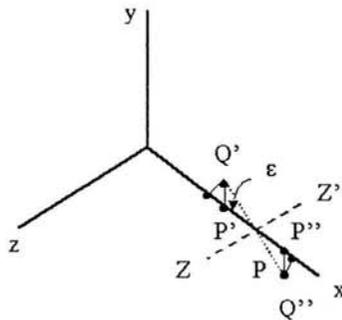


Figura 1. 6 Rotación del medio durante el paso de una onda S.

Cuando la onda llega a P, hace que el medio en la proximidad de P gire respecto al eje $Z'Z''$ (paralelo al eje z). Así a medida que la onda viaja a lo largo del eje x , el medio se desplaza transversalmente a la dirección de la propagación.

1.8) Condiciones de Frontera

Cuando una onda llega a una superficie que separa dos medios que tienen diferentes propiedades elásticas, da origen a las ondas reflejadas y refractadas. Se pueden encontrar las relaciones entre diversas ondas a partir de las relaciones entre los esfuerzos y las deformaciones a cada lado de la interfase. En la frontera entre los dos medios, los esfuerzos y las deformaciones deben ser continuos.

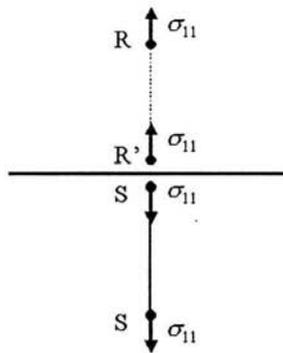


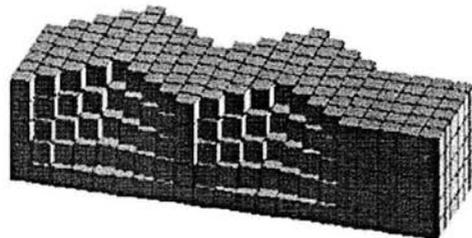
Figura 1. 7 Continuidad del esfuerzo normal

Dos puntos cercanos, R y S, que están en lados opuestos de la frontera como se ve en la figura 1.7, por lo general tendrán diferentes valores de esfuerzo normal. Esta diferencia produce una fuerza neta que acelera la capa entre ellos. Sin embargo, si se eligen puntos cada vez más cercanos entre sí, los valores de esfuerzo se deben aproximar entre sí y el límite cuando los dos puntos coinciden en la frontera, estos dos esfuerzos deben ser iguales.

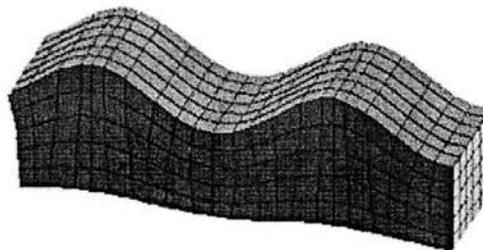
1.9) Ondas de Rayleigh y Ondas de Love.

En un medio isotrópico, homogéneo, infinito, solo existen ondas P y S. Pero siempre que hay una superficie que separe medios de diferentes propiedades elásticas, pueden existir un tercer tipo de ondas llamadas ondas superficiales, debido a que solo se propagan por las capas más superficiales de la Tierra, decreciendo su amplitud con la profundidad. Dentro de este tipo de ondas se pueden diferenciar dos modalidades, denominadas ondas de Rayleigh y ondas de Love en honor a los científicos que demostraron teóricamente su existencia.

Las ondas de Rayleigh se forman en la superficie de la Tierra y hacen que las partículas se desplacen según una trayectoria elíptica retrógrada (fig. 1.8 b). En cambio las ondas de Love se originan en la interfase de dos medios con propiedades mecánicas diferentes; en este caso el movimiento de las partículas es perpendicular a la dirección de propagación de la perturbación, similar a las ondas S, pero solo ocurre en el plano de la superficie terrestre (fig. 1.8 a).



(a)



(b)

Figura 1. 8. a) Ondas de Love; b) Ondas de Rayleigh

La onda superficial más importante en la exploración sísmológica es la onda Rayleigh que se propaga a lo largo de la superficie de un sólido.

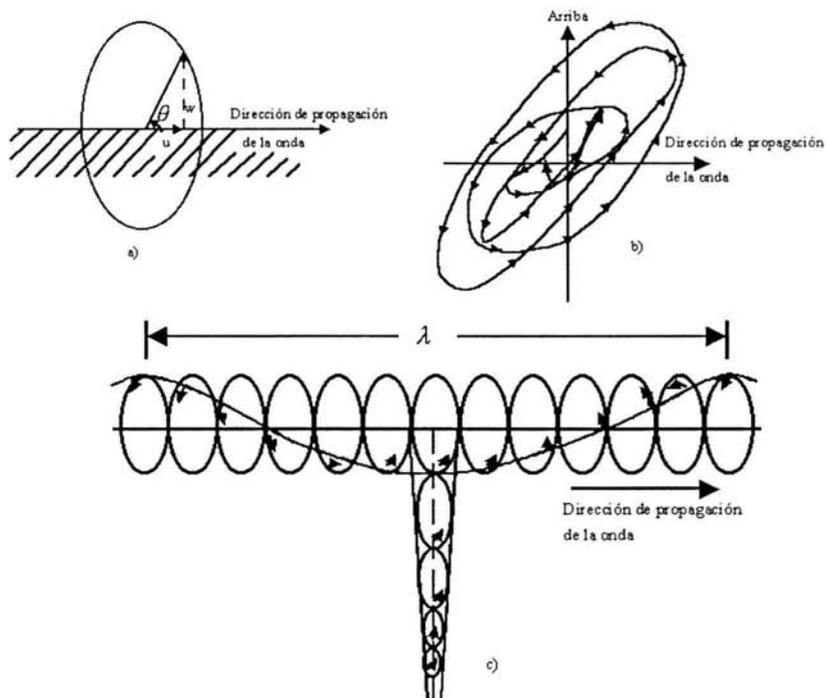


Figura 1. 9 Ondas de Rayleigh. a) Movimiento predicho de una partícula sobre la superficie de un sólido semiinfinito; b) movimiento real de una partícula sobre la superficie de la Tierra; c) sección transversal que muestra el movimiento de partículas sobre la superficie y a profundidad para un sólido semiinfinito.

Las ondas de Rayleigh son ondas de baja velocidad y baja frecuencia con un espectro sin picos pronunciados, y por lo tanto, poseen una amplia gama de longitudes de onda.

1.10) Ley de Snell

Cuando una onda incide en una interfase que separa dos medios elásticos, una parte de su energía se refleja, es decir, se queda en el medio original del cual provino la onda

incidente (ondas reflejadas), y el equilibrio, o sea la otra parte de la energía, atraviesa al otro medio (ondas refractadas).

La Ley de Snell, describe el comportamiento y la relación que guardan las ondas que intervienen.

Las leyes de reflexión y refracción combinadas en un solo planteamiento, nos dan la forma generalizada de Snell; en una interfase la cantidad $p = (\text{sen } \theta_i) V_i$ tiene el mismo valor para las ondas incidentes, reflejadas y refractadas. La cantidad p se llama parámetro de la trayectoria del rayo.

Cuando el medio consta de cierto número de capas paralelas, la ley de Snell requiere que la cantidad p tenga el mismo valor en todas partes para todos los rayos reflejados y refractados resultantes de un rayo inicial dado.

La ley de Snell, es útil para determinar las trayectorias de los rayos y los tiempos de llegada, y para derivar la posición del reflector a partir de los tiempos de arribo observados, pero no proporciona información acerca de las amplitudes de las ondas reflejadas y transmitidas.

2) Antecedentes

En el capítulo anterior, se vieron conceptos básicos de Sismología, como las ecuaciones de Navier, las ondas P y S, que se retomarán en este capítulo. Serán de gran utilidad para comprender más a fondo la propagación de onda en medios elásticos, las técnicas que mas se utilizan para el estudio de difracción de ondas elásticas, así como sus ventajas y desventajas.

2.1) Introducción a los efectos locales y a la difracción de ondas elásticas

La ingeniería sísmica estudia el movimiento del suelo. Los movimientos sísmicos dependen principalmente de tres factores:

- 1) Las características de ruptura de la falla (fuente)
- 2) Las trayectorias que realizan las ondas sísmicas emitidas (camino)
- 3) Las condiciones locales (sitio)

Cuando la energía liberada se propaga, produce los fenómenos de reflexión, refracción y difracción. Estos fenómenos dependen de las condiciones específicas de cada lugar. Las condiciones locales son muy importantes en la respuesta sísmica de las grandes estructuras de interés social como, puentes e instalaciones energéticas.

A lo largo de la historia en todo el mundo han ocurrido terremotos de gran intensidad, que por el daño que han ocasionado son llamados “destructores”. Uno de los mas importantes ha sido el sismo de México en 1985 de magnitud $M_s = 8.1$. El epicentro fue en Michoacán, en donde los daños fueron leves. Sin embargo, el mayor daño causado se localizó en la Cd. de México que se encuentra a 400 Km. de donde se originó el sismo. Estos efectos se han presentado también en otros grandes terremotos en otras partes del mundo.

Estos fenómenos han llevado a los investigadores a tratar de entender la respuesta sísmica local del valle aluvial en donde se asienta la Cd. de México.

2.2) Propagación de ondas en un medio Elástico

Como se mencionó en el capítulo 1, la ecuación de Navier describe la propagación de ondas elásticas en un medio continuo, elástico homogéneo e isotrópico. Para un campo vectorial cualquiera \vec{A} se puede aplicar el teorema de Helmholtz:

$$\nabla^2 \vec{A} = \nabla(\nabla \cdot \vec{A}) + \nabla_x(\nabla_x \vec{A}) \quad (2.2)$$

Considerando que u cumple dicho teorema, la ecuación de Navier se puede expresar como:

$$(\lambda + 2\mu)\nabla(\nabla \cdot \bar{u}) + \mu\nabla_x(\nabla_x \bar{u}) = \rho \frac{\partial^2 \bar{u}}{\partial t^2} \quad (2.3)$$

Si el desplazamiento de u ocurre sin cambio de volumen, entonces su divergencia es cero y la ecuación queda como:

$$\nabla^2 \bar{u} = \frac{1}{\beta^2} \frac{\partial^2 \bar{u}}{\partial t^2} \quad (2.4)$$

$$\nabla \cdot \bar{u} = 0$$

Donde $\beta = (\mu / \rho)^{1/2}$, es la velocidad de propagación de las ondas equivolumétricas o cortantes. Estas se llaman ondas S y se dividen en SH y SV, como se explicó anteriormente. Las ondas polarizadas en el plano horizontal perpendicular a la dirección vertical son las ondas SH y las SV son las paralelas al plano vertical, como se muestra en la siguiente figura.

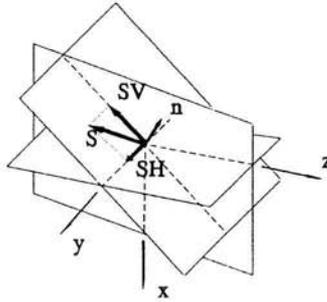


Figura 2.1 Polarización de la onda S

Si el desplazamiento es sin rotación de partículas entonces la rotación del vector de desplazamiento es nulo y la ecuación de Navier queda como:

$$\nabla^2 \bar{u} = \frac{1}{\alpha^2} \frac{\partial^2 \bar{u}}{\partial t^2} \quad (2.5)$$

$$\nabla_x \bar{u} = 0$$

donde $\alpha = [(\lambda + 2\mu) / \rho]^{1/2}$ es la velocidad de propagación de ondas compresionales, o sea ondas P.

Las ondas P y S se presentan en los registros del movimiento del suelo en un terremoto.

2.3) Difracción de ondas elásticas por irregularidades superficiales

La difracción se presenta en un movimiento ondulatorio, cuando en el trayecto de la onda se encuentra con obstáculos o aberturas con dimensiones comparables con sus longitudes de onda. De acuerdo con el principio de Huygens. Los bordes del obstáculo actúan como fuentes de ondas secundarias propagándose en todas las direcciones. Así las ondas difractadas llegan a las zonas que son inaccesibles considerando la propagación rectilínea de la onda incidente.

El desplazamiento hecho por las ondas incidentes y difractadas deben satisfacer las ecuaciones de Navier y las condiciones de contorno en la superficie

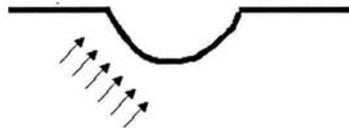


Figura 2.2 Irregularidad superficial e incidencia de ondas elásticas

Considerando un semiespacio E con una inclusión elástica R bajo la incidencia de las ondas elásticas.

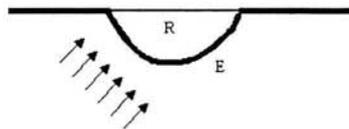


Figura 2.3 Semiespacio E e inclusión R bajo la incidencia de ondas elásticas

Se considera el movimiento del semiespacio como la superposición de las ondas difractadas y las de campo libre:

$$u_i^E = u_i^0 + u_i^d \quad (2.6)$$

donde u_i^0 es el desplazamiento de campo libre en la dirección i , es decir, la solución en el semiespacio E en ausencia de la irregularidad.

Las condiciones de contorno implican que sobre la superficie libre del semiespacio E las tracciones sean nulas, es decir:

$$t_i^0 + t_i^d = 0 \quad (2.7)$$

donde t_i^0 y t_i^d son las i -ésimas componentes de las tracciones correspondientes al campo libre difractado, respectivamente. Sobre la superficie libre de la inclusión R debe verificarse que las tracciones también sean nulas

$$t_i^r = 0 \quad (2.8)$$

donde t_i^r son las tracciones en la dirección i debida a las ondas refractadas.

Todas estas condiciones de continuidad y desplazamientos en la frontera común de E y R implican

$$\begin{aligned} u_i^0 + u_i^d &= u_i^r \\ t_i^0 + t_i^d &= t_i^r \end{aligned} \quad (2.9)$$

2.4) Soluciones numéricas en la propagación de ondas elásticas

Las técnicas que más se utilizan en el estudio de difracción de ondas elásticas son los siguientes:

Métodos de dominio

Estos métodos, también llamados directos, se dividen en tres:

- a) Método de diferencias finitas.
- b) Método pseudoespectral.
- c) Método de elementos finitos.

Son muy eficientes porque permiten el tratamiento de estructuras muy complejas. La desventaja es que se llevan mucho tiempo y memoria en su proceso.

Teoría de rayos y Haces Gaussianos

Este método se utiliza para frecuencias muy grandes, básicamente se divide en dos procesos:

- a) encontrar los caminos que recorre la señal, es decir, los rayos y
- b) encontrar la intensidad de la señal que se propaga a lo largo de cada rayo

El costo de este método es bajo en comparación de otros, pero presenta varias limitantes que hacen que el método no tome en cuenta datos importantes.

Métodos de contorno

En este método se evita la introducción de fronteras ficticias en el medio, como en los métodos de dominio, reduciendo así las dimensiones del problema. Esto disminuye en gran medida el tiempo de cómputo. Este método tiene tres técnicas basadas en:

- a) el uso de sistemas completos de soluciones
- b) la representación del número de onda discreto
- c) el uso de ecuaciones integrales

3) Método Indirecto de Elementos de Frontera

Ahora, con lo estudiado en los capítulos 1 y 2, podemos comprender mejor el Método Indirecto de Elementos de Frontera, ya que involucra todos los conceptos de Geología y Sismología antes vistos, como condiciones de frontera, la ley de Hooke, etc. También, para que sea más fácil la interpretación de las fórmulas y lo que representan.

Este método se utiliza para calcular la respuesta sísmica de varios tipos de estructuras geológicas. El método se basa en la representación integral de los campos de ondas elásticas usando fuentes de capa simple en la frontera. Es indirecto porque se obtienen las densidades de fuerza en el contorno del dominio.

Considerando las condiciones de contorno para esfuerzos y desplazamientos, se obtiene un sistema de ecuaciones integrales para las fuentes de la frontera. Después se discretiza basándose en las funciones de Green para desplazamientos y tracciones de problemas bi y tridimensionales.

3.1) Representación Integral usando fuentes en las fronteras

Considérese un espacio tridimensional euclídeo y una superficie continua S . Las coordenadas $x_1(x)$ y $x_2(y)$ forman el plano de la superficie libre, $x_3(z)$ representa la profundidad; $u(u_1)$, $v(u_2)$ y $w(u_3)$ son los desplazamientos producidos en cualquier punto en las direcciones x (x_1), y (x_2) y z (x_3) respectivamente. Si un material elástico ocupa este espacio, el campo de desplazamientos se pueden expresar por medio de la integral

$$u_i(x) = \int_S \phi_j(\xi) G_{ij}(x, \xi) dS_\xi \quad (3.1)$$

donde $u_i(x)$ = componente i -ésima del desplazamiento en x , $G_{ij}(x, \xi)$ = función de Green y $\phi_j(\xi)$ = densidad de fuerza en ξ en la dirección j .

