

## **TRABAJO ESPECIAL DE GRADO**

### **IMPLEMENTACIÓN DE UN ALGORITMO PARA RESOLVER SISTEMAS GRANDES DE ECUACIONES LINEALES MEDIANTE EL MÉTODO DE ELEMENTO A ELEMENTO**

Presentado ante la ilustre  
Universidad Central de  
Venezuela

Por los bachilleres:  
Bouzas Rodríguez, José Manuel  
Wallis Iglesias, Roberto  
para optar al Título  
De Ingeniero Mecánico.

Caracas, 2004

## **TRABAJO ESPECIAL DE GRADO**

### **IMPLEMENTACIÓN DE UN ALGORITMO PARA RESOLVER SISTEMAS GRANDES DE ECUACIONES LINEALES MEDIANTE EL MÉTODO DE ELEMENTO A ELEMENTO**

TUTOR ACADÉMICO: Prof. Antonio Barragán.

Presentado ante la ilustre  
Universidad Central de  
Venezuela  
Por los bachilleres:  
Bouzas Rodríguez, José Manuel  
Wallis Iglesias, Roberto  
para optar al Título  
De Ingeniero Mecánico.

Caracas, 2004

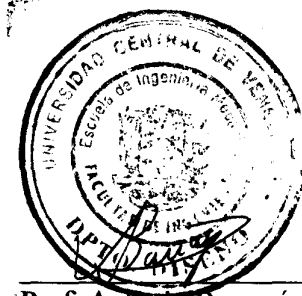
Caracas, 01/07/2004

Los abajo firmantes, miembros del jurado designado por el consejo de la Escuela de Ingeniería Mecánica, para evaluar el Trabajo Especial de Grado presentado por los bachilleres, titulado:

**“IMPLEMENTACIÓN DE UN ALGORITMO PARA RESOLVER SISTEMAS GRANDES DE ECUACIONES LINEALES MEDIANTE EL MÉTODO DE ELEMENTO A ELEMENTO”**

Consideran que el mismo cumple con los requisitos exigidos por el plan de estudios conducente al Título de Ingeniero Mecánico.

Prof. Manuel Martínez  
Jurado



Prof. Antonio Barragán  
Tutor Académico

Prof. Jorge Cruz  
Jurado

*A nuestros familiares y amigos.*

## **Agradecimientos.**

A Dios por guiar mis pasos.

A mis padres por haberme dado el ser, estar a mi lado en todo momento, confiar en mí y apoyarme en mis decisiones.

A mi hermana Carolina por ser tan especial.

A la ilustre Universidad Central de Venezuela y su personal docente por haberme formado como profesional.

A nuestro tutor Ing. Antonio Barragán, por su dedicación, orientación y amistad en todo momento de mi carrera, no tengo palabras para agradecerle todo lo que su ayuda significativo para mí.

A mi compañero de tesis Roberto Wallis (1<sup>st</sup> Horseman) por ser amigo incondicional.

Al 2<sup>nd</sup> Horseman por ayudarme en todo momento, al 3<sup>th</sup> Horseman por hacernos reír en los momentos más difíciles y ser mis amigos incondicionales.

A mi amigo Jorge Bermúdez (Honorific Horseman), por todos sus años de amistad.

A mis amigos de la facultad los cuales los llevo en mi corazón y por hacer de la escuela de ingeniería mecánica mi segundo hogar.

A la gente que no creía en mí, por darme la fortaleza de lograr mis metas.

*José Manuel Bouzas Rodríguez*

## **Agradecimientos.**

A Dios y a mis padres, por todo lo que me han dado.

A todos mis amigos, por apoyarme, escucharme y por las interminables horas de parranda que hemos compartido. ¡Ellos saben quienes son!

A las que me han amado, por ser una luz en mi vida.

Al Ing. Antonio Barragán, nuestro tutor, por su amistad, apoyo y comprensión, no sólo durante la realización de este trabajo, sino durante la mayoría del tiempo que estuve en la Escuela de Ingeniería Mecánica.

A mi compañero y co-autor de este trabajo José Manuel Bouzas, por tantos años de amistad.

A mis compañeros, Gustavo González, Bernardo Porte, Jorge Bermúdez y Estefanía López, porque sin ellos no hubiese llegado hasta aquí.

A los que me apoyaron, por creer siempre en mí.

A los que se opusieron, porque sin ellos no hubiese tenido la fuerza para demostrarles que estaban equivocados.

*Roberto Wallis Iglesias.*

**Bouzas R., José M. /y/ Wallis I., Roberto**

**IMPLEMENTACIÓN DE UN ALGORITMO PARA RESOLVER  
SISTEMAS GRANDES DE ECUACIONES LINEALES  
MEDIANTE EL MÉTODO DE ELEMENTO A ELEMENTO.**

**Tutor académico: Prof. Antonio Barragán.  
Tesis. Caracas. U.C.V. Facultad de Ingeniería.  
Escuela de Ingeniería Mecánica. 2004. 171 pág.**

Palabras Claves: Elementos Finitos, Gradiente Conjugado, Elemento a Elemento.

En el presente trabajo especial de grado, se hizo una recopilación de información sobre los algoritmos de solución del sistema de ecuaciones generado en la formulación del Método de los Elementos Finitos en problemas de elasticidad lineal. En la resolución de los sistemas dispersos de ecuaciones lineales, se analizaron los métodos directos de solución y los métodos iterativos de solución. Se investigaron distintos métodos de solución y se implementó el más eficiente, basándose en las referencias. El método iterativo de solución que se utilizó fue el Método del Gradiente Conjugado. Para optimizar el sistema de resolución se utilizó el preconditionador de Elemento a Elemento. Se recopiló información acerca de los pseudocódigos de estos dos métodos. Este método de resolución optimizado en pseudocódigo, se tradujo a lenguaje de programación C y se implementó en un programa existente de elementos finitos. Luego se procedió a hacer una comparación entre el programa resultante (*MefetGC*) y programas comerciales de elementos finitos. Resultando en la validación del programa *MefetGC*. Se observó que el programa *MefetGC* obtuvo un menor margen de error que los programas comerciales. Se concluyó que el método iterativo optimizado utilizado tiene un mejor rendimiento que los sistemas de solución antiguos.

**Índice.**

Introducción .....	1
Antecedentes .....	3
1. El método de los Elementos Finitos.....	11
1.1 Qué es el método de los Elementos Finitos. ....	11
1.2 Principios de Cálculo Variacional.....	13
1.2.1 Concepto de Funcional.....	13
1.2.2 Punto Estacionario de un Funcional.....	14
1.2.3 Métodos Directos del Cálculo Variacional. ....	15
1.2.4 El método de Rayleigh-Ritz.....	16
1.3 Principio del Trabajo Virtual. ....	18
1.3.1 Trabajo y Energía.....	19
1.3.2 El Principio del trabajo virtual para un sólido continuo. ....	21
1.4 Energía potencial Total. ....	23
1.5 Principio de la Energía Potencial Mínima. ....	25
1.6 El Método de los Elementos Finitos. ....	27
1.7 Formulación del método de elementos finitos. ....	28
1.7.1 Expresión matricial de la energía potencial total. ....	29
1.7.2 Funciones de interpolación. ....	31
1.7.3 Aproximación por elementos finitos.....	35
1.7.4 Matriz gradiente. ....	38
1.7.5 Matriz de rigidez y vector de cargas nodales equivalentes. ....	40
1.7.6 Características de la matriz de rigidez global. ....	43
1.8 Convergencia de los Resultados. ....	44
1.9 Exactitud de la solución. ....	46
1.10 Aplicación del método de los elementos finitos al análisis de problemas de elasticidad bidimensional. ....	47
1.10.1 Coeficientes de rigidez para un elemento de esfuerzo-deformación para un elemento triangular general.....	47
1.10.2 Equilibrio en los nodos.....	66
2. Métodos directos de solución de sistemas de ecuaciones. ....	71
2.1 Descomposición LU.....	71
2.2 Eliminación de Gauss.....	73
2.3 Eliminación de Gauss-Jordan.....	76
3. Métodos iterativos de solución de sistemas de ecuaciones.....	79
3.1 Formas cuadráticas.....	80



---

3.2 El método del descenso más rápido.....	84
3.3 El método de las direcciones conjugadas.....	89
3.4 El Método del Gradiente Conjugado.....	97
3.4.1 Método del Gradiente Conjugado.....	98
3.5 Precondicionamiento.....	101
3.6 Algoritmos en pseudo-código.....	103
4. El Precondicionador de Elemento a Elemento.....	106
4.1 Precondicionador de Elemento a Elemento.....	108
5. Implementación.....	111
5.1 Pseudo-código.....	111
5.2 Implementación en lenguaje C.....	112
6. Parte Experimental.....	116
6.1 Factor de concentración de esfuerzos.....	116
6.2 Procedimiento de validación (Metodología).....	119
6.3 Comparación de los programas.....	123
6.4 Resultados.....	125
7. Análisis de resultados.....	134
8. Conclusiones.....	137
9. Recomendaciones.....	138
A. Matrices Dispersas (Tipo Sparse).....	140
A.1 Matrices Dispersas.....	140
B. Factorización de Cholesky.....	142
Teorema.....	144
C. Número de Condición.....	146
C.1 Condicionamiento de un sistema.....	146
Ejemplo.....	147
C.2 Número de condición de una matriz.....	149
Ejemplo.....	150
D. Matriz Hermítica.....	152
Referencias/Bibliografías.....	154

## Índice de Figuras.

Figura 1-1 a) Circunferencia de radio $r$ ; b) Discretización de la circunferencia en ocho elementos finitos; c) Un elemento finito desconectado del dominio con los nodos 4 y 5; d) Un elemento genérico. ....	11
Figura 1-2 Aproximación de $u(x)$ con elementos lineales. ....	32
Figura 1-3. Funciones de forma locales $N_{e_i}$ lineales en cada elemento. ....	33
Figura 1-4. Aproximación en dos dimensiones. La parte de abajo es el dominio bidimensional y la parte de arriba es la función de forma. ....	34
Figura 1-5. Discretización de un dominio $\Omega$ . ....	35
Figura 1-6. Desplazamientos nodales para el triángulo de 3 nodos. ....	37
Figura 1-7 Área del elemento calculada en términos de las coordenadas nodales. ....	49
Figura 1-8 Condiciones de esfuerzo plano y fuerzas puntuales únicamente en los límites. ....	54
Figura 1-9 Estructura del ejemplo ilustrativo. ....	57
Figura 1-10 Elemento 1. ....	58
Figura 1-11 Elemento 2. ....	63
Figura 1-12 Elemento 3. ....	64
Figura 1-13 Equilibrio en el nodo 1. ....	66
Figura 1-14 Equilibrio en el nodo 2. ....	67
Figura 1-15 Equilibrio en el nodo 3. ....	68
Figura 2-1 Diagrama de descomposición LU. ....	72
Figura 2-2 Diagrama de método de Gauss. ....	75
Figura 2-3 Diagrama del método de Gauss-Jordan. ....	78