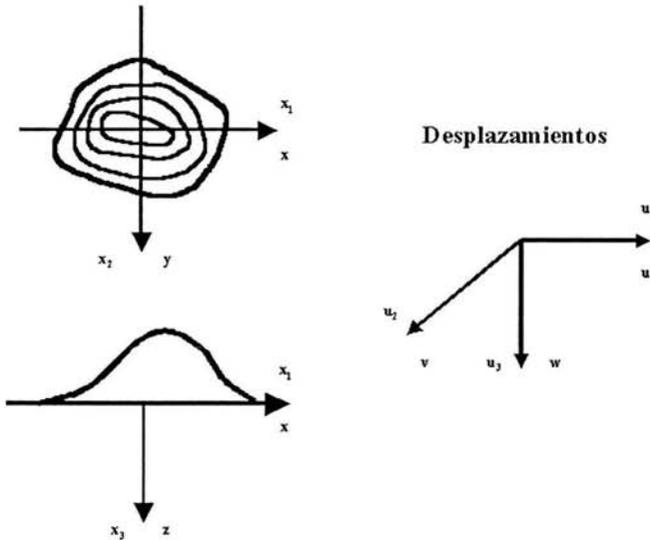


Figura 3.1 Sistema de coordenadas en una montaña 3D y notación usada para los desplazamientos

La representación integral de la ecuación, (fig.3.1), permite el cálculo de los esfuerzos y tracciones mediante la aplicación directa de la ley de Hooke, excepto en las

singularidades de la frontera. Considerando condiciones de equilibrio alrededor de la frontera es posible escribir:

$$t_i(x) = c\phi_j(\bar{x}) + \int_S \phi_j(\xi) T_{ij}(x, \xi) dS_\xi \quad (3.2)$$

donde t_i = iésima componente de la tracción en la frontera, $c = 0.5$ si x se aproxima a S por dentro de la región y $c = -0.5$ si x se aproxima a S desde fuera de la región; $T_{ij}(\bar{x}, \bar{\xi})$ = función de Green de la tracción.

3.2) Difracción de ondas debido a estructuras topográficas superficiales

Considerando una irregularidad topográfica como la mostrada en la figura 3.2. El movimiento total en el semiespacio es la superposición del denominado campo de ondas difractadas con la solución de campo libre.

$$u_i^E = u_i^{(0)} + u_i^{(d)} \quad (3.3)$$

donde $u_i^{(0)}$ es el desplazamiento de campo libre en la dirección i , es decir, la solución en el semiespacio en ausencia de la irregularidad y para este caso es la suma del campo incidente más el reflejado.

El campo de ondas difractado puede escribirse como:

$$u_i^{(d)}(x) = \int_{\partial E} \phi_j(\xi) G_{ij}(x, \xi) dS_\xi \quad (3.4)$$

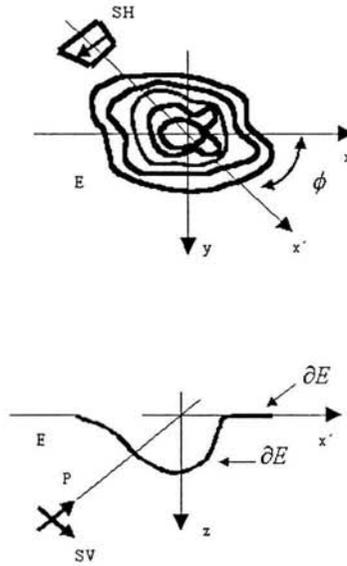


Figura 3.2 Irregularidad topográfica superficial en un semiespacio E e incidencia de ondas elásticas P y S

La condición de frontera libre para las tracciones implica que:

$$t_i^{(0)} + t_i^{(d)} = 0 \quad (3.5)$$

utilizando la ecuación (3.2), se puede expresar como:

$$0.5\phi_i(x) + \int_{\partial E} \phi_j(\xi) T_{ij}(x, \xi) dS_\xi = -t_i^{(0)} \quad (3.6)$$

Esta ecuación resulta ser la que produce el campo difractado. Las ecuaciones (3.4) y (3.6) forman un sistema de ecuaciones integrales para fuentes en la frontera. Estas ecuaciones serán discretizadas a lo largo de una porción finita de la frontera ∂E .

3.3) Discretización

Para resolver sistemas de ecuaciones integrales es necesario discretizarlos. Estas integrales recorren la superficie del contorno del espacio elástico, por lo que esta discretización se realiza para la superficie de la frontera. La elección del esquema de discretización depende del problema y de la formulación matemática que utilice. Si el problema es de 2D, la superficie se construye con segmentos iguales rectilíneos de longitud ΔS con pendientes distintas, esto se muestra en la siguiente figura:

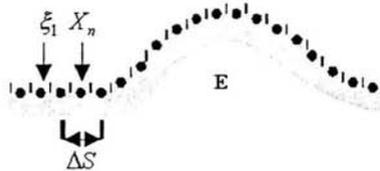


Figura 3.3 Discretización de la superficie de una irregularidad topográfica superficial bidimensional en segmentos rectilíneos con pendientes distintas.

Para poder integrar las funciones de Green es necesario conocer el valor de ΔS y el vector normal n correspondientes a cada elemento. En este caso, o sea, bidimensional, cada longitud ΔS y cada vector normal se conoce desde la construcción de las fronteras por elementos.

Para el caso de 3D se utiliza un esquema simplificado, se utilizan círculos de varios tamaños para discretizar las superficies de tal forma que se cubre toda la frontera. Esto puede parecer un tanto burdo, pero permite una formulación simple y una implementación relativamente fácil. Se eligió este esquema porque las integrales de Green sobre circunferencias se pueden obtener fácilmente.

La discretización se realiza dividiendo las superficies de bandas y rellenando cada banda con circunferencias de radios similares. La suma total de las áreas para todas las circunferencias es igual al área de la banda correspondiente.

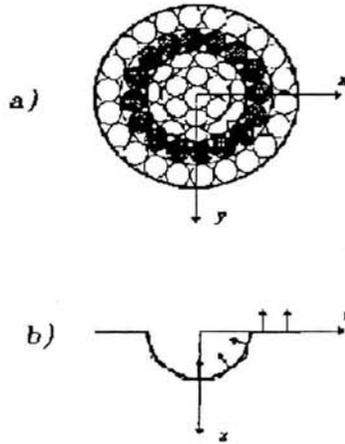


Figura 3.4 Cañón semiesférico tridimensional a) vista desde arriba y distribución de elementos en la superficie. Se muestran sombreados los círculos correspondientes a una banda. b) perfil del plano x-z con los vectores normales representados en algunos de los elementos.

En la discretización se asume que las densidades de fuerza $\phi_j(\xi)$ son constantes en cada uno de los elementos de frontera, se obtiene un sistema de ecuaciones lineales de la forma:

$$\sum_{l=1}^N \phi_l(\xi_l) t_{ij}(x_n, \xi_l) = -t_i^{(0)}(x_n) \quad (3.7)$$

donde

$$t_{ij}(x_n, \xi_l) = \frac{1}{2} \delta_{ij} \delta_{nl} + \int_{\Delta S_l} T_{ij}(x_n, \xi) dS_\xi \quad (3.8)$$

estas integrales se evalúan numéricamente, excepto para el caso $n = l$. En este caso, se tiene:

$$t_{ij}(x_n, \xi_n) = \frac{1}{2} \delta_{ij} \quad (3.9)$$

El sistema de ecuaciones integrales discretizadas se representan como $M\phi = t$, donde M es la matriz de coeficientes, t es el vector de términos independientes y ϕ son las

incógnitas. Este es un sistema de ecuaciones lineales con coeficientes complejos que se puede resolver haciendo una descomposición de la matriz $M = L \cdot U$, donde la matriz L es una matriz diagonal superior y U una matriz diagonal inferior. Una vez que los valores de $\phi_j(\xi_l)$ son conocidos, el campo difractado se calcula por medio de:

$$u_i^{(d)} = \sum_{l=1}^N \phi_l(\xi_l) g_{ij}(x, \xi_l) \quad (3.10)$$

donde

$$g_{ij}(x, \xi_l) = \int_{\Delta S_l} G_{ij}(x, \xi) dS_\xi \quad (3.11)$$

Estas integrales también se calculan numéricamente, excepto en el caso en que x cae en una vecindad de ξ_l , para el cual se hace uso de expresiones analíticas. Este tipo de aproximación es adecuado si la longitud de onda mínima es de al menos cuatro veces el diámetro de los elementos circulares considerados en la discretización.

4) Formulación IBEM/Gradiente conjugado

El IBEM da como resultado un sistema de ecuaciones (ec.3.7, cap.3), el cual es representado por $M\phi = t$, como se explicó en el capítulo anterior. Se presenta una dificultad importante con la matriz de coeficientes asociada (M), debido a que su tamaño crece con la frecuencia. En el caso de problemas tridimensionales, este crecimiento es del orden del cuadrado de la frecuencia y llega a dar origen a sistemas de ecuaciones lineales demasiado grandes que no se pueden resolver en una computadora personal por la gran cantidad de memoria que se necesita para ello.

Para resolver este problema, se propone el truncamiento de la matriz, es decir, convertirla en rala o porosa, para esto se utiliza la llamada Tecnología de matrices ralas o porosas. El método iterativo del gradiente conjugado, se utiliza para resolver sistemas truncados.

El sistema de ecuaciones integrales representado en (3.7) de la sección anterior, se puede escribir en forma matricial como:

$$M\phi = t \quad (4.1)$$

El número total de ecuaciones es igual a tres veces la cantidad de elementos de frontera considerados en la discretización. Las amplitudes de los coeficientes complejos de la matriz M , son mayores para los términos que se encuentran en la diagonal principal. La forma de la matriz de coeficientes, donde los elementos no nulos aparecen en tonalidades de gris, se muestra en la figura 4.1.

El método IBEM da origen a grandes sistemas de ecuaciones. La cantidad considerable de operaciones requeridas para resolver estos sistemas y de memoria para su almacenamiento, hace atractivo el empleo de criterios de truncamiento matricial en combinación con métodos de solución para matrices porosas. De esta forma, es posible reducir el número de operaciones de punto flotante a solamente una fracción de las originalmente utilizadas para invertir el sistema de ecuaciones completo.

La inversión matricial fue acelerada empleando criterios de truncamiento. En otras palabras, las funciones de Green en tres dimensiones exhiben un importante decaimiento espacial y los esfuerzos asociados pueden tener muy poca influencia para distancias relativamente grandes. Sin embargo, se debe hacer notar que los esfuerzos crecen como una función lineal de la frecuencia.

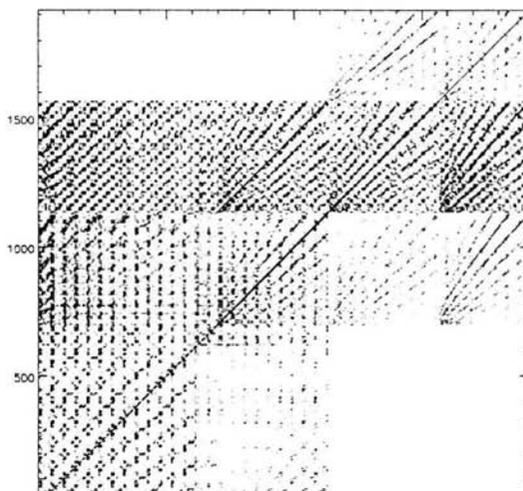


Figura 4.1 Estructura de la matriz de coeficientes para el método IBEM, aplicado al estudio de la respuesta sísmica de una estructura irregular tridimensional.

Un sistema de ecuaciones lineales se considera poroso cuando un porcentaje significativo de los elementos de la matriz de coeficientes tiene amplitud cero. En general, los elementos no nulos en la matriz de coeficientes original son aproximadamente el 80% del número total de entradas. Esto es, el sistema de ecuaciones original no es un sistema raro o poroso. El patrón de ubicación de los elementos no nulos en el sistema original se muestra en la figura anterior.

La llamada tecnología de matrices porosas permite almacenar y resolver grandes sistemas de ecuaciones lineales en forma muy eficiente. Su eficiencia radica en que se consideran en el proceso de solución exclusivamente las entradas no nulas de la matriz de coeficientes.

Con objeto de transformar el sistema original en una aproximación porosa, se utiliza un criterio de truncamiento a partir de relaciones simplificadas basadas en la física del problema en cuestión. Estas relaciones pretenden sacar ventajas del fuerte decaimiento espacial de las funciones de Green en tres dimensiones. De acuerdo a este nuevo esquema, se ignoran todas las contribuciones debidas a elementos localizados más allá de una cierta distancia, medida a partir del punto de evaluación de las tracciones. Dicha distancia se calcula en forma aproximada mediante la regla:

$$r \approx \frac{\pi\eta}{\varepsilon N_b} \quad (4.2)$$

donde N_b es el número total de elementos de frontera, η es la frecuencia normalizada y ε es un parámetro de control que se expresa como un porcentaje de la densidad de fuerzas correspondiente, φ .

Una ventaja de esta estrategia, respecto a criterios de truncamiento basados en la amplitud de los elementos de la matriz de coeficientes, es que la dependencia con la frecuencia puede ser removida. Esto se consigue fijando, para todo el intervalo de frecuencias a evaluar, el valor de la frecuencia normalizada como $\eta = \eta_{max}$, donde η_{max} es el máximo valor permitido por la discretización. De este modo, los patrones de porosidad se conservan para el rango completo de frecuencias y es posible conseguir ahorros adicionales en memoria de almacenamiento y tiempos de cómputo.

4.1) Descripción del algoritmo del Gradiente Conjugado

La resolución de ecuaciones de diferencias en derivadas parciales, tanto por métodos de elementos finitos como por métodos de diferencias finitas, suelen requerir la resolución de grandes sistemas de ecuaciones lineales dispersos. Normalmente, la resolución de dichos sistemas resulta muy costosa si se realiza de forma directa, por lo que se suelen utilizar métodos iterativos.

El término "método iterativo" hace referencia a un amplio conjunto de técnicas que, mediante la realización de aproximaciones sucesivas, van obteniendo una estimación cada vez más precisa de la solución de un determinado problema matemático. Estos métodos resultan muy adecuados para la resolución de grandes sistemas de ecuaciones lineales, pudiendo además ser paralelizados de manera más inmediata que los métodos directos.

Los métodos iterativos se pueden clasificar en dos tipos: métodos estacionarios y métodos no estacionarios. Los métodos estacionarios son más simples y fáciles de implementar, pero su velocidad de convergencia suele ser muy baja, de forma que su uso

sólo está justificado con algunos tipos de matrices concretas. Los métodos no estacionarios se diferencian de los anteriores en que hacen uso de información, evaluada en cada iteración, que les permite obtener la solución de modo dinámico. En general, estos métodos suelen ser más eficientes para la mayor parte de los problemas.

El Gradiente Conjugado es un método iterativo no estacionario que permite resolver sistemas lineales bajo la restricción de que la matriz de coeficientes A , sea simétrica y definida positiva. El método del Gradiente Conjugado se basa en realizar una proyección ortogonal en el subespacio de Krylov $K(r_0, A)$, donde r_0 es el residuo inicial. Este método no requiere una factorización de la matriz, ni es dependiente de parámetros definidos por el usuario, como sucede con otros métodos iterativos.

En las sucesivas iteraciones del algoritmo, se obtienen unas soluciones aproximadas al sistema de ecuaciones que se denominan iterandos. Además, para cada iteración, también se calcula el residuo correspondiente a las distintas soluciones y una dirección de búsqueda, que es utilizada para actualizar los valores de ambas variables. El iterando i -ésimo, que denotaremos por x_i se actualiza por el vector que determina la dirección de búsqueda p_i , que a su vez está ponderado por una variable de escalado α_i :

$$X_i = X_{i-1} + \alpha_i P_i \quad (4.3)$$

De igual manera, el residuo del sistema dado por: $r_i = b - Ax$, es actualizado de la forma:

$$r_i = r_{i-1} - \alpha_i q_i \quad \text{donde} \quad q_i = Ap_i$$

La elección de $\alpha_i = r_i^T r_i / p_i^T A p_i$ minimiza el valor de $r_i^T \cdot A^{-1} r_i$. El vector que determina la dirección de búsqueda se actualiza mediante los residuos, según la expresión:

$$p_i = r_i + \beta_{i-1} P_{i-1} \quad (4.4)$$

donde la elección de $\beta_{i-1} = r_i^T r_i / r_{i-1}^T r_{i-1}$ asegura que p_i y Ap_{i-1} , o equivalentemente r_i y r_{i-1} sean ortogonales. De hecho, se puede demostrar que esta elección de β_i , hace a p_i y a r_i , ortogonales a todos los Ap_j y r_j , previos, respectivamente.

Para cada iteración del algoritmo básico (sin preconditionador ni condición de convergencia) es necesario realizar un producto matriz-vector, tres actualizaciones vectoriales y dos productos escalares entre vectores.

La velocidad de convergencia de este algoritmo es difícil de evaluar, sin embargo, se puede demostrar que para una matriz de coeficientes A, se verifica que:

$$\|x_i - \hat{x}\| \leq 2 \left(\frac{\sqrt{k_2} - 1}{\sqrt{k_2} + 1} \right)^i \|x_0 - \hat{x}\| \quad (4.5)$$

donde x_i es la solución obtenida para la iteración i -ésima, \hat{x} es la solución exacta del sistema y k_2 es el número de condición espectral. Este valor depende de los autovalores de A mediante la relación $k_2 = \lambda_{\max}(A) / \lambda_{\min}(A)$, en donde $\lambda_{\max}(A)$ y $\lambda_{\min}(A)$ son, respectivamente, el mayor y menor autovalor de la matriz.

Así pues, la convergencia de este método dependerá, en gran manera, de las propiedades espectrales de la matriz de coeficientes. Mediante el uso de un preconditionador la matriz de coeficientes se transforma en otra equivalente, en el sentido de que el sistema de ecuaciones transformado mantiene la misma solución. La ventaja que se obtiene mediante su uso es que la nueva matriz de coeficientes presenta unas propiedades espectrales más favorables, que permiten una convergencia más rápida. La transformación realizada al sistema tiene la forma:

$$M^{-1}Ax = M^{-1}b \quad (4.6)$$

donde M es la matriz del preconditionador. El diseño de un preconditionador que dé buenos resultados es un problema complejo, que depende directamente de la distribución de la matriz y de las rutinas de paralelización. En nuestro caso hemos tomado M como la matriz identidad, lo que es equivalente a no hacer uso del preconditionador, dado que este trabajo se centra en la obtención de códigos paralelos eficientes, y no en la obtención de un método de resolución.

4.2) Almacenamiento de matrices dispersas

En la mayor parte de las aplicaciones en las que el método del Gradiente Conjugado resulta de interés, el sistema de ecuaciones lineales con el que se trabaja tiene asociada una matriz de coeficientes dispersa. Las matrices dispersas se caracterizan por tener una elevada proporción de elementos nulos. Existe una gran variedad de formatos de almacenamiento para este tipo de matrices como pueden ser el CRS (Compressed Row Storage), CCS (Compressed Column Storage), BCRS (Block Compressed Row Storage), CDS (Compressed Diagonal Storage), JDS (Jagged Diagonal Storage), SKS (Skyline Storage), etc. En todos ellos se intenta contener únicamente las posiciones relativas de los elementos no nulos reduciendo, de ese modo, el volumen de datos almacenado.

La tecnología de matrices dispersas requiere almacenar y manipular listas de “items” o elementos donde cada “item” puede ser un número entero, real, complejo o una entrada con una estructura más complicada, como una matriz o sea, un arreglo. Las operaciones que comúnmente se representan con listas son: adición de un elemento o dato al final de la lista, borrar un elemento del final de la lista, insertar o borrar un elemento en la mitad o al principio de la lista búsqueda de la posición de cierto elemento o la siguiente, clasificación, ordenación, etc. La selección del esquema de almacenamiento depende de las operaciones que se deseen realizar.

La estructura de datos más simple es el arreglo. La manera de representarlos es por ejemplo: $A(I)$, $B(I,J)$, etc. los números pueden ser almacenados directamente en un arreglo. Un arreglo puede contener punteros a un elemento o a una estructura más compleja. Todos los elementos de un arreglo se accesan directamente en un tiempo el cual es independiente del tamaño del arreglo.

Una lista ligada lineal es un conjunto de celdas ligadas o enlazadas juntas en algún orden, cada celda contiene un elemento de la lista y un puntero, el cual indica donde se localiza la siguiente celda. Por ejemplo, considerando que se desea almacenar los números a , b , c y d , en ese orden, en el arreglo $A(I)$. La forma de almacenamiento sería como sigue:

$posición = 1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8$
 $A(I) = x \ b \ x \ d \ a \ x \ c \ x$
 $NEXT(I) = x \ 7 \ x \ 0 \ 2 \ x \ 4 \ x$
 $IP = 5$
 $final = 0$

El arreglo $A(I)$ contiene el elemento actual, donde $NEXT(I)$ indica la posición del siguiente elemento. Una lista llamada IP es necesaria para almacenar el puntero del primer elemento. En este caso comienza con 5. En la posición 5 se encuentra el elemento $A(5) = a$ y $NEXT(5) = 2$ indica que la posición del siguiente elemento es 2. En $final$ se almacena la última celda para indicar el fin de la lista, en el ejemplo $final$ vale 0.

Por ejemplo, se desea insertar el número e entre b y c . Se asume que la celda 3 está vacía y b está en la posición 2.

$A(3) \leftarrow e$
 $NEXT(3) \leftarrow NEXT(2)$
 $NEXT(2) \leftarrow 3$

Para borrar, por ejemplo, el elemento c de la posición 7 de la lista ligada original es muy simple;

$NEXT(2) \leftarrow NEXT(7)$

obviamente, se tiene que saber que el elemento anterior era b y estaba en la posición 2, con este procedimiento es borrado b e insertado c en su lugar.

La siguiente estructura de datos resulta cuando las posiciones vacías del ejemplo son ligadas y IE es la cabecera de la nueva lista:

$posición = 1\ 2\ 3\ 4\ 5\ 6\ 7\ 8$
 $A(I) = x\ b\ x\ d\ a\ x\ c\ x$
 $NEXT(I) = 3\ 7\ 6\ 0\ 2\ 8\ 4\ 0$
 $IP = 5$
 $IE = 1$
 $final = 0$

La lista ligada se convierte en lista ligada circular cuando en la última posición se guarda el puntero de la posición inicial en lugar del final. Una lista circular no tiene principio ni fin, pero, sin embargo, requiere de una cabecera de lista almacenada aparte.

Una pila es una lista almacenada y manipulada de una manera simple. En una pila, los elementos son almacenados en localidades consecutivas. Se usa un puntero para apuntar a la localidad del último elemento llamado el *top* de la pila. La pila solo se utiliza cuando el elemento se puede agregar o borrar solo en el *top* de la pila. Para agregar (*push*) un nuevo elemento a la pila, se incrementa el puntero en una unidad, se debe verificar si hay suficiente espacio para agregar el elemento y almacenar la localidad a la cual el puntero apunta. Para borrar (*pop*) el último elemento en la localidad solo hay que decrementar el puntero una unidad. Una pila está vacía cuando el puntero vale cero.

Una cola es una lista de elementos almacenados en localidades consecutivas, tiene una terminación por el frente (*front*) y otra por la parte posterior (*rear*) de la cola, los elementos son adicionados solo por el frente y borrados solo por la parte posterior. Una cola está vacía cuando el puntero de “*rear*” está inmediatamente después de la localidad indicada por el puntero de “*front*”.

4.2.1) Almacenamiento de listas de enteros

Las listas de enteros tienen una gran importancia en la tecnología de matrices dispersas, por su gran facilidad y sencillez para ser almacenadas. Una lista de enteros es un subconjunto de $1, 2, \dots, n$. Este tipo de lista se puede almacenar como lista ligada, ya sea como lista lineal o circular.

La importancia de este tipo de almacenamiento es que puede ser fácil representar operaciones como meter una lista dentro de dos listas o concatenar dos listas en una.

4.2.2) Representación y almacenamiento de grafos (diagramas)

Un grafo consiste en un conjunto de vértices (o nodos) y un conjunto de arcos (o aristas). Cada arco se especifica mediante un par de nodos. Las figuras 4.2 a y 4.2 b representan grafos:

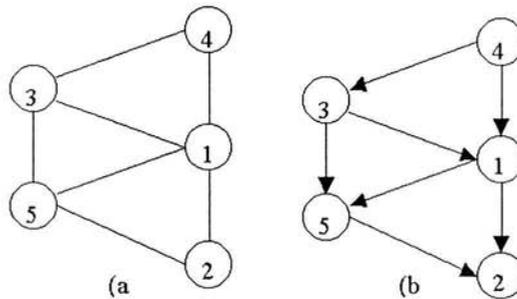


Figura 4.2 Representación de una gráfica no dirigida (a) y una dirigida (b)

Cuando no hay distinción entre el par (u,v) y (v,u) se dice que representa un par no ordenado y que la gráfica es no dirigida. Si los pares de nodos que constituyen los arcos son pares ordenados, se dice que el grafo es dirigido (o dígrafo, fig.b).

La punta de cada flecha representa el segundo nodo del par ordenado de nodos que constituyen un arco, la cola de la flecha representa el primer nodo del par.

Para representar una gráfica en la memoria de la computadora debe ser almacenada, para cada vértice, la lista de vértices, los cuales son adyacentes para ellos. Si las listas se almacenan en forma compacta, se guardan en un arreglo llamado LIST, uno de punteros llamado IP, también es necesario un indicador de la posición donde cada lista comienza.

Otro esquema común de almacenamiento es el de Tabla conectada (connection table). Si la gráfica tiene n vértices y m es el máximo número de vértices adyacentes de cualquier vértice (su grado) entonces, la tabla es un arreglo con n renglones y m columnas

con la lista adyacente de vértices i almacenados en un renglón i . La grafica de (b) quedaría como sigue:

		1	2
1		2	5
2		0	0
3		1	5
4		1	3
5		2	0

Una matriz de adyacencia o conectividad, puede ser utilizada para representar una grafica. Para una grafica con n vértices, la matriz es cuadrada de orden n , y es definida como: $A_{ij}=1$, si y solo si (i,j) es un nodo, de otra forma, $A_{ij}=0$, la grafica para la figura (b) sería:

		1	2	3	4	5
1		0	1	0	0	1
2		0	0	0	0	0
3		1	0	0	0	1
4		1	0	1	0	0
5		0	1	0	0	0

4.2.3) Almacenamiento diagonal de matrices de Banda

Las matrices de banda, representan una estrategia simple y muy utilizada para explotar los ceros de una matriz. Una matriz A es matriz de banda, cuando todos esos elementos no ceros son confinados dentro de una banda formada por diagonales paralelas de la diagonal principal.

La semibanda superior consiste en todos los elementos de la parte superior en la banda, la semibanda inferior consiste en todos los elementos de la parte inferior.

El almacenamiento de la diagonal de una matriz de banda simétrica A en un arreglo $AN(I,J)$ es mostrado a continuación:

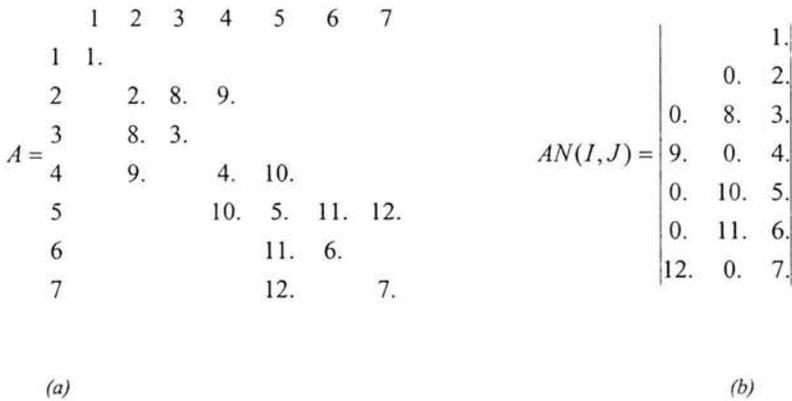


Figura 4.3 Una matriz de banda simétrica de 7x7 (a) con banda ancha igual a 5, y ese almacenamiento diagonal 7x3 (b) en el arreglo rectangular AN(I,J)

La diagonal principal es almacenada en la última columna y las codiagonales inferiores son almacenadas justificadas hacia abajo en las columnas restantes. Para una matriz asimétrica se requiere un arreglo con n por $2\beta + 1$, y las codiagonales superiores son almacenadas justificadas hacia arriba en la parte derecha del arreglo.

Las matrices de banda tienen una propiedad importante, el ancho de banda depende del orden en el cual los renglones y columnas son acomodados. Un ancho de banda pequeño significa menores requerimientos de almacenamiento, también, esto usualmente requiere menor trabajo cuando la matriz es utilizada para computación.

4.2.4) Almacenamiento “envolvente” de matrices simétricas

Jennings (1966) propone un esquema muy eficiente y simple de matrices simétricas. Este es conocido como esquema envolvente o de banda variable. Para cada renglón i de una matriz simétrica A define:

$$\beta_i = i - j_{\min}(i)$$

donde $j_{\min}(i)$ es el mínimo índice de la columna en el renglón i , para el cual $A_{ij} \neq 0$. Así, el primer renglón no cero i sitúa las posiciones β_i a la izquierda de la diagonal y la mitad de banda-ancha se define como:

$$\beta = \max_i(\beta_i)$$

La “envoltura” de A en el conjunto de elementos A_{ij} tal que $0 < i - j \leq \beta_i$. Para un cierto renglón i , todos los elementos con índices de columnas en el rango a $j_{\min}(i)$ a $i - 1$ pertenecen a la “envoltura”, un total de β_i elementos. Los elementos de la diagonal no pertenecen a la “envoltura”. *Profile* de A es el número de elementos en la “envoltura”:

$$profile(A) = \sum_i \beta_i$$

En el esquema de almacenamiento de Jennings, todos los elementos los cuales pertenecen a la “envoltura” son almacenados ordenadamente, renglón por renglón, incluyendo ceros en un vector, llamado AN. Los elementos de la diagonal son almacenados en el final de cada renglón. La longitud de AN es igual a *profile* más el orden de A . También se necesita un arreglo de punteros, llamado IA, las entradas son punteros de las localidades de los elementos de la diagonal en AN. Así, los elementos del renglón i , cuando $i > 1$, son en las posiciones $IA(i-1)+1$ a $IA(i)$. Solamente los elementos del renglón 1 como A_{11} son almacenados en $AN(1)$. Por ejemplo la matriz de la figura 4.3a con *profile*=7, el almacenamiento envolvente quedaría como sigue:

<i>posición</i> =	1	2	3	4	5	6	7	8	9	10	11	12	13	14
<i>AN</i> =	1.	2.	8.	3.	9.	0.	4.	10.	5.	11.	6.	12.	0.	7.
<i>IA</i> =	1	2	4	7	9	11	14							

4.2.5) Esquemas de almacenamiento dispersos ligados.

Los esquemas anteriormente mencionados son muy útiles en una gran cantidad de aplicaciones prácticas. La desventaja podría ser que están diseñados para problemas

pequeños o medianos, porque conforme se van complicando, el esquema se va haciendo más difícil de programar.

Otro esquema de almacenamiento disperso, propuesto por Knuth (1968), dice que los elementos no ceros son almacenados en la memoria de la computadora, en cualquier orden y en forma compacta, en un vector llamado AN.

Dos vectores adicionales paralelos, llamados I y J, contienen los índices de renglones y columnas de cada no-cero correspondiente. Así, para cada $A_{ij} \neq 0$, la triada (A_{ij}, i, j) es guardada en memoria. Además, en orden para poder encontrar fácilmente los elementos en un cierto renglón y columna de la matriz, esto llega a ser necesario para almacenar un par de punteros juntos con cada triada y cabeceras de renglón y columna, las cuales son llamadas en la entrada de cada lista de renglón y columna.

Una variante del esquema de Knuth fue utilizada por Larcombe (1971) para matrices simétricas definidas positivas con elementos los cuales son números o submatrices. Esto es llamado el esquema “columna delantera renglón trasero”. Solo la diagonal y el triángulo superior de la matriz son almacenados y solo es necesario un arreglo de cabecera de línea. Desde cada elemento de la diagonal, inician dos listas: una describiendo el renglón correspondiente a la derecha de la diagonal en la dirección delantera, la otra describiendo la correspondiente columna en la diagonal en la dirección opuesta. Esto explica el nombre en la figura se muestra un ejemplo:

$$\left| \begin{array}{cccc} A_{11} & & A_{13} & A_{14} \\ & A_{22} & & A_{25} \\ & & A_{33} & A_{35} \\ \text{SIMETRIC} & & & A_{44} \\ & & & & A_{55} \end{array} \right|$$

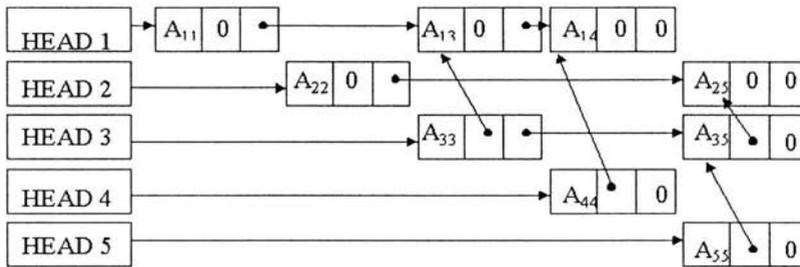


Figura 4.4 Versión de Larcombe de Knuth del esquema de almacenamiento de matrices con elementos no cero sobre la diagonal.

4.2.6) El formato row-wise disperso

El formato row-wise disperso (Chang, 1969; Curtis y Reid, 1971; Gustavson, 1972) descrito aquí es uno de los esquemas del almacenamiento normalmente usados para las matrices dispersas. El esquema tiene los requisitos de almacenamiento mínimos y al mismo tiempo ha demostrado ser muy conveniente para varias operaciones importantes como la suma, multiplicación, permutación y transposición de matrices dispersas, la solución de ecuaciones lineales con una matriz dispersa de coeficientes por los métodos directos o iterativos, etc. En este esquema, los valores de los elementos de la matriz no cero son guardados por las filas, junto con su índice de la columna correspondiente, en dos series, llamadas AN y JA, respectivamente. Un arreglo de punteros, llamados IA, también se proporciona para indicar las locaciones en AN y JA dónde la descripción de cada fila empieza. Una entrada extra en IA contiene un indicador a la primera posición vacía en JA y AN. Un ejemplo es conveniente en este punto. Considere la matriz:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 0 & 0 & 1. & 3. & 0 & 0 & 0 & 5. & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 & 7. & 0 & 1. & 0 & 0 \end{pmatrix}$$

A se representa como sigue:

<i>posición</i> =	1	2	3	4	5	6	
	<i>IA</i> =	1	4	4	6		RR(C)O
	<i>JA</i> =	3	4	8	6	8	
	<i>AN</i> =	1.	3.	5.	7.	1.	

La descripción de la fila 1 de A empieza en la posición $IA(1) = 1$ de AN y JA. Desde la descripción de la fila 2 empieza a $IA(2) = 4$, esto significa que la fila 1 de A se describe en las posiciones 1, 2 y 3 de AN y JA. En el ejemplo:

- IA(1) = 1 primera fila empieza en JA(1) y AN(1).
- IA(2) = 4 segunda fila empieza en JA(4) y AN(4).
- IA(3) = 4 tercera fila empieza en JA(4) y AN(4). Esta es la misma posición en que la fila 2 empieza, esto significa que la fila 2 está vacía.
- IA(4) = 6 ésta es la primera posición vacía en JA y AN. La descripción de la fila 3 finaliza en posición $6 - 1 = 5$ de JA y AN.