Figura 3-1 Sistema $Ax=b$ .....	81
Figura 3-2 Formas cuadráticas del sistema $Ax=b$ .....	82
Figura 3-3 Formas cuadráticas en forma de curvas de nivel. ....	83
Figura 3-4 Búsqueda de línea. ....	85
Figura 3-5 El gradiente en el punto más bajo es ortogonal al gradiente del paso anterior.....	86
Figura 3-6 a) Esta parábola es la intersección de las superficies. b) Pendiente de la parábola y magnitud de la proyección del gradiente en la línea. ....	86
Figura 3-7 Solución del ejemplo. Nótese el patrón en forma de Zig-Zag. ....	88
Figura 3-8 Direcciones de búsqueda ortogonales. ....	89
Figura 3-9 Minimización de $\ e\ _A$ . ....	94
Figura 3-10 Debido a que las direcciones de búsqueda $d_{(0)}, d_{(1)}$ se construyen a partir de los vectores $u_0, u_1$ , éstos crean el subespacio $D_2$ (el plano coloreado en gris). El término del error $e_{(2)}$ es conjugado con $D_2$ , el residuo $r_{(2)}$ es ortogonal a $D_2$ , y una nueva dirección de búsqueda $d_{(2)}$ es construida de $u_2$ para ser conjugado con $D_2$ . Los puntos finales de $u_2$ y $d_{(2)}$ se encuentran en un plano paralelo a $D_2$ , debido a que $d_{(2)}$ se construye a partir de $u_2$ por la conjugación de Gram-Schmidt. ....	95
Figura 3-11 Conjugación de Gram-Schmidt de dos vectores. Se empieza con dos vectores linealmente independientes $u_0$ y $u_1$ . Se hace $d_{(0)} = u_0$ . El vector $u_1$ posee dos componentes, que es conjugado con $d_{(0)}$ , y $u^+$ , que es paralelo a $d_{(0)}$ .....	96
Figura 3-12 En el método del Gradiente Conjugado, cada nuevo residuo es ortogonal a todos los residuos anteriores y direcciones de búsqueda; y cada nueva dirección de búsqueda es construida para ser conjugado a todos los residuos y direcciones de búsqueda previos. Los puntos finales de $r_{(2)}$ y $d_{(2)}$ se encuentran en un plano paralelo a $D_2$ (el espacio gris). En este método, $d_{(2)}$ es una combinación lineal de $r_{(2)}$ y $d_{(1)}$ .....	99

---

Figura 6-1 <i>Disposición de la probeta.</i> .....	116
Figura 6-2 <i>Probeta idealizada.</i> .....	117
Figura 6-3 <i>Mallado realizado en forma manual de 144 nodos y 55 elementos.</i> .....	118
Figura 6-4 <i>Resultados del programa Mechanical Desktop</i> .....	127
Figura 6-5 <i>Resultados del programa Visual Nastran 4D.</i> .....	128
Figura 6-6 <i>Programa Engineers Toolbox v2.0.</i> .....	130
Figura 6-7 <i>Resultado del <math>K_t</math> teórico.</i> .....	131
Figura 6-8 <i>Curva del factor de concentración de esfuerzos para placa finita con agujero circular sometida a tensión.</i> .....	133
Figura A-0-1 <i>Patrón general. Cada cuadro oscurecido representa un coeficiente o submatriz si más de una cantidad está siendo considerada en los nodos.</i> .....	141
Figura C-0-1 <i>Sistema bien condicionado y sistema mal condicionado.</i> .....	147

***Índice de Tablas.***

*Tabla 1 Resultados de los esfuerzos en el elemento 1 para el programa MefetGC. Se presenta el esfuerzo máximo de Von Misses. .... 126*

*Tabla 2 Resultados de los esfuerzos en el elemento 1 para el programa Hexafron [5]. Se presenta el esfuerzo máximo de Von Misses. .... 129*

## Introducción

Dentro del campo de la Ingeniería Mecánica, un área de gran importancia es el diseño de piezas o componentes mecánicos, elementos capaces de soportar de manera confiable los esfuerzos que se generan en ellos al ser sometidos a condiciones de funcionamiento. Durante mucho tiempo, la evaluación de esos diseños se ha realizado mediante técnicas experimentales, lo cual resulta ser de un costo sumamente elevado, además de una gran inversión en infraestructura (laboratorios, equipo de ensayo).

En el diseño de componentes mecánicos intervienen diferentes factores que afectan la distribución de esfuerzos en los mismos, como son: la forma, condiciones de apoyo, características de los materiales involucrados, cargas aplicadas. La conjunción de estos factores hace que el proceso de análisis sea complejo, exigiendo métodos sofisticados, capaces de simular diferentes condiciones de trabajo.

En la actualidad existen métodos computacionales de análisis que cumplen estos requisitos. Basados en técnicas matemáticas aproximadas, estos métodos permiten simular con grado variable de aproximación el comportamiento de piezas mecánicas bajo condiciones de trabajo. El uso de estos métodos se basa en el computador digital, el cual permite lograr resultados útiles en poco tiempo a bajo costo, siendo posible determinar los efectos de diferentes condiciones de cargas, geometrías y materiales. Fundamentalmente, los métodos utilizados son: El Método de los Elementos Finitos y El Método de los Elementos de Contorno.

Recientemente, la evolución de los computadores personales pone al alcance del usuario, equipos de altas prestaciones a un costo relativamente bajo, lo que nos permite la simulación y análisis de problemas grandes.

En la Escuela de Ingeniería Mecánica (EIM) se han venido desarrollando un grupo de programas basados en el método de los elementos finitos con la finalidad de tener una plataforma computacional para la solución de algunos tipos de problemas de la física matemática (análisis de esfuerzos, transferencia de calor). Estos problemas se presentan al simular el comportamiento de piezas o componentes como parte del proceso de diseño de los mismos.

El objetivo fundamental del presente trabajo consiste en incorporar a dicho programa, un sistema de solución de ecuaciones de alta capacidad, con lo cual se espera tener la capacidad de resolver problemas en los cuales el tamaño de la malla es grande.

El tipo de algoritmo que se ha considerado utilizar se basa en una combinación de métodos basados en el gradiente conjugado, conjuntamente con algoritmos de solución llamados Elemento a Elemento (EBE, en inglés), siendo su atractivo el que no requieren el ensamblaje de la matriz de rigidez completa, sino que más bien ensamblan y calculan la solución en base a las contribuciones individuales de los nodos de cada elemento. De allí su nombre “Elemento a Elemento”.

## Antecedentes

El método de elementos finitos hoy en día es una herramienta sumamente extendida y empleada tanto en investigación y desarrollo en la mayor parte de los ámbitos científicos y tecnológicos, como en numerosos sectores productivos de la sociedad actual, preocupados por la mejora de la calidad de sus productos y procesos.

No es extraño encontrar aplicaciones del método de los elementos finitos en áreas tan distantes entre si como el diseño estructural, campo en el que el método se desarrolló originalmente y del que nutrió de múltiples conceptos físicos, y la meteorología en donde se resuelven actualmente los problemas de simulación numérica, quizás de mayor tamaño.

El uso del computador permite resolver de manera aproximada, problemas de ingeniería tales como los de elasticidad, transferencia de calor y mecánica de los fluidos. La adaptabilidad a regiones de geometría compleja, ha hecho que su uso sea mayor en estas últimas décadas.

Debemos citar el esfuerzo de las universidades y centros de investigación de USA, los países Europeos y latinoamericanos, tanto en la formación de investigadores en esta área de simulación como en la difusión de las ventajas de los métodos numéricos.

A continuación se resume una serie de investigaciones, las cuales han sido referencia para el presente trabajo.



**Ordaz y Pulido [28]**, en este trabajo se muestran soluciones a ciertos problemas geométricos asociados a la generación automática de mallas tridimensionales de elementos finitos, como lo son:

- Determinación de la posición de un punto respecto a un sólido.
- Determinación de la posición de un sólido respecto a un plano orientado.
- Intersecciones entre plano y plano, plano y sólido, recta y sólido, recta y plano.

Tales soluciones fueron logradas mediante la implantación de un programa en lenguaje C, el cual trabaja con estructuras dinámicas de datos.

**Alonzo y Oramas [1]**, elaboraron un programa para representación gráfica tridimensional, aplicable a componentes mecánicos y de bioingeniería. Este programa es capaz de realizar transformaciones de los objetos ya creados, tales como: rotaciones, traslaciones, reflexiones, proyecciones, cambios de escalas, visibilidad y aproximaciones a curvas, de todo el objeto o parte de él. Los objetos realizados en este programa no solo pueden de formas geométricas preestablecidas, sino también pueden ser derivados de una combinación o de forma libre.

**Rosales [33]**, consistió en la elaboración e implementación de un programa de computación que resuelve estructuras planas tomando en cuenta la posibilidad que las relaciones esfuerzo-deformación de los materiales constituyentes no sean lineales. El programa fue basado en el método de los elementos finitos. Los elementos que se consideraron para el modelaje fueron elementos planos de cuatro lados que poseen de cuatro a ocho nodos. El método de análisis no lineal implementado fue el método incremental.

**Fernández y Gómez [13]**, utilizaron el método de los elementos finitos para el desarrollo e implementación de un sistema o ‘paquete’ que sea una herramienta única, capaz de resolver diferentes tipos de problemas que se pueden presentar a un diseñador, como lo son el estudio del fenómeno de elasticidad lineal y no lineal, y el estudio del fenómeno de transferencia de calor tanto en régimen permanente como en régimen transitorio. Para ello se implementó un programa computacional que permite calcular los desplazamientos y esfuerzos tanto en modelos de comportamiento lineal como no lineal en el caso de fenómenos de elasticidad, y además calcular la distribución de temperaturas y esfuerzos térmicos para el estudio del fenómeno de transferencia de calor.