En general, la fila r de A se describe en las posiciones $IA(r)$ a $IA(r + 1) - 1$ de JA y AN, excepto cuando $IA(r + 1) = IA(r)$ en cuyo caso la fila r está vacía. Si la matriz A tiene m filas, entonces IA tiene $m + 1$ posiciones.

Se dice que esta representación está completa porque toda la matriz A se representa, y es ordenada porque se guardan los elementos de cada fila en el orden ascendente de su índice de la columna. Es así una representación Row-Wise Completa y Ordenada, o RR(C)O.

Los arreglos IA y JA representan la estructura de A, dado como la lista adyacente del gráfico asociado con A. Si un algoritmo es dividido en una sección simbólica y una sección numérica, los arreglos que IA y JA son procesados por la sección simbólica, y el arreglo AN por la sección numérica.

Gustavson (1972) también propuso una variante de almacenamiento row-wise, conveniente para aplicaciones que requieren ambas operaciones fila y columna. A es almacenada como el descrito row-wise, y además la estructura de A^T se procesa y también se almacena como row-wise. Una representación row-wise de la estructura de A^T es

idéntica a una representación column-wise de la estructura de A. Puede obtenerse por la transposición de la estructura row-wise de A. Por ejemplo, este esquema se ha usado para las aplicaciones de programación lineal.

Un esquema row-wise más simple se propuso por Key (1973) para matrices asimétricas. Los no ceros se sostienen en un arreglo bidimensional de tamaño n por m donde n es el orden de la matriz y m es número máximo de no ceros en el renglón. Este esquema es fácil de manipular pero tiene la desventaja que m no puede ser predecible y puede resultar muy grande.

4.2.7) Representaciones ordenadas y desordenadas

Las representaciones de matrices dispersas no es necesario que sean ordenadas, en el sentido que los elementos de cada renglón pueden ser almacenados en cualquier orden de tiempo, preservando el orden de los renglones. La matriz A del ejemplo bien puede ser una representación row-wise completa, pero desordenada:

$$\begin{array}{rcccccc} \text{posición} = & 1 & 2 & 3 & 4 & 5 & \\ IA = & 1 & 4 & 4 & 6 & & \\ JA = & 8 & 3 & 4 & 8 & 6 & \\ AN = & 5. & 1. & 3. & 1. & 7. & \end{array} \quad \text{RR(U)}$$

Las representaciones desordenadas pueden ser muy convenientes. Los resultados de la mayoría de operaciones de matrices son obtenidos en representaciones desordenadas, y esto, computacionalmente, es muy costoso para ordenarlos. De otra manera, con pocas excepciones, los algoritmos de matrices dispersas, no requieren que las representaciones sean ordenadas. Las representaciones column-wise son también utilizadas, pero ellas pueden ser consideradas como representaciones row-wise de las matrices transpuestas. Por ejemplo, la matriz A puede ser representada como sigue:

$$\begin{array}{l}
 \text{posición} = 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \quad 8 \quad 9 \quad 10 \quad 11 \\
 IAT = 1 \quad 1 \quad 1 \quad 2 \quad 3 \quad 3 \quad 4 \quad 4 \quad 6 \quad 6 \quad 6 \\
 JAT = 1 \quad 1 \quad 3 \quad 1 \quad 3 \\
 ANT = 1 \quad 3 \quad 7 \quad 5 \quad 1
 \end{array}
 \qquad \text{CR(C)O}$$

Lo cual puede ser considerado como un RR(C)O de A^T . Si la matriz dada es simétrica, es suficiente para representar solo esta diagonal y triangulo superior. Considerando el siguiente ejemplo:

$$B = \begin{array}{c}
 \begin{array}{cccc}
 & 1 & 2 & 3 & 4 \\
 1 & | 2. & 0 & 0 & 1. | \\
 2 & | 0 & 1. & 0 & 1. | \\
 3 & | 0 & 0 & 3. & 0 | \\
 4 & | 1. & 1. & 0 & 3. |
 \end{array}
 \end{array}$$

La representación row-wise, diagonal y superior, ordenada de B es:

$$\begin{array}{l}
 \text{posición} = 1 \quad 2 \quad 3 \quad 4 \quad 5 \quad 6 \quad 7 \\
 IB = 1 \quad 3 \quad 5 \quad 6 \quad 7 \\
 JB = 1 \quad 4 \quad 2 \quad 4 \quad 3 \quad 4 \\
 BN = 2. \quad 1. \quad 1. \quad 1. \quad 3. \quad 3.
 \end{array}
 \qquad \text{RR(DU)O}$$

Si además, la mayoría de los elementos de la diagonal de la matriz simétrica dada son diferentes de cero (como en el caso de matrices simétricas definidas-positivas), entonces ellos pueden ser almacenados en un arreglo separado BD, y solo el triangulo superior de B es representado en formato disperso:

$$\begin{array}{l}
 \text{posición} = 1 \quad 2 \quad 3 \quad 4 \quad 5 \\
 IB = 1 \quad 2 \quad 3 \quad 3 \quad 3 \\
 JB = 4 \quad 4 \\
 BN = 1. \quad 1. \\
 BD = 2. \quad 1. \quad 3. \quad 3.
 \end{array}
 \qquad \text{RR(U)O}$$

Esta es una representación muy densa y conveniente.

4.2.8) Compresión de Sherman

El esquema de almacenamiento comprimido, propuesto por Sherman (1975), el cual es una variante del esquema row-wise, es utilizado para almacenar matrices dispersas triangulares obtenidas por factorización de Gauss. Para este esquema, primero hay que definir una matriz triangular inferior como una matriz solo con elementos no cero en ella y en la diagonal principal, y una matriz triangular superior solamente con elementos no cero en ese triángulo superior y diagonal principal, por ejemplo:

Matriz triangular inferior $A_{ij} = 0$ si $j > i$

Matriz triangular superior $A_{ij} = 0$ si $j < i$

La aplicación más importante de las matrices triangulares es la factorización de una matriz A:

$$A = LDU,$$

donde L es una matriz triangular inferior, U es una matriz triangular superior y D es la diagonal. Si A es dispersa, entonces L y U son dispersas.

L, D y U son obtenidos de A por medio de algún proceso de eliminación, como por ejemplo, la eliminación de Gauss, y a menudo pasa que el conjunto de renglones de U o de L, tiene una estructura similar. Frecuentemente, pasa que los renglones i y j de U, $i < j$, tienen estructura idéntica a la derecha de la posición j. Sherman propuso un almacenamiento comprimido de índices, los cuales requieren un arreglo adicional de punteros.

4.2.9) Almacenamiento de matrices de bloque-particiones

Los bloques pueden ser tratados como si fuera una matriz de matrices. Las particiones juegan un importante papel en la tecnología de matrices dispersas, porque muchos algoritmos designados, primero para matrices de números, pueden ser generalizados para operar sobre matrices de matrices. Las particiones pueden ser consideradas simplemente como una herramienta de manejador de datos, la cual ayuda a organizar la transferencia de información entre la memoria principal y los dispositivos

auxiliares. Las particiones de renglones y columnas no necesariamente son las mismas, si eso pasa, los bloques de la diagonal son cuadrados.

4.3) Tecnología de matrices ralas

No todos los sistemas de ecuaciones que se plantean en la práctica son lineales. Sin embargo, existen técnicas que consisten en reemplazar el sistema original no lineal por una aproximación lineal del mismo. Bajo circunstancias adecuadas, las soluciones del sistema lineal aproximan a las soluciones del sistema de partida. Es decir, al precio de perder exactitud se obtiene el beneficio de pasar de un sistema difícil de resolver a otro más sencillo.

El inconveniente de estas técnicas es que producen sistemas lineales con una gran cantidad de ecuaciones e incógnitas. Y mientras mejor se quiera aproximar, tanto mayor será el tamaño de los sistemas lineales. Esto es un problema porque, por ejemplo, un sistema con 100 ecuaciones y 100 incógnitas tiene una matriz de 10,000 entradas y uno de 1,000 por 1,000 una matriz de 1,000,000 de entradas. Por razones de espacio y tiempo, no es posible resolver estos sistemas con métodos tradicionales de reducción de matrices.

Existe sin embargo una hipótesis adicional que normalmente se cumple en estos casos y es que las matrices que se obtienen poseen escasas entradas diferentes de cero. Si bien en todas las filas hay elementos no nulos, cada fila tiene un número comparativamente chico de coordenadas no nulas. Una matriz de estas características se denomina rala. No existe una definición precisa que determine cuantos coeficientes no nulos debe poseer cada fila para que la matriz sea considerada rala, por lo que esta noción no tiene una definición matemática que la describa. La idea es sin embargo bien simple y significa que en cada fila, el número de elementos no nulos es de un “orden de magnitud” inferior a su tamaño.

Por ejemplo, si una matriz de 1,000 por 1,000 tiene a lo sumo 5 entradas no nulas en cada fila, en realidad la información que contiene no es la de un millón de entradas sino la de 5,000.

Las matrices asociadas con los sistemas lineales se clasifican como densas y esparcidas. Las matrices densas tienen pocos elementos nulos, y el orden de las matrices tiende a ser relativamente pequeño quizá de orden 100 o menos. Con frecuencia es más

eficiente manejar problemas en los cuales intervengan tales matrices por métodos directos. Casi siempre surgen de intentos para resolver ecuaciones diferenciales por métodos de diferencias finitas. El orden de tales matrices puede ser muy grande, y son idealmente adecuadas para solución por métodos iterativos que saquen ventaja de la naturaleza esparcida de la matriz implicada.

4.4) Métodos Directos e Iterativos

Los métodos numéricos para resolver sistemas lineales se pueden dividir en 2 tipos, directos e iterativos. Los métodos directos son aquellos en los cuales, en ausencia de errores de redondeo u otros errores, se obtiene la solución exacta en un número finito de operaciones aritméticas elementales. En la práctica, puesto en computador trabaja con una lógica de palabra finita, los métodos directos no conducen a soluciones exactas. En realidad, los errores que provienen de redondeo, inestabilidad y pérdida de cifras significativas pueden conducir a resultados extremadamente pobres o inútiles. Una gran parte del análisis numérico tienen que ver con el porqué y el cómo del surgimiento de tales errores y con la búsqueda de métodos para minimizar la totalidad de estos errores.

El método fundamental usado para la solución directa es la eliminación de Gauss, pero aún dentro de esta clase hay una variedad de métodos que varían en eficiencia computacional y en exactitud.

Los métodos iterativos son aquellos que parten de una aproximación inicial y que, aplicando un algoritmo convenientemente escogido, conducen a aproximaciones sucesivamente mejores. Aun si el proceso converge, podemos esperar solo la obtención de una solución aproximada por métodos iterativos. Los métodos iterativos varían de acuerdo al algoritmo escogido y su rapidez de convergencia. Algunos métodos iterativos pueden ser de convergencia tan lenta que son inútiles computacionalmente. Las ventajas importantes de los métodos iterativos son la simplicidad y uniformidad de las operaciones que se realizan, las cuales los hacen adecuados para usarlos en computadoras y su relativa insensibilidad al crecimiento de errores de redondeo.

Los métodos iterativos dejan quieta a la matriz de entrada y hacen variar el vector de soluciones. En estos métodos se empieza proponiendo una solución y se comienza a

iterar de modo que en cada paso la solución propuesta es reemplazada por otra que aproxima mejor a la solución real del sistema. Después de un número de iteraciones, el valor calculado por el algoritmo se toma como “solución aproximada”. Lo bueno de este método es que en todo momento trabaja con la misma matriz rara, lo malo es que no calcula la solución real del sistema, sino una aproximación.

4.5) Esquemas de truncamiento matricial

Enseguida se discutirán algoritmos que encuentran la solución x del sistema

$$Ax=b \quad (4.7)$$

donde A es una matriz dispersa real no singular cuadrada de $n \times n$ y b un vector lleno. Los algoritmos pueden agruparse en dos categorías: los métodos directos y los métodos iterativos. Como se explicó en la sección anterior, los métodos directos son basados en la eliminación de Gauss: se modifican las ecuaciones o las incógnitas del sistema en los pasos sucesivos hasta que la solución se encuentre. En los métodos iterativos, una suposición inicial es normalmente hecha para x , y esta suposición se mejora entonces hasta que la exactitud suficiente se obtenga. Ambos métodos tienen ventajas y desventajas en cada caso particular y es difícil de declarar las reglas generales acerca de cual es el más conveniente.

Nuestra atención se restringe a métodos directos para resolver la ecuación (4.7) con A una matriz cuadrada no singular real de orden n . Un sistema complejo de ecuaciones lineales puede transformarse en uno real de dos veces el orden. La solución de ecuación (4.7) puede escribirse: $x=A^{-1}b$, pero el uso explícito de esta forma sólo es hecho en el caso de n pequeño (dígase $n \sim 100$ o menos). A^{-1} es una matriz densa aún cuando A es dispersa, y su proceso implica perder las ventajas de dispersión. Para una ventaja del sistema grande se toma del hecho que A^{-1} no se necesita explícitamente: A^{-1} puede expresarse en la forma de un producto de matrices tal que la acción en b es fácil procesar. Por supuesto, tal procedimiento es conveniente porque los factores de las matrices son esparcidos y todos pueden ser almacenados en mucho menos espacio que requeriría para A^{-1} solo, y también porque el producto del factor de matrices y b requiere pocas operaciones que el producto

$A^{-1}b$ solo podría requerir. La manera usual por resolver la Ecuación (4.7) por un método directo es procesar la factorización triangular de A:

$$A = LU \quad (4.8)$$

donde U es una matriz triangular superior (su triángulo inferior contiene sólo ceros) con una diagonal de unos, y L es triangular inferior (su triángulo superior contiene sólo ceros). Entonces $A^{-1} = U^{-1}L^{-1}$ y $x = U^{-1}w$ donde $w = L^{-1}b$. Así, la solución de Ecuación (4.7) se encuentra resolviendo los sistemas lineales:

$$Lw = b \quad (4.9)$$

para w, y

$$Ux = w \quad (4.10)$$

para x. Si A es dispersa, ambos L y U son dispersas, aunque normalmente menos dispersa que A. Los sistemas como (4.9) y (4.10), son muy simple de resolver.

La descomposición (4.8) es única si A es representada como un producto y es no singular. Frecuentemente una matriz diagonal D se factoriza fuera de L de tal manera que $L = L' T$ donde L' es triangular inferior con la diagonal de la unidad. La ecuación (4.8) se vuelve:

$$A = L'DU \quad (4.11)$$

Si A es simétrico, entonces $L' = U^T$. Otras descomposiciones de A son posibles. Es importante examinar las posibilidades porque ellas proporcionan la flexibilidad en el plan de algoritmos y esquemas del almacenamiento.

Las computadoras introducen redondeo-fuera de los errores al realizar los funcionamientos aritméticos. Los procesos de la balanza grandes como la solución de un sistema (4.7) de orden alto, un análisis del error cuidadoso y el uso de técnicas especiales exigen garantizar un resultado final significativo con una exactitud especificada. En la eliminación de Gauss, la acumulación del error excesiva se previene normalmente

descascarando la matriz original, seleccionando los pivotes convenientes en el curso de la eliminación, y supervisando el crecimiento del elemento en matrices reducido sucesivo.

La estructura cero-no-cero de A depende de la fuente de donde el sistema (4.7) se obtuvo; por ejemplo, la matriz se asoció con un problema de elementos finitos y eso asociado con una red eléctrica tiene las estructuras completamente diferentes. Una consecuencia de esto es la existencia de dos categorías de algoritmos: aquellos de que toman ventaja de las propiedades particulares de A y son así ideales para las aplicaciones específicas, y aquellos que trabajan igualmente bien con cualquier matriz dispersa general. Normalmente, cuando más restricciones se ponen, el resultado de los algoritmos es más eficaz. Hay gran interés en los algoritmos en vías de desarrollo para las aplicaciones específicas y la investigación en este campo es actualmente muy activa.

La estructura de la matriz de un sistema de ecuaciones lineales puede modificarse si las incógnitas se reenumeran y las ecuaciones permutan. Pero esto hace que disminuya la eficacia de la mayoría de los algoritmos iterativos, se obtienen los efectos importantes cuando un método directo se usa. Esto es porque en el curso de eliminación se generan nuevos coeficientes no-cero y la matriz del sistema se pone menos dispersa: sufre de llenado. Es un hecho bien establecido que la cantidad de llenado depende drásticamente de la manera en que se piden las incógnitas y las ecuaciones. Desde el almacenamiento y procesamiento requerido para resolver el sistema, y la exactitud del resultado final depende a su vez de la cantidad de llenado, es importante considerar clasificaciones que reducen el llenado tanto como sea posible.

Un método directo alternativo para resolver las ecuaciones lineales es la factorización ortogonal. La Ortogonalización da mejor estabilidad numérica que la eliminación de Gauss, pero el llenado es mucho mayor. No se recomiendan métodos basados en la ortogonalización para los sistemas dispersos generales. Sin embargo, ellos pueden ser beneficiosos en algunos casos, por ejemplo en los de mínimos-cuadrados lineales y la programación lineal, y en el algoritmo de AQ (Borland, 1981) para casos donde la matriz de coeficientes puede permutarse cerca de la forma más inferior de Hessenberg.

No sólo pueden aplicarse eliminación y métodos de factorización a matrices de números sino también a matrices de submatrices, dando lugar para bloquear los métodos.

Éstos son algebraicamente similares a la eliminación ordinaria, sólo que los funcionamientos como la suma, la multiplicación e inversión deben ser reemplazadas por la suma de la matriz, multiplicación de la matriz e inversión de la matriz.

4.5.1) Algunas definiciones y propiedades

Un vector columna es una matriz con una única columna, un vector renglón es una matriz con un único renglón y se puede expresar como el transpuesto de un vector columna.

Se puede sumar o restar matrices con el mismo número de columnas y renglones.

Una matriz B puede multiplicar a la izquierda (premultiplicarse) a otra matriz A, si el número de columnas de A es igual al número de renglones de B.

Si $BA = I$ o $AB = I$, donde I es una matriz identidad, entonces $B = A^{-1}$.

Si A es simétrica definida positiva, existe una única factorización de Cholesky, $A = U^T U$, donde U es triangular superior con elementos positivos en la diagonal. Además, U puede ser escrita $U = D' U'$, donde D' es la diagonal con elementos positivos y U' es triangular superior con unos diagonal unitaria. Así,

$$A = U'^T D' U'$$

donde $D = D'^2$, es también una única factorización de A.

4.5.2) Descomposición LU

El esquema de descomposición LU es una transformación de una matriz A como producto de dos matrices,

$$A = LU$$

donde L es una matriz triangular inferior y U es una matriz triangular superior. Cuando uno debe resolver varios conjuntos de ecuaciones lineales en los que todas las matrices de coeficientes son iguales pero los términos no homogéneos (lado derecho) son distintos, la solución de las ecuaciones utilizando la descomposición LU tiende a ser más eficiente que la eliminación de Gauss.

La descomposición LU para una matriz de 3x3 se vería de esta manera.

$$\begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ l_{2,1} & 1 & 0 \\ l_{3,1} & l_{3,2} & 1 \end{bmatrix} \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ 0 & u_{2,2} & u_{2,3} \\ 0 & 0 & u_{3,3} \end{bmatrix} \quad (4.12)$$

Se observa que los elementos de la diagonal valen 1.

Para evaluar u y l en la ecuación 4.12 sin pivoteo, primero se multiplica el primer renglón de L por cada columna de U y se compara el resultado con el primer renglón de A . Se tiene entonces que el primer renglón de U es idéntico al de A :

$$u_{1,j} = a_{1,j}, \quad j=1 \text{ a } 3 \quad (4.13)$$

Multiplicando el segundo y tercer renglones de L por la primera columna de U respectivamente, y lo comparamos con el lado izquierdo para obtener

$$a_{2,1} = l_{2,1}u_{1,1}, \quad a_{3,1} = l_{3,1}u_{1,1}$$

o en forma equivalente

$$l_{2,1} = a_{2,1} / u_{1,1}, \quad l_{3,1} = a_{3,1} / u_{1,1} \quad (4.14)$$

Multiplicando el segundo renglón de L por la segunda y tercera columnas de U y las comparamos con el lado izquierdo para obtener

$$a_{2,2} = l_{2,1}u_{1,2} + u_{2,2}, \quad a_{2,3} = l_{2,1}u_{1,3} + u_{2,3}$$

o bien

$$u_{2,2} = a_{2,2} - l_{2,1}u_{1,2}, \quad u_{2,3} = a_{2,3} - l_{2,1}u_{1,3} \quad (4.15)$$

Se multiplica el tercer renglón de L por la segunda columna de U y se obtiene

$$a_{3,2} = l_{3,1}u_{1,2} + l_{3,2}u_{2,2}$$

o, en forma equivalente,

$$l_{3,2} = [a_{3,2} - l_{3,1}u_{1,2}] / u_{2,2} \quad (4.16)$$

Finalmente, $l_{3,3}$ se obtiene multiplicando la última columna de U por el último renglón de L y lo igualamos a $a_{3,3}$ como sigue

$$l_{3,1}u_{1,3} + l_{3,2}u_{2,3} + u_{3,3} = a_{3,3}$$

o bien

$$u_{3,3} = a_{3,3} - l_{3,1}u_{1,3} - l_{3,2}u_{2,3} \quad (4.17)$$

El esquema general de la descomposición LU para una matriz de orden N es el siguiente:

El primer renglón de U, $u_{i,j}$ para $j = 1$ hasta N, se obtiene por medio de

$$u_{1,j} = a_{1,j}, \quad j = 1 \text{ hasta } N \quad (4.18)$$

La primera columna de L, $l_{i,1}$ para $i = 2$ hasta N, se obtiene por medio de

$$l_{i,1} = a_{i,1} / u_{1,1}, \quad i = 2 \text{ hasta } N \quad (4.19)$$

El segundo renglón de U se obtiene como

$$a_{2,j} = a_{2,j} - l_{2,1}u_{1,j}, \quad j = 2 \text{ hasta } N \quad (4.20)$$

La segunda columna de L se obtiene mediante

$$l_{i,2} = [a_{i,2} - l_{i,1}u_{1,2}] / u_{2,2} \quad i = 3 \text{ hasta } N \quad (4.21)$$

El n-ésimo renglón de u se obtiene de

$$u_{n,j} = a_{n,j} - \sum_{k=1}^{n-1} l_{n,k}u_{k,j}, \quad j = n \text{ hasta } N \quad (4.22)$$

La n-ésima columna de L se obtiene de

$$l_{i,n} = \left[a_{i,n} - \sum_{k=1}^{n-1} l_{i,k}u_{k,n} \right] / u_{n,n}, \quad i = n+1 \text{ hasta } N \quad (4.23)$$

En el proceso anterior, no se calculan los elementos de la diagonal de L, es decir, $l_{i,i}$, puesto que todos valen 1.

Como se habrá observado, los elementos de la parte triangular superior de L son iguales a cero. También los elementos de la parte triangular inferior de la matriz U se anulan. Por lo tanto, los elementos de L y U se pueden guardar en un arreglo con el fin de ahorrar espacio en la memoria. Por ejemplo, las matrices L y U de la ecuación (4.12) se pueden combinar en un arreglo como

$$\begin{array}{ccc} u_{1,1} & u_{1,2} & u_{1,3} \\ l_{2,1} & u_{2,2} & u_{2,3} \\ l_{3,1} & l_{3,2} & u_{3,3} \end{array}$$

En este arreglo, los elementos de la diagonal de L no se guardan porque valen 1. Para reducir aún más el uso del espacio en la memoria, los resultados de la factorización se escriben encima del espacio de memoria de A. Esto es posible debido a que cada elemento $a_{i,j}$ de A se utiliza solo una vez para calcular $l_{i,j}$ o $u_{i,j}$ en toda la factorización. Por lo tanto, al utilizar $a_{i,j}$, su espacio de memoria se puede utilizar para guardar $l_{i,j}$ o $u_{i,j}$.

Ahora, se estudian las formas para resolver un conjunto de ecuaciones lineales.

La ecuación $Ax=y$ se puede escribir como

$$LUX=y \quad (4.24)$$

donde $LU=A$. La ecuación (4.24) se resuelve como sigue. Sea

$$Ux=z \quad (4.25)$$

La ecuación (4.24)

$$Lz=y \quad (4.26)$$

La solución de la ecuación (4.26) para z es fácil, debido a la forma triangular de L. Una vez que se conoce z, se resuelve la ecuación (4.25) en términos de x.

En caso de una matriz de 3x3, por ejemplo, se puede escribir la ecuación (4.26) como

$$\begin{bmatrix} 1 & 0 & 0 \\ l_{2,1} & 1 & 0 \\ l_{3,1} & l_{3,2} & 1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad (4.27)$$

La solución se calcula en forma recursiva como

$$\begin{aligned} z_1 &= y_1 \\ z_2 &= [y_2 - z_1 l_{2,1}] \\ z_3 &= [y_3 - z_1 l_{3,1} - z_2 l_{3,2}] \end{aligned} \quad (4.28)$$

Escribimos la ecuación (4.25) en forma más explícita

$$\begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} \\ 0 & u_{2,2} & u_{2,3} \\ 0 & 0 & u_{3,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}$$

y la solución es

$$\begin{aligned} x_3 &= \frac{z_3}{u_{3,3}} \\ x_2 &= \frac{z_2 - u_{2,3}x_3}{u_{2,2}} \\ x_1 &= \frac{z_1 - u_{1,2}x_2 - u_{1,3}x_3}{u_{1,1}} \end{aligned}$$

Para una matriz de orden N , las eliminaciones hacia delante y hacia atrás se resumen de la forma siguiente.

Paso de la eliminación hacia delante:

$$\begin{aligned} z_1 &= y_1 \\ z_i &= y_i - \left[\sum_{j=1}^{i-1} l_{i,j} z_j \right], \quad i = 2, 3, \dots, N \end{aligned}$$

Paso de la eliminación hacia atrás:

$$x_N = \frac{z_N}{u_{N,N}}$$

$$x_i = \frac{\left[z_i - \sum_{j=i+1}^N u_{i,j} x_j \right]}{u_{i,i}}, \quad i = N-1, N-2, \dots, 3, 2, 1$$

Hasta este punto no se ha utilizado el pivoteo con el fin de hacer mas sencilla la explicación. Sin embargo, el pivoteo es importante, por la misma razón que la eliminación de Gauss. Se debe recordar que el pivoteo en la eliminación de Gauss es equivalente a mezclar las ecuaciones en el conjunto. En forma matricial, quiere decir que los renglones de coeficientes se mezclan entre si junto con el término del lado derecho. Esto indica que el pivoteo se puede aplicar a la descomposición LU siempre que la forma de mezclar se aplique a los términos de ambos lados en la misma forma. Al hacer el pivoteo en la descomposición LU, se registran los cambios en el orden de los renglones. Después se aplica el mismo reordenamiento a los términos del lado derecho, antes de comenzar a resolver de acuerdo con los pasos a) y b) señalados arriba.

4.5.3) Eliminación de Gauss

La eliminación de Gauss es el método que se utiliza en forma más amplia para resolver un conjunto de ecuaciones lineales.

Un conjunto de N ecuaciones es de la forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1N}x_N &= y_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2N}x_N &= y_2 \\ &\vdots \\ a_{N1}x_1 + a_{N2}x_2 + a_{N3}x_3 + \dots + a_{NN}x_N &= y_N \end{aligned} \quad (4.29)$$

donde los a_{ij} son coeficientes, lo x_i son la incógnitas y los y_i son los términos llamados términos libres o independientes. En este caso, el número de incógnitas es igual al número de ecuaciones, que es la forma más usual de un conjunto de ecuaciones lineales.

Cuando al menos uno de los términos libres de la ecuación (4.29) es distinto de cero, se dice que el conjunto es no homogéneo. La eliminación de Gauss se aplica solo al caso de los conjuntos no homogéneos de ecuaciones. No siempre puede ser fácil la solución de un conjunto de ecuaciones lineales, debido al hecho de que quizá no tenga solución única, la solución calculada puede ser inexacta en el caso de un problema mal condicionado.

La eliminación de Gauss consiste en: a) la eliminación hacia delante, y b) la sustitución hacia atrás. La eliminación hacia delante se lleva a cabo de la manera siguiente.

La primera ecuación se multiplica por $a_{2,1}/a_{1,1}$ y se le resta a la segunda ecuación para eliminar el primer término de la segunda; de la misma forma, el primer término de las ecuaciones restantes, $i>2$, se elimina restando la primera ecuación multiplicada por $a_{i,1}/a_{1,1}$. Así, las ecuaciones deben de quedar de la siguiente forma:

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1N}x_N &= y_1 \\ a'_{22}x_2 + a'_{23}x_3 + \dots + a'_{2N}x_N &= y'_2 \\ &\vdots \\ a'_{N2}x_2 + a'_{N3}x_3 + \dots + a'_{NN}x_N &= y'_N \end{aligned} \tag{4.30}$$

donde

$$a'_{ij} = a_{ij} - (a_{i1}/a_{11})a_{1j}$$

Conviene observar que la primera ecuación no ha cambiado.

En seguida, el segundo término de cada una de las ecuaciones, desde la 4.31 hasta la última, $i>2$, se elimina restando la segunda ecuación multiplicada por $a'_{i,2}/a'_{2,2}$. Después de terminar este paso, se eliminan los terceros términos de las demás ecuaciones, de la cuarta a la última. Al finalizar este proceso de eliminación hacia delante, el conjunto de ecuaciones se verá de la forma siguiente:

$$\begin{aligned}
 a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1N}x_N &= y_1 \\
 a'_{22}x_2 + a'_{23}x_3 + \dots + a'_{2N}x_N &= y'_2 \\
 a''_{33}x_3 + \dots + a''_{3N}x_N &= y''_3 \\
 &\vdots \\
 a^{(N-1)}_{NN}x_N &= y^{(N-1)}_N
 \end{aligned}
 \tag{4.31}$$

Los términos principales de cada una de las ecuaciones anteriores reciben el nombre de pivotes.

El procedimiento de sustitución hacia atrás comienza con la última ecuación. Se obtiene la solución de x_N en la última ecuación:

$$x_N = y_N^{(N-1)} / a_{NN}^{(N-1)}$$

Sucesivamente,

$$\begin{aligned}
 x_{N-1} &= [y_{N-1}^{(N-2)} - a_{N-1,N}^{(N-1)}x_N] / a_{N-1,N-1}^{(N-2)} \\
 &\vdots \\
 x_1 &= \left[y_1 - \sum_{j=2}^N a_{1,j}x_j \right] / a_{11}
 \end{aligned}
 \tag{4.32}$$

Con esto se completa la eliminación de Gauss.

La eliminación de Gauss-Jordan es una variante de la eliminación de Gauss, comparte con ésta el proceso de eliminación hacia delante, pero difiere en el proceso hacia atrás, el cual es llamado eliminación hacia atrás.

4.5.4) Factorización de Cholesky de una matriz definida positiva simétrica

Cuando A es definida positiva simétrica, allí existe una única descomposición

$$A = U^T U \tag{4.33}$$

donde U es triangular superior con los elementos de la diagonal positivos. En esta sección se examina cómo la Factorización de (4.33) puede obtenerse directamente. También se deducen algunos límites que juegan un papel importante en el análisis del error de la eliminación simétrica. Se escribe la matriz dada A en la forma dividida:

$$A = \begin{vmatrix} \alpha & b^T \\ b & B \end{vmatrix} \quad (4.34)$$

dónde $\alpha = A_{11} > 0$ y B es simétrico de orden $n - 1$.

Ahora, $A = A^{(1)}$ pueden ser factorizados como sigue:

$$A = \begin{vmatrix} \alpha^{1/2} & 0^T \\ b/\alpha^{1/2} & I \end{vmatrix} \begin{vmatrix} 1 & 0^T \\ 0 & A^{(2)} \end{vmatrix} \begin{vmatrix} \alpha^{1/2} & b^T/\alpha^{1/2} \\ 0 & I \end{vmatrix} \quad (4.35)$$

donde 0 es el vector de la columna nulo, y

$$A^{(2)} = B - \frac{b}{\alpha^{1/2}} \frac{b^T}{\alpha^{1/2}} \quad (4.36)$$

es una matriz de orden $n - 1$.

Claramente, $A^{(2)}$ es simétrica; también se puede demostrar que es definida positiva. Para este propósito, se utiliza la forma (4.34) para A y se escribe la ecuación con:

$$y = \begin{vmatrix} y_1 \\ z \end{vmatrix} \quad (4.37)$$

donde z es cualquier vector no nulo de orden $n - 1$. El resultado es:

$$y^T Ay = \alpha y_1^2 + 2y_1 z^T b + z^T Bz > 0 \quad (4.38)$$

Si se toma el $y_1 = -z^T b/\alpha$ y se usa la ecuación (4.36):

$$z^T A^{(2)} z > 0 \quad (4.39)$$

y como z es cualquier vector, esto demuestra que $A^{(2)}$ es definida positiva.

Como $A^{(2)}$ es definido simétrico positivo, puede ser factorizado de la misma manera que A , ecuación (4.35). Se obtiene una matriz definida simétrica positiva $A^{(3)}$ de orden $n - 2$, la cual puede, a su vez, ser factorizada. El procedimiento se repite hasta que se obtiene una matriz $A^{(n)}$ de orden 1. Todas las matrices de la secuencia $A^{(1)} \equiv A, A^{(2)}, \dots, A^{(n)}$ son definidas positivas simétricas, y $A^{(k)}$ es de orden $n - k + 1$. Al final del procedimiento se obtiene la siguiente expresión para A :

$$A = L_1^c L_2^c \dots L_n^c U_n^R \dots U_2^R U_1^R \quad (4.40)$$

La aplicación práctica del algoritmo de Cholesky para matrices dispersas es, sin embargo, bastante diferente. Se discutirá en la siguiente sección. La discusión en esta sección fue dada porque se presta a encontrar los límites para los elementos de las matrices $A^{(k)}$. Estos límites son a su vez necesarios para el análisis del error de factorización de Cholesky, y para la justificación de la propiedad más importante de algoritmo de Cholesky: ese pivoteo diagonal es numéricamente estable para matrices simétricas definidas positivas.

Ahora se tienen que encontrar los límites de los valores absolutos de los elementos de matrices $A^{(k)}$. $A^{(k)}$ es definida simétrica positiva:

$$|A_{ij}^{(k)}| \leq a_k \quad \text{para todo } i, j \quad (4.41)$$

donde a_k es el elemento diagonal más grande de $A^{(k)}$. Ahora, de la Ecuación (4.36), es evidente que los elementos de la diagonal de cada $A^{(k)}$, los cuales son positivos, son construidos substrayendo los números non-negativos de los elementos diagonales correspondientes de $A^{(k-1)}$, que también son todos positivos. Así:

$$a_1 \geq a_2 \geq \dots \geq a_n, \quad (4.42)$$

donde el a_k es el elemento diagonal más grande de $A^{(k)}$. Finalmente, se puede declarar el hecho que $a \equiv a_i$ es el límite superior para todos los elementos de todo $A^{(k)}$:

$$|A_{ij}^{(k)}| \leq a \quad \text{para todo } i,j,k \quad (4.43)$$

La ecuación (4.43) dice que ningún elemento de cualquier matriz reducida $A^{(k)}$ puede exceder en magnitud el valor de a , de la diagonal de elementos más grande de la matriz original A .