**Morales y Sánchez [26]**, muestra una aplicación del método de los elementos finitos a la solución del problema de flujo de calor en régimen transitorio. Para lo cual se implementa un programa de computación que calcula la distribución de temperaturas y esfuerzos térmicos, en donde se manejan condiciones de borde de primera y segunda clase.

**Montero, Montenegro y Rodríguez [25]**, presentan una visión general de técnicas avanzadas para la resolución de grandes sistemas de ecuaciones lineales con matriz dispersa (sparse). En primer lugar se introducen diferentes algoritmos de reordenación orientados a mejorar el efecto del preconditionamiento de un sistema. Seguidamente, se define el concepto de preconditionamiento y se formula algunos de los preconditionadores más usados en la actualidad, destacando los desarrollos recientemente basándose en la inversa de una matriz. Por otro lado, se consideran algunos métodos iterativos basados en los subespacios de Krylov para la resolución de sistemas de ecuaciones lineales. Para el caso simétrico se propone el Gradiente Conjugado, mientras que para el no simétrico, existen varias alternativas que se pueden clasificar en tres grandes familias de métodos, los de ortogonalización, los de

bi-ortogonalización, y los basados en la Ecuación Normal. Esta estrategia de resolución, que combina las tres técnicas anteriores, parece la más eficiente desde el punto de vista computacional como lo muestran los experimentos que se encuentran en el trabajo.

**Sánchez y Moraga [35]**, realiza un estudio numérico de mecánica de fluidos 2D en una contracción brusca para fluidos no newtonianos tipo ley de potencia de Oswald-de-Waele. Se emplean dos métodos de resolución, el método de los volúmenes finitos y el método de los elementos finitos. En el primer método se utiliza el algoritmo SIMPLE, para la resolución de las ecuaciones de momentum lineal y de continuidad. Para el método de elementos finitos se utiliza un código propio en lenguaje Fortran 90. En ambos métodos se estudia el comportamiento del fluido para diferentes índices de potencia. Los cálculos se realizan con malla variable rectangular (método de volúmenes finitos) y con una malla apegada al contorno (método elementos finitos)

**Daydé, L'Excellent y Gould [9]**, consideran la solución de un sistema  $n$  por  $n$  de ecuaciones lineales.  $Ax = b$ , donde  $A$  es estructurado y disperso así que se puede expresar como:

$$A = \sum_{i=1}^p A_i$$

Los sistemas lineales estructurados dispersos se utilizan en muchas aplicaciones. Las matrices elementales  $A_i$  son de bajo rango, y son usualmente dispersos, de modo que sus entradas no nulas se representan como un pequeño bloque denso.

Asumieron que  $A$  es una matriz simétrica grande y normalmente positiva. La técnica de solución considerada es el método del gradiente conjugado utilizando un rango de

precondicionadores de Elemento a Elemento (EBE) que fueron introducidos por [16] y [29], que han sido utilizados con éxito en muchas aplicaciones de ingeniería y física.

**Daydé, Decamps y L'Excellent [10]**, estudiaron la solución de problemas de gran escala, no lineales y sin restricciones, utilizando técnicas que explotan la estructura, común en estos problemas, de separabilidad parcial. Muestra cómo los preconditionadores pueden ser calculados para diseñar métodos iterativos utilizando la separabilidad parcial. La técnica de optimización se basa en el algoritmo de Newton truncado, el cual se presta para la solución de problemas de gran escala. El sistema lineal que provee las direcciones de búsqueda es resuelto utilizando el gradiente conjugado combinado con un rango de preconditionadores de Elemento a Elemento.

**Hestenes y Stiefel [15]**, introdujeron un método iterativo para resolver sistemas de ecuaciones lineales con  $n$  incógnitas llamado Método del Gradiente Conjugado. Se muestra que este método es un caso especial de un método muy general que incluye la eliminación Gaussiana. Estos algoritmos generales son esencialmente algoritmos para hallar un elipsoide  $n$  dimensional.

**Carey y Jiang [4]**, experimentó implementando el Método del Gradiente Conjugado con una modificación, se remodeló el método en una forma que permita calcular elemento a elemento secuencialmente con métodos de elementos finitos. Esta estrategia se implementó para resolver los sistemas lineales provenientes de la aproximación de elementos finitos para la ecuación de Laplace. También se aplicó para aproximar una clase representativa de problemas no lineales.

**Hughes, Levit y Windget [16]**, propusieron un algoritmo para resolver sistemas de gran escala de ecuaciones provenientes de elemento finitos provenientes de mecánica de sólidos y mecánica estructural mediante el uso de una técnica de factorización aproximada de elemento a elemento, el cual no necesita de una matriz global de coeficientes.

**Jennings y Malik [18]**, realizaron una comparación entre el Método del Gradiente Conjugado ,los métodos de relajación, y la iteración acelerada de Jacobi de Chebychev, cuando se aplica a la solución de grandes grupos de ecuaciones lineales que tengan una matriz de coeficientes dispersa, definida positiva y simétrica. Llegó a le en el peor de los casos la rata de convergencia de GC no va a ser peor que la de los otros métodos, de hecho va a ser considerablemente mejor.

**Dickinson y Forsyth [11]**, Realizaron una comparación entre el rendimiento del Método del Gradiente Conjugado con mallas tetraédricas y formas cuadráticas, y el rendimiento de un método directo de solución. Se consideraron problemas hasta de 70000 grados de libertad y tamaño de elemento considerablemente pequeño.

**Chapman y Cox [7]**, desarrollaron una metodología para identificar la condición única del preconditionador EBE mediante un algoritmo. Además introdujeron una modificación a la estructura vectorial iterativa del preconditionador EBE; la implementaron y discutieron.

**Klisinski y Runesson [20]**, presentaron una valoración de la eficiencia de un nuevo grupo de rutinas para la solución de sistemas de ecuaciones lineales, comparándolas con una selección de paquetes de solucionadores comerciales. Consideraron problemas simétricos y no simétricos. Llegaron a la conclusión de que la efectividad del compilador es igual de importante que el código en sí. En la

comparación determinaron que los solucionadores recientemente desarrollados mostraron un rendimiento mayor.

**Vignjevic, Morris y Belagundu [41]**, presentaron un procedimiento sistemático para el uso confiable y preciso del Método de los Elementos Finitos. Hicieron énfasis en valorar los efectos de las suposiciones del modelado en las respuestas estructurales. El procedimiento es basado en una descomposición del proceso de idealización.

**Zörner y Plasmeijer [44]**, investigaron la aplicabilidad de lenguajes funcionales para la descripción e implementación eficiente de algoritmos en álgebra lineal numérica. Implementaron dos algoritmos para resolver sistemas lineales en Clean y los compararon con los mismos algoritmos en C y Matlab. Hicieron una comparación de los tiempos de corrida y demostraron el conflicto entre códigos elegantes y códigos eficientes.

**Lu, Connell y Tullberg [23]**, utilizaron la técnica de la Realidad Virtual para hacer un análisis de elementos finitos interactivo. Hicieron una comparación entre el análisis de elementos finitos por realidad virtual y los pre- y post- procesadores de los paquetes tradicionales de elementos finitos.

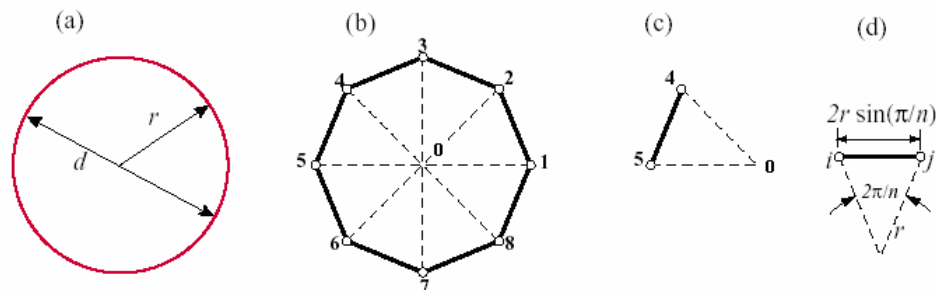
**Taghavi [40]**, desarrolló un programa de generación automática de mallas llamado HEXAR. Genera las mallas a partir de los datos producidos por paquetes CAD.



## 1. El método de los Elementos Finitos.

### 1.1 Qué es el método de los Elementos Finitos.

El método de los Elementos Finitos es una técnica numérica para resolver ecuaciones diferenciales en derivadas parciales discretizando el dominio en sus dimensiones espaciales. La discretización se lleva a cabo en pequeñas regiones de geometría simple y con forma arbitraria (el elemento finito). En la Figura 1-1 se muestra la discretización de una circunferencia en ocho elementos finitos de forma triangular. El resultado es un conjunto de ecuaciones que se expresan en forma de matrices, que relacionan la entrada en puntos específicos del elemento (los nodos) con la salida en estos mismos puntos. Un dominio que sea de nuestro interés se representará como un ensamblaje de elementos finitos. Las funciones de aproximación en elementos finitos son definidas en términos de valores nodales del campo físico que es buscado. Un problema físico continuo es transformado en un problema discretizado de elementos finitos con valores nodales desconocidos. Para un problema lineal, se debe resolver con un sistema lineal de ecuaciones.



**Figura 1-1**<sup>1</sup> a) Circunferencia de radio  $r$ ; b) Discretización de la circunferencia en ocho elementos finitos; c) Un elemento finito desconectado del dominio con los nodos 4 y 5; d) Un elemento genérico.

<sup>1</sup> Tomado de [17]



Dos características del método de elementos finitos que valen la pena mencionar se detallan a continuación:

La aproximación por elementos finitos de problemas de elasticidad lineal en casos dos y tres dimensiones produce grandes matrices dispersas<sup>2</sup>, para un problema discretizado. Esto ayuda a solucionar problemas con un número muy grande de incógnitas nodales. En la mayoría de los casos el costo del análisis de elementos finitos está dominado por el costo de la solución del gran sistema resultante disperso.

Una aproximación por piezas de los fenómenos físicos, mediante elementos finitos, provee una buena precisión aún con funciones simples de aproximación (con el aumento del número de elementos se puede mejorar la precisión).

En la interpretación física, el concepto básico es la partición o descomposición de un sistema mecánico complejo en un sistema simplificado de componentes disjuntos llamados elementos finitos o simplemente elementos. La respuesta mecánica de un elemento está caracterizada por un número finito de grados de libertad. Éstos se representan como los valores de las funciones desconocidas de un conjunto de nodos. La respuesta del elemento se define por ecuaciones algebraicas que se construyen a partir de los argumentos matemáticos o experimentales. La respuesta del sistema original se considera aproximada al del modelo discreto conectando o ensamblando todos los elementos.

Los problemas de equilibrio estático de sistemas discretos pueden ser modelados por los valores estacionarios de una función de una o varias variables; esto es dominio del cálculo ordinario. Cuando estos principios estacionarios son extendidos a una

---

<sup>2</sup> Refiérase al anexo A.

situación más general de análisis de tensiones y otro tipo de problemas en mecánica continua, entonces el cálculo variacional es la herramienta más adecuada.

## ***1.2 Principios de Cálculo Variacional.***

El cálculo variacional se concentra principalmente en los valores *extremos* (máximos y mínimos) o *estacionarios* de unas ciertas integrales definidas, las cuales se suelen llamar funcionales. Estas integrales involucran funciones desconocidas al principio del análisis del problema físico, y el objetivo es el de determinar cuales condiciones se generan cuando los funcionales toman valores estacionarios. Al satisfacer estas condiciones logramos producir las funciones en sí. La motivación que nos conduce a la utilización del cálculo variacional, es que los llamados “métodos directos” nos proveen de muchas técnicas poderosas para obtener resultados numéricos para problemas complejos, lo cual tiene una importancia significativa en el área de ingeniería.

### **1.2.1 Concepto de Funcional.**

En el cálculo de variaciones, se buscan los valores *extremos* (máximos y mínimos) o *estacionarios* de los **funcionales**. Un funcional se define como una cantidad cuyo valor depende de la forma completa de algunas funciones y no del número de variables discretas.

Tomemos por ejemplo la longitud de una curva  $y = f_{(x)}$ , en el intervalo de  $x = 0$  a  $x = l$ :

$$L_{(y)} = \int_0^l \sqrt{1 + (y')^2} dx$$

Esta expresión de  $L_{(y)}$  acepta cualquier  $f_{(x)}$  arbitrario, siempre y cuando sea una función continua y que acepte primera derivada. Sin embargo, el valor de  $L_{(y)}$  depende de la forma de  $f_{(x)}$ .

### 1.2.2 Punto Estacionario de un Funcional.

Sea un funcional del tipo:

$$I_{(y)} = \int_a^b F(x, y, y') dx \quad (1.1)$$

Buscamos un  $y = f_{(x)}$  que haga  $I_{(y)}$  estacionario. Necesitamos que  $f_{(x)}$  satisfaga todas las condiciones de borde y de continuidad, además necesitamos que  $f'_{(x)}$  y  $f''_{(x)}$  sean continuas y diferenciales con respecto a  $x$ ,  $y$  y  $y'$ .

El procedimiento a seguir es introducir una función:

$$u_{(x)} = f_{(x)} + \varepsilon \eta_{(x)} \quad (1.2)$$

La cual se aproxima a  $f_{(x)}$ ;  $\eta'_{(x)}$  y  $\eta''_{(x)}$  son continuos en  $\eta_{(a)} = \eta_{(b)} = 0$  de tal forma que  $\varepsilon \eta_{(x)}$  sea una variación admisible.

Entonces:

$$I_{(u)} = \int_a^b F(x, y + \varepsilon\eta, y' + \varepsilon\eta') dx$$

Se hace estacionario en  $\varepsilon = 0$ , por lo tanto:

$$\left. \frac{d}{d\varepsilon} \int_a^b F(x, y + \varepsilon\eta, y' + \varepsilon\eta') dx \right|_{\varepsilon=0} = 0 \quad (1.3)$$

Para un  $I_{(u)}$  estacionario,  $\frac{d}{d\varepsilon} I_{(u)} = \delta I = 0$ . Donde  $\delta I$  es llamado la primera variación de  $I_{(u)}$ .

Si una función  $f_{(x)}$  es continua en un intervalo  $a \leq x \leq b$  y:

$$\int_a^b f_{(x)} \eta_{(x)} dx = 0 \quad (1.4)$$

es cierto para todas las funciones  $\eta_{(x)}$ , las cuales son continuas y derivables, y  $\eta_{(a)} = \eta_{(b)} = 0$ , entonces  $f_{(x)} = 0$ . Éste es el teorema fundamental del cálculo variacional.

### 1.2.3 Métodos Directos del Cálculo Variacional.

Se ha visto que, aplicando el proceso formal del cálculo de variaciones a la determinación de las condiciones estacionarias de un funcional, se obtienen como

resultado ecuaciones diferenciales con sus correspondientes condiciones de borde, cuya solución permite determinar las funciones que hacen estacionario el funcional.

Muchas aproximaciones pueden ser usadas para transformar la formulación física del problema a su análogo discreto de Elementos Finitos. Si el problema puede ser formulado como una minimización de la energía potencial, entonces una *formulación variacional* de las ecuaciones de Elementos Finitos es usada generalmente. En el caso de mecánica de sólidos la *formulación variacional* es lo más usado.

Una solución alternativa, que permite obtener soluciones aproximadas, consiste en sustituir en el funcional, soluciones asumidas que involucren parámetros ajustables, para luego determinar las condiciones de estacionaridad del funcional con respecto a los parámetros ajustables. A este respecto, existen varios métodos de los cuales el método conocido como Rayleigh-Ritz es el más utilizado en la mecánica de sólidos.

De hecho el Método de los Elementos Finitos en su formulación usual para medios continuos, es simplemente la técnica de Rayleigh-Ritz discretizado por partes.

#### 1.2.4 El método de Rayleigh-Ritz.

La idea general es sustituir el medio continuo por un sistema discreto, caracterizado por tener un número finito de grados de libertad. En relación al Cálculo Variacional, supóngase que se desea encontrar la función  $y = f(x)$  que hace estacionario al funcional:

$$I_{(y)} = \int_a^b F(x, y, y') dx$$

con las condiciones de borde:  $f_{(a)} = f_{(b)} = 0$  .

El procedimiento de Rayleigh-Ritz consiste en asumir que  $y = f_{(x)}$  puede ser aproximada por una combinación lineal de funciones de prueba arbitrarias, de forma de:

$$y_n(x) = c_1\phi_1(x) + c_2\phi_2(x) + \dots + c_n\phi_n(x)$$

$$y_n(x) = \sum_{i=1}^n c_i\phi_i(x) \quad (1.5)$$

Donde las constantes  $c_i$  son desconocidas.

Cuando  $y(x)$  es aproximada por  $y_n(x)$ , el funcional  $I_{(y)}$  se hace función de las constantes  $c_i$ . Entonces, el problema se transforma en conseguir el valor estacionario de una función de un número finito de parámetros  $c_k$ , los cuales se consiguen a partir de las siguientes condiciones:

$$\frac{\delta I}{\delta c_k} = 0, \quad k = 1, 2, \dots, n \quad (1.6)$$

Resultando en un grupo de ecuaciones lineales donde las incógnitas son las  $c_i$ . El procedimiento a seguir es restringir el rango de funciones que estamos comparando a la familia finita de funciones  $\phi_1, \phi_2, \phi_3, \dots, \phi_n$ , luego, la mejor solución es aquella que hace a  $I(c_1, c_2, c_3, \dots, c_n)$  estacionario.

Las funciones de prueba  $\phi_i(x)$  son conocidas y su forma es elegida arbitrariamente, además se eligen de forma de satisfacer las condiciones de borde para cualquier valor de  $c_i$ .

### ***1.3 Principio del Trabajo Virtual.***

En la sección anterior analizamos el Cálculo Variacional y las variaciones. Podemos relacionar los desplazamientos nodales con las variaciones de los funcionales. Cuando trabajamos con mecánica de sólidos, estamos trabajando con fuerzas y desplazamientos. Estos desplazamientos los consideramos como las variaciones de las cantidades reales. El proceso de aproximar el comportamiento de un medio continuo por el Método de los Elementos Finitos, se puede introducir mediante aplicaciones físicas específicas o como un concepto matemático. Hemos introducido una serie de aproximaciones, pero generalmente no es fácil asegurar que las funciones de desplazamientos elegidas satisfacen las condiciones de continuidad. El procedimiento más sencillo es suponer un desplazamiento arbitrario, el cual llamaremos *desplazamiento virtual*, a los nodos e igualar el trabajo exterior realizado por las fuerzas nodales al trabajo efectuado en el interior por las tensiones y fuerzas distribuidas. **El principio del Trabajo virtual** se puede enunciar de la siguiente forma:

“Un cuerpo se encuentra en equilibrio si el trabajo interno virtual es igual al trabajo externo virtual, para cualquier campo de desplazamientos admisible cinemáticamente.”

### 1.3.1 Trabajo y Energía.

El trabajo que una fuerza  $F$  realiza en una partícula en un desplazamiento infinitesimal desde  $A$  hasta  $B$ ,  $\Delta u = \Delta u_{\hat{i}} + \Delta u_{\hat{j}} + \Delta u_{\hat{k}}$  se define como:

$$\begin{aligned}\Delta W &= F \cdot \Delta u \\ \Delta W &= F_x \Delta u + F_y \Delta v + F_z \Delta w \\ W &= \int_A^B F \cdot du = \int_A^B (F_x du + F_y dv + F_z dw)\end{aligned}\tag{1.7}$$

El trabajo total de la partícula se define como la suma del trabajo interno y externo, de tal modo que:

$$W = W_{\text{int}} + W_{\text{ext}}\tag{1.8}$$

Supongamos que un sistema se compone de  $N$  partículas, y además sufre un desplazamiento infinitesimal  $\Delta u_s = \Delta u_{s\hat{i}} + \Delta v_{s\hat{j}} + \Delta w_{s\hat{k}}$ ,  $s = 1, 2, 3, \dots, N$ , el trabajo total sería:

$$\Delta W = \sum_{s=1}^N F_s \cdot \Delta u_s = \sum_{s=1}^N (F_{x_s} \Delta u_s + F_{y_s} \Delta v_s + F_{z_s} \Delta w_s)\tag{1.9}$$

Siendo  $N$  el número de grados de libertad. Si un sistema mecánico tiene un número de grados de libertad finito, su configuración puede ser especificada por un número finito ( $N$ ) de variables reales llamadas *coordenadas generalizadas*. El concepto se basa en que en cualquier instante el sistema se puede describir completamente mediante el valor particular de estas coordenadas, esto puede ser considerado como definir un punto en un espacio  $N$  dimensional. El movimiento del sistema es



equivalente al movimiento de este punto. Estas coordenadas generalizadas las describiremos mediante unos vectores  $q_i$  de dimensión  $N$ .