4.5.5) Implementación práctica de la factorización de Cholesky

En la práctica, A es simétrica, sólo su triángulo diagonal y superior se guarda en la memoria. Las matrices $A^{(k)}$ nunca se construyen explícitamente. El factor triangular superior que U se desarrolla directamente en el almacenamiento separado, en lugar de sobre escribiendo en A , porque normalmente es menos dispersa que A y requiere una estructura de datos especial. Aunque no se puede hablar del método de Cholesky como "eliminación", porque de hecho, no se está eliminando nada, se verá que el algoritmo es casi idéntico a la eliminación de Gauss.

Viendo la ecuación (4.35) comprendemos que, una vez la raíz cuadrada de $A_{11}^{(1)}$ se obtiene y los elementos restantes de la fila 1 son divididos por este valor, el renglón 1 de U está completo. Entonces, se pueden calcular los elementos del renglón 1 de $A^{(2)}$, usando la ecuación (4.36), e inmediatamente se obtiene la raíz cuadrada de $A_{11}^{(2)}$ y se dividen los elementos restantes de la primera fila de $A^{(2)}$ por este valor; esto nos da la segunda fila completa de U . Note que los elementos de $b^T/\alpha^{1/2}$ ya se han procesado y se ha guardado como la primera fila de U ; así sólo multiplicaciones y subtracciones son necesarias cuando se usa la ecuación (4.36). También se nota que este paso es precisamente análogo a la eliminación de Gauss de A_{21} , utilizando A_{11} como pivote, aunque A_{11} no es unidad en este caso.

Ahora, se procede a el renglón 3 de A , y utilizando la ecuación (4.36) para obtener la fila 2 de $A^{(2)}$. Esto es análogo a la eliminación de A_{31} . Luego, para la fila 3 de A , se

realizan los procedimientos necesarios para obtener la fila 1 de $A^{(3)}$; esto simplemente es análogo a la eliminación de A_{32} . Tomando la raíz cuadrada del elemento diagonal obtenido y dividiendo los elementos restantes de la fila por este valor. Esto produce la fila 3 completa de U en su forma final. Procediendo de esta manera se obtiene U exactamente de la misma manera como si se hubiera realizado la eliminación de Gauss, la única diferencia es que, al final de cada paso, en lugar de guardar el recíproco del elemento diagonal y multiplicando los elementos fuera de la diagonal por ese valor, se guarda el recíproco de la raíz cuadrada del elemento diagonal y se multiplican los elementos de fuera de la diagonal de la fila por esta raíz cuadrada. La tabla de factores, representando la factorización de Cholesky (4.33), consiste en el triángulo superior de U y el recíproco de los elementos diagonales de U . No se guarda ni se opera en los elementos del triángulo mas bajo de A . Si llamamos $D^{1/2}$ diagonal de la matriz formada con los elementos (positivos) diagonales de U , podemos escribir $U = D^{1/2}U'$ donde U' es triangular superior con diagonal de unos (U diferente); la ecuación (4.33) puede escribirse entonces:

$$A = U'^T D U' \quad (4.44)$$

donde $D = (D^{1/2})^2$ es diagonal con los elementos de la diagonal positivos. Ahora es fácil comparar la factorización de Cholesky con la eliminación de Gauss. De hecho, la ecuación (4.44), se obtuvo por la factorización de Cholesky. La diagonal de la tabla de Cholesky es la matriz $D^{-1/2}$, el triángulo superior es el triangulo superior de U . La diagonal de Gauss es D^{-1} , el triángulo superior es de U' .

5) Respuesta Sísmica 3D de la Estructura del Sótano Profunda de la Cuenca de Granada (Sudeste de España)

5.1) Resumen

En este capítulo, se aplica el IBEM a un problema real, en este caso, se estudia la Cuenca de Granada. Se explica claramente el método y la discretización, que se vio en el capítulo 3 y se realiza todo el procedimiento con datos reales que se analizan para saber la respuesta sísmica que la estructura de la cuenca.

Esta Cuenca se localiza en el sudeste de España, en el sector central de las Cordilleras de Betic, en un área de alto riesgo sísmico en la Península Ibérica. En este trabajo de investigación se utiliza el método indirecto de elementos frontera (IBEM) para calcular la respuesta sísmica tridimensional (3D) de la estructura profunda de la Cuenca de Granada para incidencia de ondas planas P y S que vienen del sur. Se han analizado los resultados en frecuencia y dominios de tiempo, y para el rango de frecuencias que se han

podido calcular (entre 0 y 0.312 Hz). Generalmente hablando, los modelos de amplificación observados de los desplazamientos para las ondas incidentes SH están los más altos en esas situaciones dónde la cuenca es más profunda. Para las ondas P y SV, las fotografías ("snapshots") muestran los efectos de amplificación de componentes radiales y verticales, cuando las ondas atraviesan la Cuenca de Granada.

En los recientes años, se han desarrollado los códigos numéricos para tratar con la propagación de la onda sísmica en modelos tridimensionales (3D). Las simulaciones más realistas son aquéllas de diferencias finitas. En particular, geometrías 3D han sido consideradas para algunas cuencas sedimentarias alrededor del mundo.

En este trabajo se ha estudiado la respuesta sísmica de la estructura profunda de la Cuenca de Granada (Fig.5.1) que se localiza en una de las regiones con el nivel de riesgo más alto en la Península Ibérica como se muestra en el código de construcción sísmico español (Norma de la Construcción Sismorresistente NCSE-94, 1995). Los registros históricos disponibles indican que a finales del año 2000, diferentes partes del sur de la Península Ibérica han sido afectados por terremotos mayores. En la reciente historia, los epicentros de la región para terremotos con magnitudes ≥ 4 del año 1970 a 2001 se presentan en la Fig.5.1.

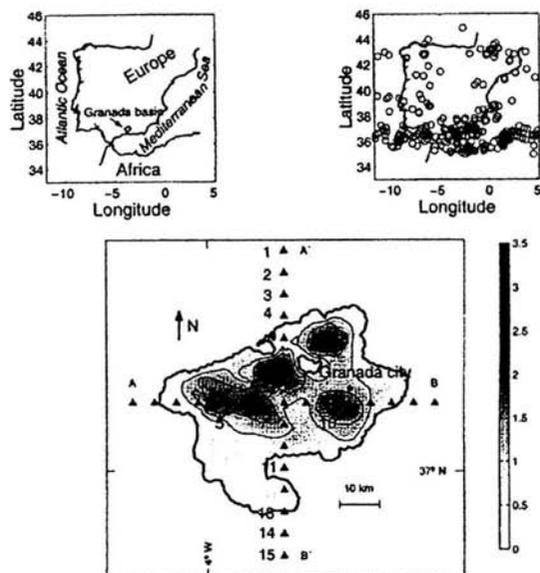


Fig. 5.1 (Arriba a la izquierda): Mapa mostrando la localización de la cuenca de Granada. (Arriba a la derecha): mapa mostrando los epicentros (círculos) de la región terremotos con magnitudes ≥ 4 en el periodo de 1970-2001. (Abajo): Profundidad del sótano de la cuenca de Granada, con la localización de receptores en los perfiles A-B y A'-B'. El contorno tiene un intervalo de 600m; el sombreado oscuro indica mayor profundidad del límite del sótano-sedimentario. Las unidades de la profundidad en la escala de la barra son en km.

5.2) La Cuenca de Granada

La Cuenca de Granada se localiza en el sureste de España, en el sector central de las Cordilleras Betic. El sur de España tiene el riesgo sísmico más alto en la Península Ibérica, como el mostrado en el mapa de la máxima aceleración horizontal para un regreso de periodo de 500 años (Norma de la Construcción Sismorresistente NCSE-94, 1995). La contribución de sismicidad para riesgo es una consecuencia de la colisión en la litosfera entre las placas africanas y Euroasiáticas en una escala regional. La Cuenca de Granada tiene el nivel más alto de actividad microsísmica (magnitud < 3.5) en la Península Ibérica. La mayoría de la energía sísmica es liberada por los terremotos con las magnitudes < 5.0 .

La estructura 3D del fondo de la Cuenca de Granada se estableció por medio de un análisis e interpretación de gravedad y datos de sísmico-reflexión. En esta Cuenca (Fig. 1),

hay depósitos centrales importantes con las profundidades >3000 m. Las dimensiones de las características superficiales de los depósitos sedimentarios son ~ 60 Km. X 45 km. Desde el punto de vista de riesgo sísmico, los efectos locales en la Cuenca de Granada han sido estudiados usando el microtemblor utilizando un evento pequeño, con magnitud de M_w 5, como una función empírica de Green para obtener registros sintéticos de un posible terremoto de magnitud M_w 7.0 en varias locaciones en la cuenca. Las condiciones geológicas del sitio fueron reflejadas en sismogramas sintéticos en que la aceleración del suelo máximo alcanzó valores de ~ 250 cm./seg.². El $S_{a_{max}}$ del sismograma simulado, ~ 0.8 , también señaló la importancia de efectos del sitio en la Cuenca de Granada en el movimiento del suelo inducido por los terremotos moderados.

5.3) Método y Discretización

Se utilizó el método indirecto de elemento de frontera (IBEM) para procesar la respuesta sísmica de la estructura profunda de la Cuenca de Granada. En el IBEM, los problemas de difracción y refracción de ondas elásticas son formulados en términos de densidades de fuerza de frontera. Las condiciones de frontera de continuidad de desplazamientos y tracciones entre las regiones y aquellos de tracciones nulas en superficies libres, una vez discretizado, llevan a un sistema de ecuaciones lineales en que las incógnitas son las densidades de fuerza de las fuentes del límite. Por lo tanto, el método puede considerarse como una realización numérica de principio de Huygens en que las fuentes secundarias se localizan en los límites de la superficie. En esta técnica la discretización de superficies es realizada utilizando elementos circulares como se explicó anteriormente.

Se presentaron varios ejemplos para utilizar el IBEM y resolver el sistema lineal de ecuaciones con la clásica descomposición LU. Sin embargo, tratando con los problemas 3D más grandes y más realistas, surgen algunas dificultades adicionales. Para guardar una proporción razonable entre la longitud de onda mínima y el tamaño de elementos, la dimensión de la matriz del coeficiente crece aproximadamente con el cuadrado de frecuencia en estos cálculos, y por consiguiente los requerimientos del cómputo pueden ser demasiado grandes para los problemas de interés práctico. En un artículo, Ortiz-Alemán

(1998), superó el problema utilizando el criterio del umbral para convertir la matriz llena en una dispersa y usando el método del gradiente biconjugado junto con un esquema iterativo para resolver el sistema lineal de ecuaciones. Por otro lado, el uso de las funciones analíticas de Green de un espacio completo homogéneo, como se hace en este trabajo, implícitamente limita la técnica cuando la respuesta para las frecuencias más altas se quiere en una cuenca de capa 3D. No obstante, con esta técnica y con las funciones de Green apropiadas, es todavía posible, por ejemplo, para tener en cuenta una estructura de capa o un medio con una velocidad gradiente, en las cuencas aluviales con geometría 2D con un gradiente vertical de velocidad.

Con el sistema de ecuaciones lineal obtenido en el IBEM se trata algo diferente. Se utiliza una matriz dividida en lugar de la matriz completa. Se reordenan las filas y columnas de la matriz para tener matrices de bloque y encontrar que el sistema tiene una estructura particular que puede construirse fácilmente como:

$$\begin{bmatrix} \mathbf{T}_{M_E M_E} & \mathbf{T}_{M_E L_E} & \mathbf{0}_{M_E L_R} & \mathbf{0}_{M_E K_R} \\ \mathbf{T}_{L_E M_E} & \mathbf{T}_{L_E L_E} & \mathbf{T}_{L_E L_R} & \mathbf{T}_{L_E K_R} \\ \mathbf{G}_{L_R M_E} & \mathbf{G}_{L_R L_E} & \mathbf{G}_{L_R L_R} & \mathbf{G}_{L_R K_R} \\ \mathbf{0}_{K_R M_E} & \mathbf{0}_{K_R L_E} & \mathbf{T}_{K_R L_R} & \mathbf{T}_{K_R K_R} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\varphi}_{M_E} \\ \boldsymbol{\varphi}_{L_E} \\ \boldsymbol{\varphi}_{L_R} \\ \boldsymbol{\varphi}_{K_R} \end{Bmatrix} = \begin{Bmatrix} \mathbf{0}_{M_E} \\ -\mathbf{t}_{L_E} \\ -\mathbf{u}_{L_R} \\ \mathbf{0}_{K_R} \end{Bmatrix} \quad (5.1)$$

en que cada término en la expresión representa un bloque de submatriz o subvector. La submatriz $T_{I_A J_B}$ y $G_{I_A J_B}$ están relacionados con la tracción y desplazamiento las funciones de Green, respectivamente, donde I y J puede ser M (el número de elementos en la superficie libre del medio-espacio E), L (el número de elementos a la interfaz compartido entre el medio-espacio E e inclusión R), K (el número de elementos en la superficie libre de la inclusión R), y en qué A y B se refieren a los dominios E o R. Los subvectores t_{I_A} y u_{I_A} son las tracciones y desplazamientos debidos al campo libre, es decir, aquellos producidos por el campo de la onda en la ausencia de cualquier irregularidad que incluye ondas incidentes y reflejadas por la superficie libre del espacio-medio. Los bloques $O_{M_E L_R}$, $O_{M_E K_R}$, $O_{K_R M_E}$, y $O_{K_R L_E}$ son el bloques cero, y los subvectores O_{M_E} y O_{K_R} son vectores cero. Es fácil ver que al aplicar la matriz el acercamiento de la condensación estático, el

sistema lineal de ecuaciones desplegado previamente se convierte en varios sistemas lineales pequeños de ecuaciones que pueden resolverse usando menos memoria que el sistema original. Sacando los bloques de ceros de los procesos, la solución final gasta un tiempo menor pero manteniendo la misma exactitud. El acercamiento de la condensación estática es basado en una factorización implícita de la matriz del coeficiente aprovechándose de la situación de los bloques de ceros cerca del ángulo superior derecho e inferior izquierdo del sistema. Considerado las primera y cuarta filas del bloque del sistema, se puede escribir

$$[\mathbf{T}_{M_E M_E}][\mathbf{P}_1] = [\mathbf{T}_{M_E L_E}] \quad (5.2)$$

$$[\mathbf{T}_{K_R K_R}][\mathbf{P}_2] = [\mathbf{T}_{K_R B L_R}] \quad (5.3)$$

donde $[\mathbf{P}_1]$ y $[\mathbf{P}_2]$ son matrices del intermedio desconocidos que pueden resolverse como un juego de ecuaciones simultáneas. Entonces,

$$\begin{aligned} [\mathbf{Q}_{11}] &= [\mathbf{T}_{L_E L_E}] - [\mathbf{P}_1][\mathbf{T}_{L_E M_E}] \\ [\mathbf{Q}_{12}] &= [\mathbf{T}_{L_E L_R}] - [\mathbf{P}_2][\mathbf{T}_{L_E K_R}] \\ [\mathbf{Q}_{21}] &= [\mathbf{G}_{L_R L_E}] - [\mathbf{P}_1][\mathbf{G}_{L_R M_E}] \\ [\mathbf{Q}_{22}] &= [\mathbf{G}_{L_R L_R}] - [\mathbf{P}_2][\mathbf{G}_{L_R K_R}] \end{aligned} \quad (5.4)$$

donde $[\mathbf{Q}_{11}]$, $[\mathbf{Q}_{12}]$, $[\mathbf{Q}_{21}]$ y $[\mathbf{Q}_{22}]$ son matrices intermedias que pueden ser fácilmente procesadas y produce un nuevo sistema compacto de ecuaciones lineales, se reduce a un cuarto el tamaño del sistema original que puede expresarse como sigue:

$$\begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \begin{Bmatrix} \boldsymbol{\varphi}_{L_E} \\ \boldsymbol{\varphi}_{L_R} \end{Bmatrix} = \begin{Bmatrix} -\mathbf{t}_{L_E} \\ -\mathbf{u}_{L_R} \end{Bmatrix} \quad (5.5)$$

Después de resolver este sistema, $\boldsymbol{\varphi}_{L_E}$ y $\boldsymbol{\varphi}_{L_R}$ son los vectores de densidad de fuerza conocidos. Así que, el paso final en el algoritmo consiste en el proceso de los únicos vectores de densidad de fuerza desconocidos, $\boldsymbol{\varphi}_{M_E}$ y $\boldsymbol{\varphi}_{K_R}$ como sigue:

$$\begin{aligned} \{\Phi_{M_E}\} &= -[P_1]\{\Phi_{L_E}\} \\ \{\Phi_{K_R}\} &= -[P_2]\{\Phi_{L_R}\} \end{aligned} \quad (5.6)$$

En contraste con el acercamiento del gradiente del umbral-biconjugado, eso trabaja bien para una frecuencia relativamente baja, este esquema no tiene ningún error de truncamiento implícito, y así, está estable en las frecuencias más altas por valles aluviales y topografías. Los procesos se hacen en un programa de la computadora muy eficaz en FORTRAN 90 con memoria dinámica que no gasta ningún recurso extra. De esta manera, es posible procesar, en la misma computadora y con el mismo problema 3D, la solución que corresponde a aproximadamente dos veces la frecuencia calculada con la forma clásica que usa la matriz llena en el IBEM.