Ahora, por comodidad, introducimos las coordenadas generalizadas  $q_i$ . Los desplazamientos pueden ser expresados en términos de  $\Delta q_i$ , los cuales son vectores linealmente independientes. Entonces, el desplazamiento puede ser expresado en términos de  $\Delta q_i$ . Podemos expresar el trabajo como:

$$\Delta W = Q_1 \Delta q_1 + Q_2 \Delta q_2 + \dots = \sum_{j=1}^n Q_j \Delta q_j \quad (1.10)$$

Los factores  $Q_j$  son llamados los componentes de las *fuerzas generalizadas*, debido a que si cada uno es multiplicado por su correspondiente incremento de coordenada generalizada el resultado es trabajo.

Usualmente en ingeniería trabajamos en términos de la función de trabajo negativo. Por ende:

$$\begin{aligned} V(q_1, q_2, \dots, q_n) &= -W(q_1, q_2, \dots, q_n) \\ Q_i &= -\frac{\delta V}{\delta q_i} \end{aligned} \quad (1.11)$$

Hasta ahora no se han introducido restricciones con respecto a las fuerzas. Se puede simplificar el análisis introduciendo restricciones a las fuerzas que se están utilizando. Se van a restringir las fuerzas a *fuerzas conservativas*. De este modo se puede introducir el concepto de *energía potencial*. Una fuerza conservativa es aquella en

que el trabajo producido por mover dos configuraciones distintas depende solamente en los puntos desplazados y no en el camino tomado para moverlos.

Como las fuerzas son conservativas, podemos decir que  $\Pi(q_1, q_2, \dots, q_n)$  es una función potencial y que su valor es la **Energía Potencial** de las fuerzas.

Los cuerpos elásticos son aquellos que recobran completamente su forma original cuando las fuerzas son retiradas. Cuando se retiran las fuerzas, las fuerzas internas producen trabajo neto nulo. Estas fuerzas son conservativas y contienen una Energía Potencial. Esta Energía Potencial es:

$$\Pi = U_e + \Omega \quad (1.12)$$

Donde  $U_e$  es la energía interna de deformación y  $\Omega$  es el potencial de las fuerzas externas.

### 1.3.2 El Principio del trabajo virtual para un sólido continuo.

El trabajo realizado por una fuerza en un desplazamiento virtual se llama *Trabajo Virtual* de forma de:

$$\begin{aligned} \delta W &= Q_1 \delta q_1 + Q_2 \delta q_2 + \dots + Q_N \delta q_N \\ \delta W &= \sum_{i=1}^N Q_i \delta q_i \end{aligned} \quad (1.13)$$

Donde:

$Q_i$  = Es la i-ésima fuerza generalizada.

$\delta q_i$  = Desplazamiento virtual en la  $i$ -ésima coordenada generalizada.

El Principio del Trabajo Virtual es el método mediante el cual se identifica la configuración real del sólido analizado. Éste enuncia que, si un sistema mecánico se encuentra en su configuración de equilibrio, el trabajo virtual de las fuerzas que sobre él actúan son cero en un desplazamiento virtual. Entonces, para una condición de equilibrio, nos queda:

$$\delta W = \sum_{i=1}^N Q_i \delta q_i = 0 \quad (1.14)$$

Como los  $\delta q_i$  son arbitrarios e independientes, esto implica que:

$$Q_i = 0 \quad i = 1, 2, \dots, n \quad (1.15)$$

Podemos concluir del Principio del Trabajo Virtual que:

*“Un sistema mecánico finito se encuentra en equilibrio si, y sólo si, las fuerzas generalizadas desaparecen idénticamente”.*

Cuando el sistema contiene miembros deformables, la ecuación (1.14) se transforma en:

$$\begin{aligned} \delta W &= \delta W_{\text{int}} + \delta W_{\text{ext}} \\ \delta W &= \sum_{i=1}^N Q_i \delta q_i + \sum_{i=1}^N P_i \delta q_i \end{aligned} \quad (1.16)$$

Donde las  $Q_i$  son las fuerzas generalizadas internas y  $P_i$  las fuerzas generalizadas externas.

Y la ecuación (1.16) se transforma en:

$$\left(Q_j\right)_{\text{int}} + \left(Q_j\right)_{\text{ext}} = 0 \quad (1.17)$$

Cuando el cuerpo que se está analizando es elástico y se cumplen las leyes de Hooke, las fuerzas internas van a tener un potencial, y si las fuerzas son conservativas, la ecuación (1.17) se transforma en:

$$\delta\Pi = \delta(U_e + \Omega) = 0 \quad (1.18)$$

### ***1.4 Energía potencial Total.***

En mecánica de sólidos, nuestro problema es determinar el desplazamiento  $u$  del cuerpo que se está estudiando, satisfaciendo las ecuaciones de equilibrio. Los esfuerzos están relacionados a las deformaciones unitarias que, a su vez, están relacionadas con los desplazamientos. Lo anterior nos conduce a resolver ecuaciones diferenciales parciales de segundo orden, a las soluciones de este conjunto de ecuaciones se les llama generalmente una solución exacta. Tales soluciones existen para geometrías y condiciones de cargas simples. Pero para problemas que posean geometrías más complejas con condiciones de frontera y con cargas generales, la obtención de tales soluciones es una tarea muy complicada, por esto los métodos de solución aproximada usualmente emplean métodos de energía potencial, los cuales imponen condiciones menos estrictas sobre las funciones.

Podemos definir el trabajo como el cambio negativo en la función  $\Pi$ , de forma que:

$$W = -\Delta\Pi \quad (1.19)$$

La función  $\Pi$ , es llamada Función Potencial (también llamada *Energía Potencial Total*). Con esta definición podemos calcular  $F$  como el gradiente de la función potencial:

$$\vec{F} = -\nabla\Pi \quad (1.20)$$

La *Energía Potencial Total*  $\Pi$  de un cuerpo elástico es definido como:

$$\Pi = (\text{Energía de deformación unitaria}) + (\text{Potencial de trabajo})$$

$$\Pi = U_e + \Omega \quad (1.21)$$

Para materiales elástico lineales la energía de deformación unitaria  $U_e$  o el Potencial de Fuerzas Internas, por unidad de volumen, en el cuerpo es  $\frac{1}{2}\sigma^T \epsilon$ . Para el material elástico lineal:

$$U_e = \frac{1}{2} \int_V \sigma^T \epsilon dV \quad (1.22)$$

$$W_i = -\Delta U_e$$

El Potencial de trabajo o la energía Potencial de las Cargas Externas ( $\Omega$ ) viene dado por la expresión:

$$\Omega = -\int_V u^T F dV - \int_S u^T T ds - \sum_i u_i^T P_i \quad (1.23)$$

$$W_e = -\Delta\Omega$$

donde  $F$  es la fuerza distribuida por unidad de volumen,  $T$  corresponde a la tracción superficial que puede darse por los valores de sus componentes en puntos superficiales.

De esta forma el potencial total para un sólido en tres dimensiones queda:

$$\Pi = \frac{1}{2} \int_V \sigma^t \varepsilon dV - \int_V u^T F dV - \int_S u^T T ds - \sum_i u_i^T P_i \quad (1.24)$$

Utilizando esto, podemos calcular el trabajo total como:

$$W = W_i + W_e = -\Delta U - \Delta \Omega = -\Delta \Pi \quad (1.25)$$

Concluyendo que:

*“El trabajo realizado por todas las fuerzas en un sistema es igual al cambio negativo en el potencial total de ese sistema”.*

### ***1.5 Principio de la Energía Potencial Mínima.***

Para sistemas conservativos, de todos los campos de desplazamientos cinemáticamente admisibles, aquellos que corresponden a condiciones de equilibrio extremizan la energía potencial total. Si la condición extrema es un mínimo, el estado de equilibrio es estable. Recordemos que el Principio del Trabajo Virtual nos dice que:

$$\delta W = \text{Trabajo Virtual} = 0$$

Para cada desplazamiento virtual si, y sólo si el sistema se encuentra en equilibrio. Cuando expresamos este principio en términos de las coordenadas generalizadas, nos queda que:

$$\delta W = Q_1 \delta q_1 + Q_2 \delta q_2 + \dots + Q_N \delta q_N$$

Donde  $Q_i$  son las fuerzas generalizadas. Recordemos también que se requiere que  $Q_i = 0$ .

Como ahora se ha definido una Función Potencial para el sistema y ya no se necesita calcular el Trabajo Virtual, podemos investigar los cambios correspondientes en la Función Potencial producido por los desplazamientos virtuales. Podemos expresar las fuerzas generalizadas como la derivada direccional de una Función Potencial Total en la dirección de la correspondiente coordenada generalizada. Esto es:

$$Q_i = -\frac{\delta \Pi}{\delta q_i} \quad (1.26)$$

Introduciendo esto en la expresión de  $\delta W$  e introduciendo la condición  $\delta W = 0$ , nos queda:

$$\begin{aligned} 0 = \delta W &= -\frac{\partial \Pi}{\partial q_1} \delta q_1 - \frac{\partial \Pi}{\partial q_2} \delta q_2 - \dots - \frac{\partial \Pi}{\partial q_N} \delta q_N \\ &= -\sum_{i=1}^N \frac{\partial \Pi}{\partial q_i} \delta q_i \\ &\underline{0 = -\delta \Pi} \end{aligned} \quad (1.27)$$

La expresión de  $\delta\Pi$  en (1.28) se le suele llamar “Primera Variación” en la Función Potencial Total  $\Pi$ .

Se puede concluir que:

*“Un sistema se encuentra en equilibrio si, y solo si, el cambio (primera variación  $\delta\Pi$ ) en el Potencial Total,  $\Pi$ , es cero para cada desplazamiento virtual”.*

Esto significa que cada una de las derivadas respecto a las coordenada generalizadas,  $\frac{\partial\Pi}{\partial q_i}$ , deben ser cero. En consecuencia la función Potencial Total debe tener un valor estacionario.

### ***1.6 El Método de los Elementos Finitos.***

El problema principal respecto a la utilización del método de Rayleigh-Ritz, corresponde a la selección de las funciones  $\phi_i(x)$ , las cuales no sólo deben satisfacer las condiciones de borde del problema, sino que también deben poder representar en forma apropiada otras características del mismo. Por otra parte, para mejorar las soluciones numéricas es necesario utilizar funciones de orden cada vez mayor. El trabajo correspondiente al manejo de funciones de orden alto puede ser considerable, razón por la cual muchas veces no es factible utilizar dichos métodos.

El Método de los Elementos Finitos se basa en utilizar dichas aproximaciones por partes, en lugar de efectuarlas de manera global. El concepto básico consiste primero en dividir el dominio bajo estudio en un número finito de subregiones de tamaño finito (denominadas elementos finitos). Luego, sobre cada una de esas zonas, y en



forma aislada, se asume un comportamiento aproximado para la incógnita del problema. Al efectuar la división en zonas, es posible adoptar dentro de cada una de ellas, funciones simples que representen de manera aproximada el comportamiento local.

La ventaja de este método consiste en que para mejorar la solución ya no es imprescindible utilizar funciones de orden mayor, sino que basta con emplear una división más fina del dominio, con un número mayor de subregiones, pero manteniendo funciones de aproximación de orden bajo.

### ***1.7 Formulación del método de elementos finitos.***

A continuación se describe el desarrollo de la formulación del Método de los Elementos Finitos en problemas de elasticidad lineal. Por razones de sencillez se describen dominios bidimensionales.

Para construir el modelo de los elementos finitos primero se realiza la división del dominio de integración (región bidimensional) en triángulos de lados rectos, como se muestra en la Figura 1-5. Donde están conectados a través de sus fronteras y en sus vértices (nodos). Cada triángulo formado en este mallado típico se llama elemento. En los problemas de elasticidad, se puede considerar al elemento como una región donde existe un campo de desplazamientos y los nodos o vértices son puntos del espacio donde se desea determinar las componentes de desplazamiento. Se permite que cada nodo se desplace en las dos direcciones  $x$  y  $y$ , por lo tanto cada nodo posee dos grados de libertad.

La idea básica del *Método de los Elementos Finitos* puede interpretarse como un método de aproximación, donde las funciones de prueba del método de Rayleigh-Ritz

se definen en forma local en cada elemento y son llamadas *funciones de interpolación*. Estas funciones de interpolación se combinan para dar lugar a una aproximación por trozos.

Empleando notación matricial, se describirá la formulación del Método de los Elementos finitos aplicado a problemas de elasticidad bidimensional. En esta formulación se incluye el ensamblaje de elementos, la imposición de condiciones de contorno, la solución del sistema de ecuaciones para obtener las cantidades nodales y el procesamiento de elementos para obtener cantidades tales como los esfuerzos.

### 1.7.1 Expresión matricial de la energía potencial total.

Consideremos un cuerpo plano que puede representarse mediante un dominio bidimensional  $A$  discretizado mediante elementos finitos. La energía potencial total  $\Pi$  de un cuerpo elástico de comportamiento lineal viene dada por la suma de la energía potencial de deformación  $U_e$  y de la energía potencial  $\Omega$  asociada al trabajo de las fuerzas externas.

$$\Pi = U_e + \Omega \quad (1.28)$$

La energía potencial de deformación  $U_e$  se puede expresar como:

$$U_e = \frac{1}{2} \int_A \boldsymbol{\varepsilon}^T D \boldsymbol{\varepsilon} h dA \quad (1.29)$$

donde  $\boldsymbol{\varepsilon}$  es el vector de deformaciones,  $D$  es la matriz constitutiva,  $\boldsymbol{\sigma}$  es el vector de esfuerzos, y  $h$  es el espesor.

$$\varepsilon = \begin{Bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{Bmatrix}, D = \frac{E}{1-\nu^2} \begin{bmatrix} 1 & \nu & 0 \\ \nu & 1 & 0 \\ 0 & 0 & \frac{(1-\nu)}{2} \end{bmatrix} \quad (1.30)$$

$$\sigma = \varepsilon^T D$$

La energía potencial  $\Omega$  asociada al trabajo de las fuerzas externas es:

$$\Omega = -\int_A f^T u h dA - \int_L t^T u h dL \quad (1.31)$$

donde  $f$  es el vector de fuerzas de volumen,  $t$  es el vector de fuerzas de superficie o tracción superficial y  $u$  es el vector de desplazamientos.

$$f = \begin{Bmatrix} f_x \\ f_y \end{Bmatrix}, t = \begin{Bmatrix} t_x \\ t_y \end{Bmatrix}, u = \begin{Bmatrix} u \\ v \end{Bmatrix} \quad (1.32)$$

Luego la energía potencial total se puede expresar como:

$$\Pi = U_e + \Omega = \frac{1}{2} \int_A \varepsilon^T D \varepsilon h dA - \int_A f^T u h dA - \int_L t^T u h dL \quad (1.33)$$

Si usamos una aproximación por elementos finitos es necesario dividir el dominio  $A$  en elementos y podemos expresar la energía potencial total como:

$$\Pi = U_e + \Omega = \sum_{e=1}^{nel} \frac{1}{2} \int_{A_e} \varepsilon^T D \varepsilon h dA - \int_{A_e} f^T u h dA - \int_{L_e} t^T u h dL \quad (1.34)$$

Siendo  $nel$  el número de elementos,  $A_e$  la región ocupada por cada elemento y  $L_e$  su contorno sometido a fuerzas de tracción superficial.

Para poder obtener una aproximación por elementos finitos debemos aplicar el método de Rayleigh-Ritz utilizando los campos de desplazamientos  $u$  formados por las funciones de interpolación de los elementos finitos.

### 1.7.2 Funciones de interpolación.

A continuación se describirán los conceptos de funciones de interpolación y su continuidad para elementos finitos. Los desplazamientos que se producen dentro de un elemento necesitan ser representados en términos de los desplazamientos nodales  $a_i$  del elemento, lo cual se expresa de la siguiente manera:

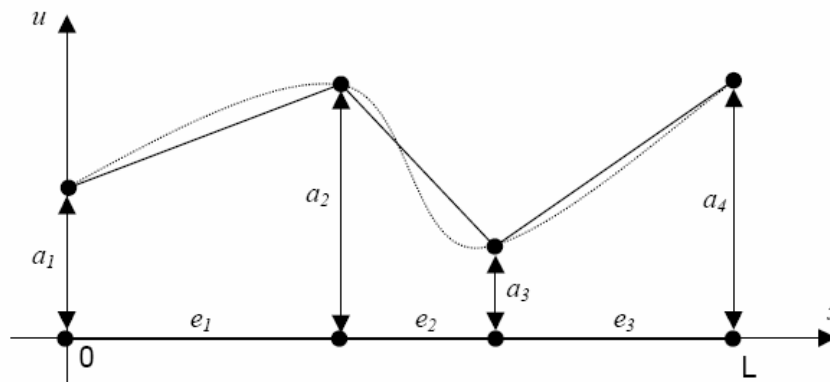
$$u = \sum_{i=1}^n a_i N_i(x, y) \quad (1.35)$$

Donde se han asumido implícitamente que las funciones de prueba  $N_i$  están definidas por una expresión simple válida en todo el dominio  $A$ . Esto puede ser posible en el caso de dominios con geometría sencilla como triángulos, rectángulos, círculos, elipses, pero no en el caso de geometrías más complejas. Para el caso de triángulos de deformación constante, las funciones de prueba resultan ser lineales sobre el elemento.

Una forma alternativa de definir las funciones de prueba consiste en subdividir el dominio  $A$  en una serie de subdominios o *elementos*  $A_e$  que no se solapen, y luego las aproximaciones  $u$  se construyen por trozos usando definiciones simples de las

funciones de prueba sobre estos subdominios. Si los subdominios son de forma relativamente simple y la definición de las funciones de prueba sobre estos subdominios puede ser hecha de manera repetitiva, es posible aproximar dominios complejos de forma bastante directa. En este método de aproximación las funciones de prueba se definen en forma local en cada elemento. Se aplica el Método de Rayleigh-Ritz con estas funciones de prueba y se logra una aproximación por trozos. Estas funciones de prueba en elementos finitos las llamaremos funciones de interpolación.

Considérese, por ejemplo, un dominio unidimensional, esto es una recta de longitud  $L$ , sobre la cual queremos aproximar una función arbitraria  $u(x)$  mediante una aproximación lineal por trozos. Para ello subdividimos esta recta en  $m$  elementos de recta vinculados por sus extremos. Nótese que hemos asociado los valores de las coordenadas generalizadas  $a_i$  a los valores de la aproximación en los extremos de los elementos. Estos puntos de cada elemento que tienen asociados valores de las coordenadas generalizadas son llamados *nodos* del elemento. Por tanto, asociaremos las funciones de prueba con puntos genéricos del elemento y las funciones de interpolación con los nodos del elemento.

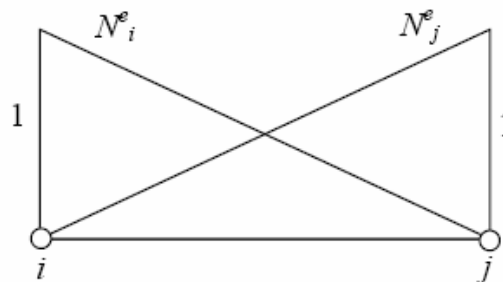


**Figura 1-2** Aproximación de  $u(x)$  con elementos lineales.

Si observamos la definición (1.35) de las funciones de aproximación es evidente que en un nodo  $i$  asociado a la coordenada generalizada  $a_i$ , la función de prueba  $N_i$  debe valer 1 y el resto de las funciones de prueba  $N_j$  ( $j \neq i$ ) deben ser nulas en ese punto. Luego, una vez identificados los puntos nodales, es muy sencillo definir las funciones de prueba  $N_i$  ya que deben valer 1 en el punto nodal asociado y 0 en el resto de los puntos nodales del elemento.