Es más, se discretizó dinámicamente toda la superficie en el modelo. Para hacer esto, se procesó una reja discreta para cada frecuencia analizada, de tal manera que siempre se tiene por lo menos 4.5 elementos del límite por la longitud de ondas S en cada región. Para la superficie irregular, se construyó una reja formada por elementos rectangulares que cubren la superficie completa. Después de esto, cada elemento rectangular planar es asumido para hacer un círculo con la misma superficie como eso del elemento rectangular. Por otro lado, en los límites llanos, eso está, en las superficies libres del espacio-medio y la cuenca, los elementos son distribuidos con las distancias constantes del origen del sistema de la coordenada local, formando los anillos adyacentes diferentes. La superficie total de un anillo es igual a la suma de las superficies de todos los elementos del anillo. La superficie libre llana de la cama de roca fue discretizada hasta una distancia radial de 55 Km. del origen de nuestro sistema coordenado, el cual se localiza en el punto común que es compartido por ambos perfiles A-B y A'-B', en el centro de la cuenca. En la Fig.5.2, puede verse un ejemplo de la distribución de elementos de frontera del sótano para una frecuencia de 0.149 Hz.

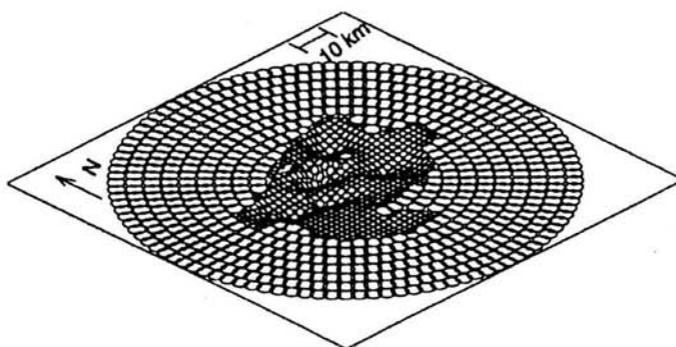


Fig. 5. 2 Ejemplo de discretización del límite del sótano-sedimentario para la frecuencia de 0.149 Hz.

6) Resultados

A continuación se presentan los resultados obtenidos al aplicar el Método Indirecto de Elementos de Frontera al problema planteado en el capítulo 5. En este capítulo se integra todo lo planteado en capítulos anteriores, los conceptos de sismología son útiles para poder interpretar los resultados, para resolver el problema se utilizó el IBEM y para la matriz, se utilizó el criterio de truncamiento de matrices expuesto en el capítulo 4.

En este trabajo se ha utilizado el IBEM en el dominio de frecuencia, y se ha procesado la respuesta de la cuenca para 77 frecuencias de 0 a 0.312 Hz, para ondas planas incidentes P, SV, y SH que propagan del sur. La frecuencia de 0 Hz corresponde al caso donde ninguna inclusión existe, es decir, la solución del espacio-medio.

Los primeros resultados obtenidos, muestran la influencia de la estructura profunda de la Cuenca de Granada en los modelos de amplificación para ondas planas SH incidentes en varias frecuencias. Se seleccionaron ángulos incidentes de 0° y 60° con respecto a la vertical. Estos resultados pueden verse en las Fig. 6.1 y 6.2, respectivamente, para el E-W (movimiento transversal), N-S (movimiento radial), y componentes verticales de desplazamiento. En ambas figuras los modelos de amplificación se despliegan a las mismas frecuencias por causa de la comparación, (0.149, 0.198, 0.248, y 0.312 Hz). Se puede

observar claramente en ambas figuras, incluso para las frecuencias relativamente bajas, las amplificaciones producidas por la estructura profunda de la Cuenca de Granada. De hecho, para el componente de E-W, la cuenca produce las amplificaciones grandes para las frecuencias escogidas. En particular, para una incidencia vertical (Fig. 6.1), la amplificación alcanza a casi 10 veces la amplitud de la onda de la fuente incidente, en la frecuencia de 0.312 Hz, a las varias situaciones de la cuenca. Para el mismo componente de desplazamiento fuera de la cuenca, la amplitud del movimiento observado es aproximadamente dos veces el incidente, como esperado. Para N-S y los movimientos verticales, el modelo de amplificación es muy interesante. Generalmente hablando, se producen los niveles máximos del desplazamiento a esos lugares dónde las profundidades de la cuenca son mayores, es decir, los conocidos “depocenters” (centros de depósito). Por otro lado, estos aumentos de los niveles máximos con la frecuencia para ambas incidencias, 0° y 60° . Para incidencia vertical y el componente de N-S, las amplitudes del desplazamiento alcanzan valores más grandes que 5.

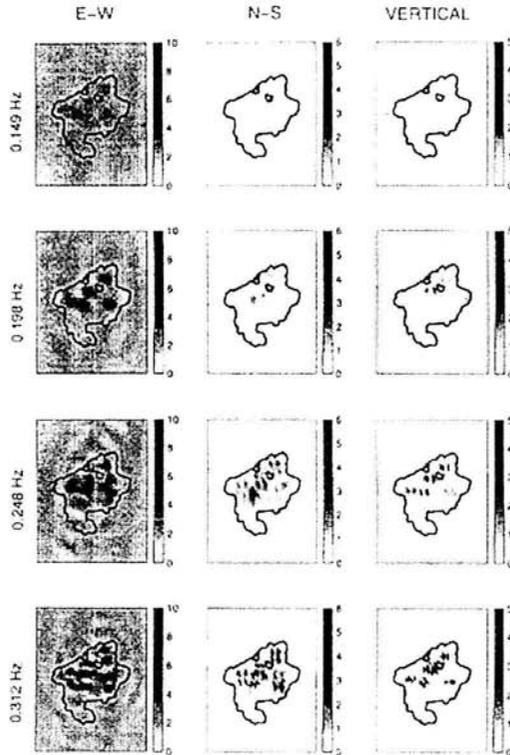


Fig 6.1 Amplificaciones, relativas para la amplitud de la onda incidente, de los tres componentes de desplazamiento en varias frecuencias en la cuenca de Granada para una onda SH viniendo desde el sur con la incidencia vertical. El contorno de la profundidad-cero de la cuenca es mostrado para referencia.

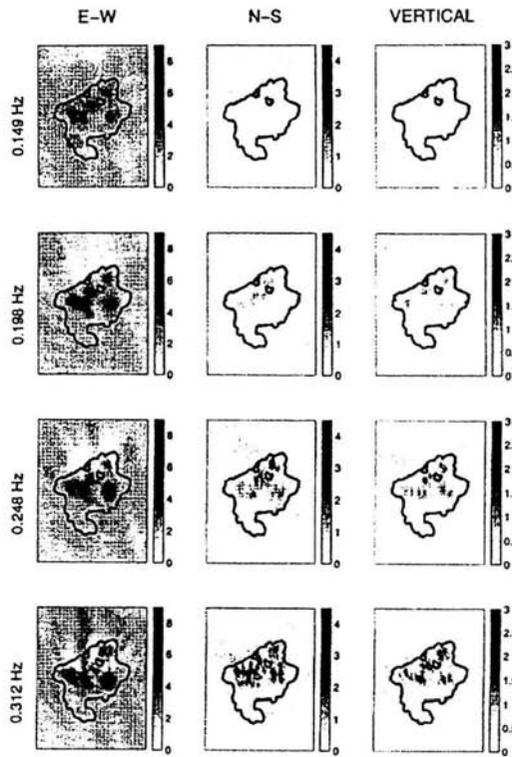


Fig. 6.2 La misma que la Fig. 6.1, pero para una onda SH viniendo desde el sur y con un ángulo incidente de 60° con respecto a la vertical.

Estos efectos que no pueden obtenerse usando modelado numérico 1D o 2D, corresponden a la naturaleza 3D del sótano profundo de la Cuenca de Granada. La Fig. 6.3 muestra la comparación de la respuesta 1D y 3D para ondas SH incidentes verticalmente a los varios receptores a través de la cuenca. La solución de 1D corresponde al modelo de una sola capa encima de un espacio-medio, con el mismo espesor como el de la profundidad del sótano en cada estación. Las estaciones seleccionadas (5,7, y 10) en el perfil A-B está al lado de los “depocenters”, con las profundidades del sótano de 2295 m, 1908 m, y 1744 m, respectivamente. La respuesta 3D de la Cuenca de Granada es más compleja que la 1D, incluso a la frecuencia baja, como puede observarse en Fig. 6.3. En el desplazamiento transverso, los niveles de amplitud de las respuestas de 1D y 3D a los receptores 7 y 10 son bastante similares, pero por otro lado, se cambia la frecuencia

resonante en la solución completa 3D en estación 7 ligeramente a 0.2338 Hz. En estación 5, la respuesta está diferente en el nivel de amplificación y el volumen de frecuencia. De hecho, la solución de 1D predice una amplificación de 4.91 a una frecuencia de 0.1743 Hz, considerando que la respuesta 3D completa tiene una amplitud cerca de 7 a 0.2480 Hz.

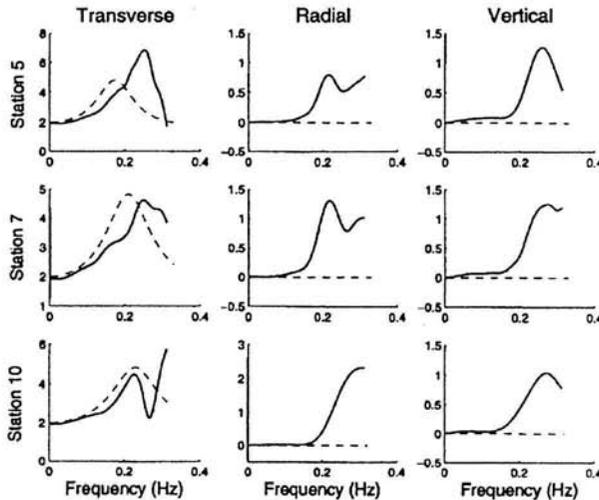


Fig. 6.3 La respuesta sísmica de 1D (línea punteada) y 3D (línea continua) para las ondas verticalmente incidentes SH en las estaciones 5, 7 y 10 de el perfil A-B mostrada en la Fig.5.1.

Para los componentes radiales y verticales, la respuesta es completamente diferente. Considerando que las amplitudes en los modelos de 1D son iguales a cero en todos los casos, para el caso de 3D, estos desplazamientos sólo están comparables a la amplitud de la onda incidente en las más bajas frecuencias. Es más, en el movimiento radial en la estación 10, la amplificación 3D está dos veces el incidente uno en las frecuencias cercanas a 0.3 Hz.

6.1) Análisis de Propagación de Onda

En esta sección, se procesa la respuesta sísmica en el dominio de tiempo usando un marcador de onda incidente permitido de Ricker (Ricker 1977). En estos procesos, se utiliza un periodo característico de $t_p = 6.5$ seg. y un retraso de tiempo de $t_s = 40$ seg. Este pulso representa el movimiento de la entrada que tiene una forma de onda que corre de $u_{\max} = -0.4432$ al $u_{\max} = 1$ unidad de desplazamiento, como mostrado en Fig.6.4. Con el objetivo de estudiar las características principales de propagación de la onda elástica dentro del modelo de la Cuenca de Granada, se utilizaron ondas planas P y SV que llegan desde el sur, con un ángulo incidente de 30° relativo al vertical.

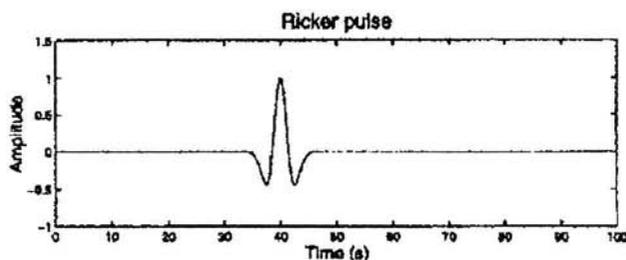


Fig.6.4 Pulso de Ricker incidente utilizado para este estudio. El periodo característico es de $t_p = 6.5$ seg. y el tiempo de retraso, de $t_s = 40$ seg.

En el caso de la fuente de la onda P incidente, se procesó la respuesta de tiempo de varias frecuencias estudiadas usando la síntesis de Fourier. Aproximadamente 18 seg. del desplazamiento radial (N-S) se representan en la Fig.6.5 por medio de fotografías en los momentos diferentes. Note que los desplazamientos observados en la superficie libre para estos componentes están en el rango de -1.8 a 2 unidades de desplazamiento, como mostrado en la escala de la barra, considerando que el movimiento incidente se produjo por el pulso de Ricker en el rango compuesto de N-S va del $u_{\min}(\text{N-S}) = \sin 30 \cdot u_{\min} = -0.2216$ a $u_{\max}(\text{N-S}) = \sin 30 \cdot u_{\max} = 0.5$ unidades del desplazamiento, es decir, en algunos tiempos y posiciones, la amplificación producida es ~ 8 . La onda P llega después de 36.8826 seg. a la parte profunda de la cuenca al sur. Estos primeros desplazamientos, representados con el

negativo (las sombras oscuras) las amplitudes dentro y fuera de los sedimentos, no produzcan efecto notable. No obstante, después de un tiempo de 40.0898 seg., la onda directa empieza a ser amplificada, como puede verse (las sombras oscuras dentro de la cuenca) y en la fotografía siguiente (las sombras ligeras dentro de la cuenca). Mientras la onda incidente está viajando a través de la cuenca, el movimiento mas fuerte de tierra se encuentra en sitios dónde la cuenca es más profunda.

Este efecto es más prominente para el componente radial que para el vertical (no mostrado). Después la onda directa se aparta de la parte norte del modelo ($t = 46.5041$ seg.), la energía mecánica que resta atrapada dentro de la cuenca continúa produciendo desplazamientos en la superficie libre. Como antes de los desplazamientos mayores para $t > 46.5041$ seg. son observados de nuevo a las situaciones sobre las partes más profundas de la cuenca y son más significantes en el componente N-S, presentado en la Fig.6.5.

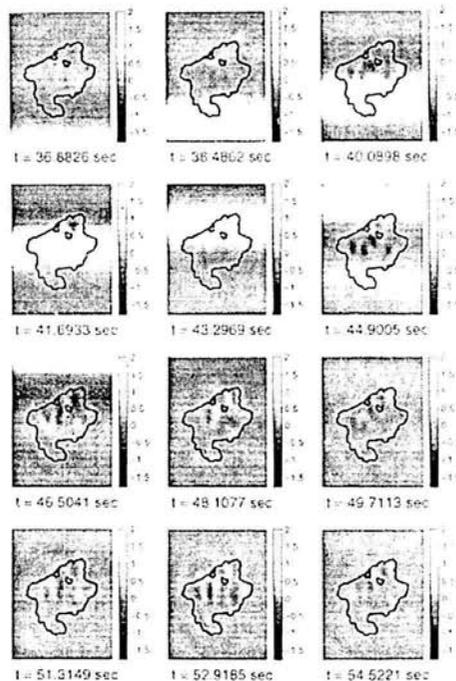


Fig.6.5 Fotografía del desplazamiento radial (N-S) de la propagación de onda en la cuenca de Granada para una simulación con una onda P de el sur y con un ángulo incidente de 30° con respecto a la vertical. La luz (oscura) sombreada describe el desplazamiento de partícula positivo (negativo). El contorno de profundidad cero de la cuenca se muestra para referencia.

Los resultados de los desplazamientos radiales producidos por la fuente de la onda incidente SV son mostrados en la Fig.6.6. Las fotografías se dan durante los mismos tiempos usados para la simulación de la onda P. Para esta onda incidente, los desplazamientos observados a la superficie libre para el componente de N-S están en el rango de -1.8 a 2.4 unidades de desplazamiento, considerando que el movimiento incidente en el componente de N-S va de $u_{\min}(\text{N-S}) = \cos 30 \cdot u_{\min} = -0.3838$ al $u_{\max}(\text{N-S}) = \cos 30 \cdot u_{\max} = 0.8660$ unidades de desplazamiento. El movimiento empieza a ser amplificado dentro de la cuenca después de 40.0898 seg., y la amplificación es más importante en los tiempos siguientes cuando la onda directa cruza el modelo de la Cuenca de Granada. Esto se ha observado en los componentes radiales y verticales. Después de esto, la energía se entrapa dentro de la estructura, produciendo desplazamientos que pueden observarse de $t = 46.5041$ seg. y los tiempos subsecuentes, cuando las salidas de la onda directa de la parte norte de la cuenca.

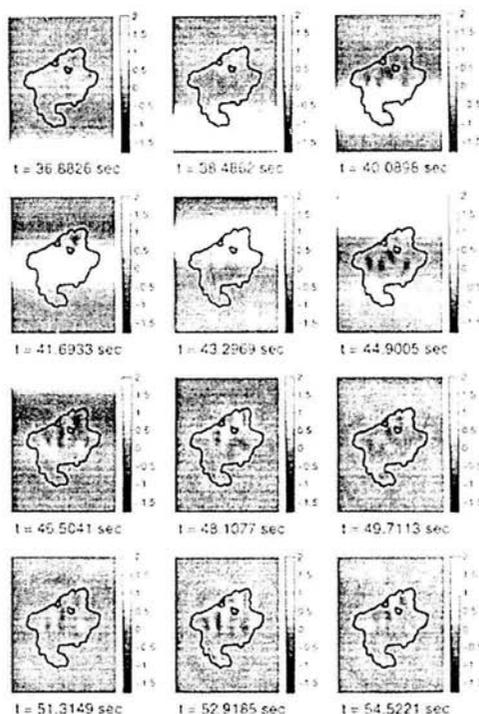


Fig.6.6 La misma que la Fig.6.4, pero para una onda incidente SV.