También puede observarse que las únicas funciones de interpolación que son diferentes de cero en cada elemento son aquellas asociadas con los nodos de ese elemento. Luego, en cada elemento  $e$  con nodos  $i, j$  la aproximación puede ser expresada simplemente en función de dos funciones de interpolación lineales del elemento  $N_{ei}$ ,  $N_{ej}$  y de los valores nodales  $a_i$ ,  $a_j$  como:

$$u^e = a_i N_i^e + a_j N_j^e \quad (1.36)$$



**Figura 1-3.** Funciones de forma locales  $N_{e_i}$  lineales en cada elemento.

La aproximación global se genera a partir de la combinación de las funciones de interpolación locales en cada elemento. Por otro lado, observemos que si tenemos dos valores nodales por elemento podemos reproducir cualquier variación lineal sobre

este elemento. En particular, si tenemos un valor constante de la aproximación  $u$  sobre el elemento  $e$  esto implica que los valores nodales deben ser iguales a este valor, esto es:

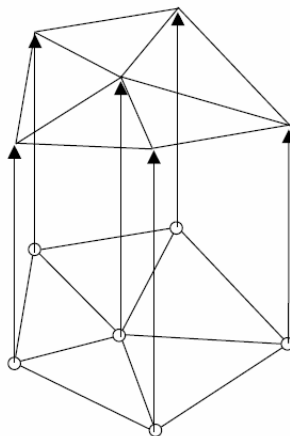
$$u^e = c = a_i N_i^e + a_j N_j^e = c(N_i^e + N_j^e) = cte \quad (1.37)$$

Luego, sobre cada elemento se debe verificar

$$(N_i^e + N_j^e) = 1 \quad (1.38)$$

Esto es, la suma de las funciones de interpolación de cada elemento debe ser igual a uno.

La extensión de estos conceptos a dos dimensiones es bastante directa. En este caso, la subdivisión del dominio se efectúa utilizando triángulos ó cuadriláteros.



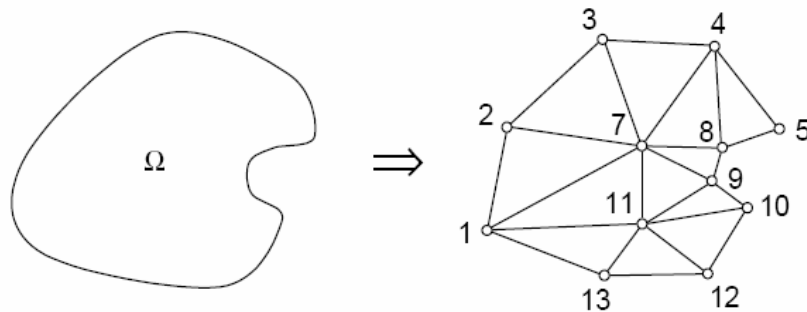
**Figura 1-4.** Aproximación en dos dimensiones. La parte de abajo es el dominio bidimensional y la parte de arriba es la función de forma.

En este caso los puntos nodales quedan asociados, en general, a los vértices de la malla y para el caso de aproximación lineal sobre cada triángulo las funciones de interpolación del elemento son planos.

### 1.7.3 Aproximación por elementos finitos.

El primer paso para obtener una aproximación por elementos finitos es realizar una *discretización* del dominio. Esto es, debemos generar una malla de elementos finitos que cubra todo el dominio. Además debemos numerar los *nodos* de la malla, que son aquellos puntos que tienen asociadas coordenadas generalizadas. Para el caso particular de análisis de esfuerzos, las coordenadas generalizadas son los *desplazamientos nodales*. Así, para el nodo  $i$  sus desplazamientos nodales serán  $u_i$  y  $v_i$ . Consideremos una aproximación por elementos finitos de los desplazamientos  $\mathbf{u} = (u, v)$ , como:

$$u = \sum_{i=1}^n u_i N_i(x, y), v = \sum_{i=1}^n v_i N_i(x, y) \quad (1.39)$$



**Figura 1-5.** Discretización de un dominio  $\Omega$



donde  $u_i, v_i$  son los desplazamientos nodales. Cada función de prueba  $N_i$  se compone de las funciones de forma asociadas al nodo  $i$  de todos los elementos que contienen ese nodo, esto es:

$$N_i(x, y) = \sum_{e=1}^{nen} N_i^e(x, y) \quad (1.40)$$

donde  $nen$  es el número de elementos que contienen al nodo  $i$ .

Además, los desplazamientos  $u_e, v_e$  en cada elemento se pueden expresar como:

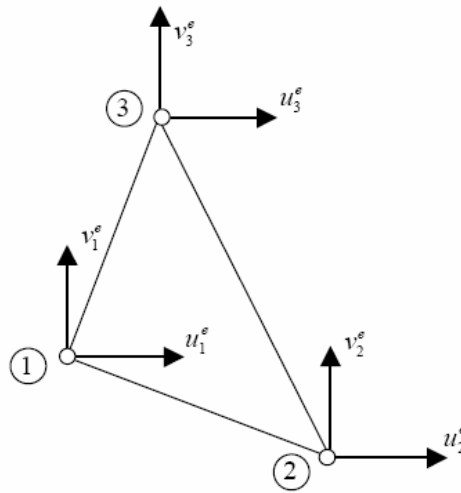
$$u^e = \sum_{j=1}^{nmod} u_j^e N_j^e(x, y), v^e = \sum_{j=1}^{nmod} v_j^e N_j^e(x, y) \quad (1.41)$$

donde  $nmod$  es el número de nodos del elemento. Nótese que en este caso el índice  $j$  se refiere a la numeración local del nodo en el elemento. Así, por ejemplo, para un triángulo de 3 nodos los desplazamientos sobre el elemento son:

$$\begin{aligned} u^e &= N_1^e u_1^e + N_2^e u_2^e + N_3^e u_3^e \\ v^e &= N_1^e v_1^e + N_2^e v_2^e + N_3^e v_3^e \end{aligned} \quad (1.42)$$

Esta ecuación puede escribirse matricialmente como:

$$u^e = N^e d^e \quad (1.43)$$



**Figura 1-6.** Desplazamientos nodales para el triángulo de 3 nodos.

siendo  $N_e$  la matriz de funciones de forma del elemento:

$$N^e = \begin{bmatrix} N_1^e & 0 & N_2^e & 0 & N_3^e & 0 \\ 0 & N_1^e & 0 & N_2^e & 0 & N_3^e \end{bmatrix} \quad (1.44)$$

y  $d_e$  es el vector de desplazamientos nodales del elemento:

$$d^{eT} = \{u_1^e \quad v_1^e \quad u_2^e \quad v_2^e \quad u_3^e \quad v_3^e\} \quad (1.45)$$

En forma particionada la matriz de funciones de forma se puede escribir como:

$$N^e = [N_1^e \quad N_2^e \quad N_3^e] \quad (1.46)$$

donde las submatrices  $N_i^e$  que están asociadas a cada nodo del elemento son:

$$N_i^e = \begin{bmatrix} N_i^e & 0 \\ 0 & N_i^e \end{bmatrix} \quad (1.47)$$

El vector  $d_e$  de desplazamientos nodales del elemento se puede también expresar en forma particionada como:

$$d^e = \begin{Bmatrix} d_1^e \\ d_2^e \\ d_3^e \end{Bmatrix} \quad (1.48)$$

Donde los vectores  $d_i^e$  que están asociados a los desplazamientos de cada nodo  $i$  del elemento son:

$$d_i^e = \begin{Bmatrix} u_1^e \\ v_i^e \end{Bmatrix} \quad (1.49)$$

siendo  $u_i^e$  los desplazamientos del nodo  $i$  del elemento.

#### 1.7.4 Matriz gradiente.

Si reemplazamos los campos de desplazamientos aproximados por elementos finitos en las expresiones de las deformaciones, en cada elemento tenemos:

$$\begin{aligned}
 \varepsilon_x &= \frac{\partial u^e}{\partial x} = \frac{\partial N_1^e}{\partial x} u_1^e + \frac{\partial N_2^e}{\partial x} u_2^e + \frac{\partial N_3^e}{\partial x} u_3^e \\
 \varepsilon_y &= \frac{\partial v^e}{\partial y} = \frac{\partial N_1^e}{\partial y} v_1^e + \frac{\partial N_2^e}{\partial y} v_2^e + \frac{\partial N_3^e}{\partial y} v_3^e \\
 \gamma_{xy} &= \frac{\partial u^e}{\partial y} + \frac{\partial v^e}{\partial x} = \frac{\partial N_1^e}{\partial y} u_1^e + \frac{\partial N_1^e}{\partial x} v_1^e + \frac{\partial N_2^e}{\partial y} u_2^e + \frac{\partial N_2^e}{\partial x} v_2^e + \frac{\partial N_3^e}{\partial y} u_3^e + \frac{\partial N_3^e}{\partial x} v_3^e
 \end{aligned} \tag{1.50}$$

y en forma matricial

$$\varepsilon = \begin{Bmatrix} \frac{\partial u^e}{\partial x} \\ \frac{\partial v^e}{\partial y} \\ \frac{\partial u^e}{\partial x} + \frac{\partial v^e}{\partial y} \end{Bmatrix} = \begin{bmatrix} \frac{\partial N_1^e}{\partial x} & 0 & \frac{\partial N_2^e}{\partial x} & 0 & \frac{\partial N_3^e}{\partial x} & 0 \\ 0 & \frac{\partial N_1^e}{\partial y} & 0 & \frac{\partial N_2^e}{\partial y} & 0 & \frac{\partial N_3^e}{\partial y} \\ \frac{\partial N_1^e}{\partial y} & \frac{\partial N_1^e}{\partial x} & \frac{\partial N_2^e}{\partial y} & \frac{\partial N_2^e}{\partial x} & \frac{\partial N_3^e}{\partial y} & \frac{\partial N_3^e}{\partial x} \end{bmatrix} \begin{Bmatrix} u_1^e \\ v_1^e \\ u_2^e \\ v_2^e \\ u_3^e \\ v_3^e \end{Bmatrix} \tag{1.51}$$

y en forma abreviada

$$\varepsilon = B^e d^e \tag{1.52}$$

donde  $B^e$  es la *matriz gradiente* (relación deformación-desplazamiento) del elemento y  $d^e$  es el vector de desplazamientos nodales el elemento. En forma particionada la matriz gradiente se puede escribir como:

$$B^e = [B_1^e \quad B_2^e \quad B_3^e] \tag{1.53}$$

Donde las submatrices  $B_i^e$  que están asociadas a cada nodo del elemento son:

$$B_i^e = \begin{bmatrix} \frac{\partial N_i^e}{\partial x} & 0 \\ 0 & \frac{\partial N_i^e}{\partial y} \\ \frac{\partial N_i^e}{\partial y} & \frac{\partial N_i^e}{\partial x} \end{bmatrix} \quad (1.54)$$

Observemos que un caso general la matriz gradiente del elemento  $B^e$  estará compuesta de tantas submatrices  $B_i^e$  como nodos tenga el elemento.