Una cantidad pequeña de energía se genera a los bordes de la cuenca. Esto no puede observarse bien en las fotografías de la Fig.6.6, porque las amplitudes de estos desplazamientos son bajas comparadas con el movimiento producido por el campo de la onda incidente. Esta cantidad pequeña de energía puede observarse bien en Fig.6.7, dónde se presentan los sismogramas sintéticos de los componentes de desplazamiento vertical (V), radial (R), y tangencial (T) en los receptores a lo largo de los perfiles A-B y A' - B' de la Fig.5.1. El movimiento de tierra dentro de la cuenca sedimentaria es generalmente más complejo que el observado en piedra dura, como se muestra en el componente tangencial de ambos perfiles dónde no existe onda directa. Comparando este componente en ambos perfiles, se observa que el que corresponde al perfil A-B presenta un modelo más complicado que el desplazamiento de perfil A'-B'. Esto probablemente es porque el perfil A-B cruza varios “depocenters” de la Cuenca de Granada, considerando que el perfil A'-B' cruza sólo uno. También note que hay más estaciones dentro de la cuenca en el perfil A-B. No obstante, las características de las fases propagando dentro de la cuenca que puede verse en los componentes tangenciales son diferentes. Considerando que en A-B la propagación de una fase identificada como una onda de Rayleigh se observa, la onda correspondiente que viaja en A'-B' puede identificarse como una onda de Love. Por otro lado, la emisión de ondas radiada fuera de la Cuenca de Granada también se observa en los componentes verticales y radiales del perfil A'-B'. Esas ondas propagadas con más energía en la dirección delantera (del sur al norte) de la onda del cuerpo incidente. Éstos pueden identificarse como las ondas de Rayleigh por medio de los resultados dados (Fig.6.8), dónde se presentan los “polarigrams” en el plano radial-vertical para la incidencia de ondas SV en 15 receptores en el perfil A'-B. Estos “polarigrams” despliegan la variación en los vectores del desplazamiento con tiempo. La onda de Rayleigh, muestra la poca energía, que puede identificarse con las llegadas mas tarde, con su característico movimiento elíptico y retrogrado. Éste no sólo es el caso para un espacio-medio, también para el modo fundamental de Rayleigh en una capa.

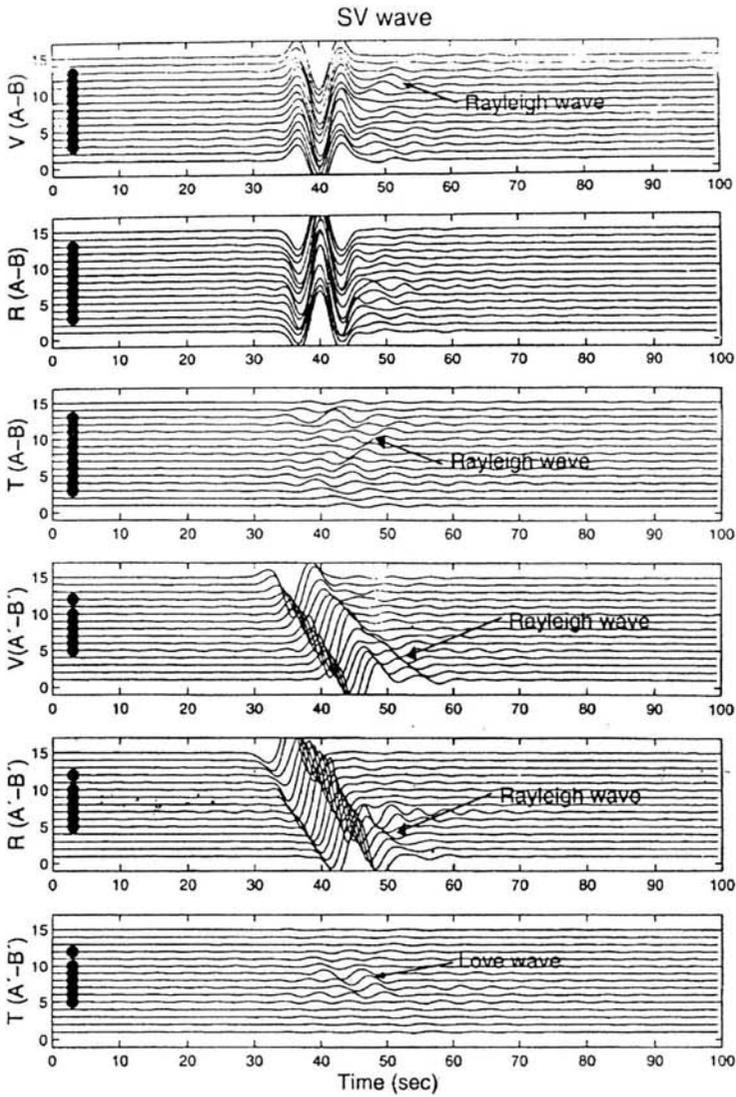


Fig. 6.7 Los sismogramas de los componentes de la vertical (V), radial (R) y tangencial (T) de desplazamientos en los 15 receptores del perfil A'-B' y los 15 receptores del perfil A-B. Los sismogramas que tienen un círculo sólido son localizados en la cuenca de Granada.

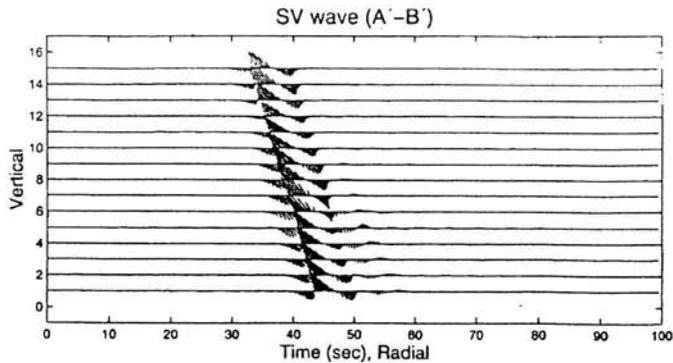


Fig. 6.8 Polarigrams en el plano radial-vertical de la incidencia de ondas SV en los 15 receptores del perfil A'-B'.

6.2) Conclusiones

Se ha procesado la respuesta sísmica de la estructura de sótano 3D del modelo de la Cuenca de Granada utilizando el IBEM para la incidencia de ondas P y S planas. En este estudio, se ha resuelto el sistema lineal de ecuaciones utilizando un método de matriz de partición que es diferente de la descomposición de LU clásica. La ventaja del método de matriz de partición utilizada aquí, es que es posible calcular, en la misma computadora y con la misma configuración, la solución que corresponde a aproximadamente dos veces la frecuencia procesada con el LU clásico.

Se han analizado los resultados en frecuencia y dominios de tiempo, y para el rango de frecuencias que pudieron computar, se ha observado que los efectos del sitio más pertinentes son debidos a las amplificaciones locales producidas dentro de los sedimentos. Para las ondas de SH, la amplificación del factor alcanza valores de casi 10 veces la amplitud de la onda incidente para un poco de frecuencias y locaciones específicas. En el otro caso, para el periodo característico considerado en este trabajo, para la onda P incidente del pulso de Ricker, la amplificación llega a algunos tiempos específicos y posiciones alrededor de 8. En general, para las ondas planas incidentes consideradas en este trabajo, los niveles de amplificación no son demasiado grandes, porque el contraste de impedancia entre los sedimentos y el lecho de roca (igual a 2.45) es relativamente bajo.

La geometría de la cuenca contribuye de una manera determinística a la respuesta. Donde las profundidades del sótano son >3 Km., se ve una respuesta sísmica característica en el rango de frecuencias repartido en este estudio que no puede predecirse utilizando un modelo de 1D. Los modelos amplificados del desplazamiento observado para ondas incidentes P, SV, SH tienen los niveles máximos a esas situaciones donde las profundidades de la cuenca son más grandes. Es más, según este modelo, en la Cuenca de Granada es probable producir y emitir las ondas de la superficie con poca energía, cuando se observa en el sismograma sintético y su polarigramas correspondiente. Estas ondas en esta cuenca no causan los desplazamientos grandes. La generación de estas ondas de la superficie es un rasgo común que se ha observado en muchas cuencas alrededor del mundo usando los archivos reales y ese se ha reproducido utilizando los métodos numéricos para modelar la respuesta sísmica de estructuras geológicas similares.

Una de las conclusiones más importantes que se puede obtener de este trabajo es que un modelo 3D simple permite señalar efectos del sitio importantes que pueden influir en el movimiento sísmico a los terremotos futuros en la Cuenca de Granada entre 0 y 0.312 Hz. Con el IBEM y las funciones del Green apropiadas, es posible tener en cuenta una estructura de capa o un medio con una pendiente de velocidad. Todo esto podría proporcionar buenas estimaciones de movimiento sísmico para comparar los archivos de aceleración de tierra, con el objetivo de entender la respuesta sísmica compleja de la Cuenca de Granada.

Por otro lado, en esta tesis se puede observar claramente la interrelación que existe entre la carrera de Matemáticas Aplicadas y Computación y otras áreas de estudio, como en este caso, la Geología, Sismología y Física. Me di cuenta que la carrera nos da buenas bases para poder entender y resolver problemas que no son estrictamente de M.A.C., ya que apliqué conocimientos de varias materias como Cálculo, Álgebra lineal, Optimización, Métodos numéricos, Programación Avanzada y Estructuras de datos.

Bibliografía

- 1) Frankel, A. (1993). Simulaciones Tridimensionales de movimientos de Tierra en el Valle de San Bernardino, California, para terremotos hipotéticos en la falla de San Andrés, *Bol. Seism. Soc. Am.*, **83**, 1020-1041.
- 2) Frankel, A. y W. Stephenson (2000). Simulaciones Tridimensionales de movimientos de Tierra en la región de Seattle para terremotos en la zona de falla de Seattle, *Bol. Seism. Soc. Am.* **90**, 1251-1267.
- 3) Hatayama, K. y H. Fujiwara (1998). Excitación de ondas secundarias Love y Rayleigh en una cuenca sedimentaria tri-dimensional evaluada por el Método Directo de Elementos de Frontera con los modos normales, *Geophys. J. Int.* **133**,2, 260-278.
- 4) Luzón, F., F. J. Sánchez-Sesma, J. L. Rodríguez-Zúñiga, A. M. Posadas, J. M. García, J. Martín, M. D. Romacho, y M. Navarro, (1997). Difracción de ondas P, S y Rayleigh para una topografía tri-dimensional, *Geophys. J. Int.* **129**, 571-578.
- 5) Olsen, K. B., y R. J. Archuleta (1996). Simulación tri-dimensional de terremotos en sistema de falla de Los Ángeles. *Bol. Seism. Soc. Am.* **86**, 575-596.
- 6) Olsen, K. B. (2000). Amplificación del Sitio en la cuenca de los Ángeles del modelado tri-dimensional del movimiento del suelo, *Bol. Seism. Soc. Am.* **90**, 6B, S77-S94.
- 7) Pissanetzky, Sergio. *Sparse Matrix Technology*. Ed. Academic, Londres, 1984
- 8) R.E. Sheriff, L.P Geldart. *Exploración Sismológica Volumen 1*. Ed. Limusa, México, 1991.
- 9) Thorne Lay, Terry C. Wallace. *Modern Global Seismology*. Ed. Academic, San Diego, 1995.

- 10) Ortiz-Alemán, C., F. J. Sánchez-Sesma, J. L. Rodríguez-Zúñiga, y F. Luzón (1998). Computing topographical 3D site effects using a fast IBEM/Conjugate gradient approach, *Bull. Seism. Soc. Am.* **88**, 393-399.
- 11) Sánchez-Sesma, F. J. (1996). Strong ground motion and site effects, in *Computer Analysis of earthquake resistant structures*, D. E. Beskos and S.A. Anagnostopoulos (Editors), Comp. Mech. Publications, Southampton, 200-229.
- 12) Sánchez-Sesma, F. J. and F. Luzón (1995). Seismic response of three dimensional alluvial valleys for incident *P*, *S* and Rayleigh waves, *Bull. Seism. Soc. Am.* **85**, 269-284.
- 13) Mason J.C. *Métodos matriciales: Teoría, problemas y aplicaciones*. Ed. Anaya Multimedia, Madrid, 1986.
- 14) Scheid, Francis. *Teoría y problemas de análisis numéricos*. Ed. McGraw-Hill, 1972.
- 15) Boyce, DiPrima; *Ecuaciones Diferenciales y problemas con valores en la frontera*. Ed. Limusa, México, 1983.
- 16) Shoichiro Nakamura. *Métodos Numéricos Aplicados con Software*. Ed. Prentice-Hall Hispanoamericana, S.A., México, 1992.

Glosario

Amplitud

Es el valor máximo de la Función de Onda y corresponde al máximo valor que alcanza la perturbación en un punto.

Coefficiente de reflexión

Coefficiente complejo que caracteriza el fenómeno de reflexión en una frontera de separación. Su modulo determina el cociente de amplitudes y su argumento el desfase de la onda reflejada respecto a la incidente.

Dirección de propagación

Es aquella en la cual se transmite la onda. A veces se utiliza el término *rayo* como línea orientada en la dirección y sentido de la propagación.

Ecuación de Ondas

Establece una relación entre las variaciones temporales y espaciales de la función de onda a través de un coeficiente que depende de propiedades asociadas al medio donde se propaga. Proviene de las leyes físicas que rigen el fenómeno en particular. Para ondas de pequeña amplitud la forma matemática de la Ecuación de Ondas implica la igualdad entre las derivadas segundas de la función de ondas respecto al tiempo y respecto al espacio mediante un factor que es el cuadrado de la velocidad con que se propaga la onda.

Fase

Es el valor que en la expresión matemática de la onda toma el argumento de la función. Existe un desfase cuando se comparan dos ángulos de distinto valor. La perturbación en dos puntos diferentes puede estar en fase si el valor que toma la función de onda es el mismo en cualquier instante. En otro caso los puntos están desfasados.

Foco, Fuente u Origen de la Onda

Es el punto donde se produce inicialmente la perturbación. Si la perturbación varía con el tiempo en forma oscilatoria la onda va a consistir en la propagación de una vibración y si es armónica estará representada por una función de onda armónica. La perturbación en el foco oscila con una frecuencia que determina la frecuencia de la onda y su valor máximo coincide con la amplitud de la onda.

Frecuencia

Es el número de veces que en la unidad de tiempo se repite el mismo valor de la perturbación en un punto. Equivale a la inversa del periodo. La frecuencia de la onda esta determinada por el foco

Frente de Ondas

Es una superficie definida por los puntos en fase. La **Onda plana** cuyo frente de ondas es un plano, es la mas utilizada por ser una buena aproximación de cualquier frente en puntos alejados del origen. Además su función de onda es unidimensional.

Función de Onda

Es la expresión matemática de la onda, es decir la función que relaciona la magnitud perturbada con el tiempo y las coordenadas espaciales de cada punto del medio. Se obtiene como solución de la ecuación de ondas. Nos referimos a ella cuando utilizamos una expresión para denominar la Onda o Movimiento Ondulatorio Cuando la Función de Onda corresponde a un seno o coseno se le llama *Onda armónica* Si además depende de una sola variable espacial se tiene una *Onda armónica* monodimensional.

Interferencia

Superposición de ondas que puede dar lugar a una intensificación o debilitamiento de la onda resultante respecto de las ondas componentes. En el primer caso tenemos interferencia constructiva y en el segundo destructiva.

Longitud de Onda

Es la distancia mínima entre dos puntos con el mismo valor de la perturbación.

Medio material, vacío

Constituye el conjunto de puntos del espacio por donde se transmite la onda pudiendo ser material si hay masa puntual o vacío en caso contrario. Las únicas ondas que se propagan en vacío son las electromagnéticas. Las propiedades del medio determinan la velocidad de propagación.

Numero de Ondas

Es el número de longitudes de onda que se pueden contar en 2π unidades de longitud. Equivale a una frecuencia espacial.

Onda

Es una perturbación en movimiento, de forma que el valor en un punto (que inicialmente sería el foco) se transmite a los puntos contiguos sucesivamente. Puesto que la perturbación tiene diferente valor dependiendo del punto y del instante considerados se utiliza como equivalente el término *Movimiento Ondulatorio*.

Onda continua

La onda es continua cuando la oscilación en el origen se repite durante tiempo indefinido.

Onda estacionaria

Onda que no avanza porque alguno de los puntos está permanentemente en un estado de perturbación nula.

Onda longitudinal

Es aquella en que la dirección de propagación coincide con la dirección en que vibra la magnitud física que constituye la onda.

Onda transversal

Es toda aquella en la que la magnitud física cambia o vibra, en dirección perpendicular a la dirección de avance de la onda.

Ondas electromagnéticas

Están constituidas por campos eléctricos y magnéticos pudiéndose propagar en el vacío.

Ondas mecánicas o elásticas

Son las que viajan en medios materiales debido a sus propiedades elásticas.

Periodo

Es el tiempo mínimo transcurrido para que en un punto se repita un mismo valor de la perturbación.

Perturbación

Es la variación de una magnitud física respecto a un determinado valor que se considera estacionario o de equilibrio. Cuando la magnitud física es un vector su dirección define una dirección de la perturbación y en cualquier otro caso consideramos que ésta sería la dirección en la cual cambia (o vibra) la magnitud física.

Pulso

La onda es un pulso cuando la perturbación en el origen se produce sólo durante un corto intervalo de tiempo.

Resonancia

Fenómeno por el que se conoce a las ondas estacionarias confinadas en una región. . Lleva implícito un almacenamiento de energía en ese espacio confinado y sólo es posible para determinadas frecuencias relacionadas con sus dimensiones.

Superposición de ondas

Suma algebraica de las ondas coincidentes en un punto como consecuencia de aplicar el principio de superposición determinado por la linealidad de la ecuación de ondas.

Velocidad de propagación de la Onda

Es aquella a la cual se transmite la perturbación. Coincide con la velocidad a la que se propaga la fase de la función de onda. Numéricamente coincide con el cociente entre la longitud de onda y el periodo.