### 1.7.5 Matriz de rigidez y vector de cargas nodales equivalentes.

Si reemplazamos los campos de desplazamientos aproximados por elementos finitos de (1.52) y (1.43) en la expresión de la energía potencial total tenemos:

$$\Pi = U_e + \Omega = \sum_{e=1}^{nel} \frac{1}{2} \int_A \overbrace{d^{eT} B^{eT}}^{\varepsilon^T} D \overbrace{B^e d^e}^{\varepsilon} h dA - \int_A \overbrace{d^{eT} N^{eT}}^{u^{eT}} f h dA - \int_L \overbrace{d^{eT} N^{eT}}^{u^{eT}} t h dL \quad (1.55)$$

Definiendo a la *matriz de rigidez* del elemento como:

$$K^e = \int_A B^{eT} D B^e h dA \quad (1.56)$$

Esta matriz es una matriz cuadrada de dimensión igual a la cantidad de desplazamientos nodales del elemento y definiendo además al *vector de cargas nodales equivalentes* del elemento como:

$$f^e = \int_A N^{eT} f h dA - \int_L N^{eT} t h dL \quad (1.57)$$

luego la energía potencial total se puede expresar como:

$$\Pi = \sum_{e=1}^{nel} \frac{1}{2} d^{eT} K^e d^e - d^{eT} f^e \quad (1.58)$$

Si empleamos la forma particionada (1.46) para las matrices gradiente del elemento  $B^e$  entonces la matriz de rigidez del elemento  $K^e$  se puede expresar en forma particionada como:

$$K^e = \begin{bmatrix} K_{11}^e & K_{12}^e & K_{13}^e \\ K_{21}^e & K_{22}^e & K_{23}^e \\ K_{31}^e & K_{32}^e & K_{33}^e \end{bmatrix} \quad (1.59)$$

Siendo

$$K_{ij}^e = \int_A B_i^{eT} D B_j^e h dA \quad (1.60)$$

La submatriz de rigidez del elemento que relaciona los nodos numerados localmente como  $i, j$  en el elemento. Si definimos al vector  $d$  de desplazamientos de la malla con  $n$  nodos como:

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_n \end{bmatrix} \quad (1.61)$$

Entonces la energía potencial de deformación se puede expresar como:

$$U_e = \sum_{e=1}^{nel} \frac{1}{2} d^{eT} K^e d^e = \frac{1}{2} d^T K d \quad (1.62)$$

siendo  $K$  la *matriz de rigidez global* formada por las submatrices  $K_{ij}$ , las cuales valen:

$$K_{ij} = \sum_{e=1}^{nel} K_{ij}^e \quad (1.63)$$

Esto es, si dos nodos están vinculados por un elemento, entonces dicho elemento debe contribuir con una submatriz a la matriz de rigidez global.

Por otro lado, la energía potencial de las fuerzas externas se puede expresar como:

$$\Omega = -d^T f \quad (1.64)$$

siendo  $\mathbf{f}$  el *vector de fuerzas externas* global cuyas componentes  $f_i$  valen

$$f_i = \sum_{e=1}^{nel} f_i^e \quad (1.65)$$

Finalmente, la energía potencial total queda:

$$\Pi = \frac{1}{2} d^T K d - d^T f \quad (1.66)$$

Aplicando Rayleigh-Ritz debemos minimizar esta expresión respecto de las coordenadas generalizadas, que en este caso son los desplazamientos nodales  $d$ , esto es:

$$\frac{\partial \Pi}{\partial d} = K d - f = 0 \quad (1.67)$$

Resultando el siguiente sistema de ecuaciones:

$$K d = f \quad (1.68)$$

Que tiene por incógnitas a los desplazamientos nodales de toda la malla. En general, algunos de estos desplazamientos tendrán valores prescritos por lo que no serán incógnitas, en este caso deberíamos eliminar la línea correspondiente a este desplazamiento de la matriz de rigidez global. Obsérvese que el primer paso para resolver este sistema de ecuaciones es el montaje de la matriz  $K$  y del vector  $f$  a partir de las contribuciones de los elementos, este proceso se denomina *ensamblaje*.

### 1.7.6 Características de la matriz de rigidez global.

Una matriz de rigidez global tiene muchas características generales que pueden usarse para comprobar la formulación de una matriz particular de rigidez. Una matriz de rigidez global debe tener las siguientes propiedades:



- *Simétrica*. Esto significa que  $k_{ij} = k_{ji}$ . Éste es siempre el caso cuando los desplazamientos son directamente proporcionales a las cargas aplicadas.
- *Cuadrada*. Esto implica que el número de filas es igual al número de columnas en la matriz.
- *Singular*. La matriz de rigidez global es singular (la determinante de la matriz es igual a cero). Debido a que ninguna de las restricciones (los desplazamientos prescritos y / o la rotación) han sido aplicadas.
- *Diagonal principal positiva*. Todos los elementos de la diagonal principal deben ser positivos. Si  $k_{ii}$  es negativo, entonces la fuerza y su desplazamiento correspondiente estarían en direcciones opuestas, lo cual es físicamente irracional. Si  $k_{ii} = 0$ , entonces el desplazamiento no produciría una fuerza de reacción, lo cual implicaría que la estructura es inestable.

### ***1.8 Convergencia de los Resultados.***

En general, se puede decir que para obtener una solución mediante el Método de los Elementos Finitos, se requiere idealizar el problema físico para obtener así una descripción mecánica, y luego, la solución de esa idealización mecánica utilizando el Método de los Elementos Finitos.

Una solución apropiada debe converger hacia la solución exacta a medida que se aumenta el número de elementos, entendiéndose por exacta, la solución de las ecuaciones diferenciales que rigen el comportamiento del sistema idealizado.

En el caso de que no sea posible obtener una solución exacta de la idealización mecánica, la convergencia del modelo basado en el Método de los Elementos Finitos puede medirse únicamente por el hecho de que todas las condiciones básicas de tipo

cinemático, estático y constitutivo deben ser satisfechas en el límite, cuando el número de elementos tiende a infinito. Al hablar de la convergencia de un modelo de elementos finitos, se entiende por error el que se produce como consecuencia de sustituir un modelo discreto por un modelo continuo, además del error proveniente de asumir las funciones de aproximación de los desplazamientos.

Entonces, se entiende por convergencia el hecho de que el error disminuya a medida que se use un número cada vez mayor de elementos más pequeños. Las pruebas matemáticas de la convergencia asumen que el proceso de afinamiento de la malla está definido por tres condiciones:

- La reducción del tamaño de los elementos debe hacerse de manera que cada punto del dominio de solución quede siempre dentro de cada nueva malla.
- Cada nueva malla debe estar contenida dentro de la anterior.
- Durante el proceso, la forma de las funciones de interpolación debe permanecer invariable.

Notación para expresar el grado de continuidad de la variable de interés:

Si la variable es continua en las fronteras entre elementos, se habla de continuidad  $C^0$ . Si, además, sus primeras derivadas deben ser continuas, se habla de continuidad  $C^1$ .

Entonces, para garantizar convergencia monotónica y para darle significado al ensamblaje de elementos, las funciones  $N_i$  (funciones de forma) deben cumplir los siguientes requisitos:

- Compatibilidad: Debe haber continuidad  $C^r$  en las fronteras entre elementos.

- Completitud: Dentro de un elemento debe haber continuidad  $C^{r+1}$ .

El requisito de completitud significa que las funciones de desplazamiento dentro del elemento deben poder representar:

- Desplazamientos de cuerpo rígido.
- Estados de deformación constante.

Los desplazamientos de cuerpo rígido son aquellos modos de desplazamiento que el cuerpo debe poder experimentar sin que se generen esfuerzos en su interior. Por otra parte, podemos entender la necesidad de estados de deformación constante, imaginando que representamos el cuerpo con un número creciente de elementos. En el límite, cuando el tamaño del elemento tiende a cero, la deformación del mismo alcanza un valor constante, lo que permite representar cualquier variación compleja del estado de deformaciones en el cuerpo.

El requisito de compatibilidad implica que debe haber continuidad de los desplazamientos dentro de los elementos, y a través de las fronteras entre los mismos.

### ***1.9 Exactitud de la solución.***

Una solución más exacta se puede obtener aumentando el número de elementos o utilizando elementos de geometría más complicada. Se ha de notar que en los nodos el Método de los Elementos Finitos provee valores exactos de  $u$ . Los elementos finitos que tengan funciones de interpolación lineales producen valores nodales exactos, si la solución buscada es cuadrática. Los elementos cuadráticos producen valores nodales exactos si la solución es cúbica, etc.

## ***1.10 Aplicación del método de los elementos finitos al análisis de problemas de elasticidad bidimensional.***

A continuación se presenta la aplicación del método de los elementos finitos a un problema de elasticidad bidimensional.

### **1.10.1 Coeficientes de rigidez para un elemento de esfuerzo-deformación para un elemento triangular general.**

La suposición más simple para expresar la deformación dentro de un elemento es una condición de tensión completamente uniforme. Esto puede ser logrado asumiendo que los desplazamientos de  $u$  y  $v$  son funciones lineales de  $x$  y  $y$ .

Por consiguiente:

$$\begin{aligned} u &= \alpha_1 + \alpha_2 x + \alpha_3 y \\ v &= \beta_1 + \beta_2 x + \beta_3 y \end{aligned} \quad (1.69)$$

Donde  $\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3$  son constantes.

El estado de deformación en el elemento correspondiente a las funciones (1.69) es:

$$\begin{aligned} \varepsilon_x &= \frac{\partial u}{\partial x} = \alpha_2 \\ \varepsilon_y &= \frac{\partial v}{\partial y} = \beta_3 \\ \gamma_{xy} &= \frac{\partial u}{\partial y} + \frac{\partial v}{\partial x} = \alpha_3 + \beta_2 \end{aligned} \quad (1.70)$$

Puede verse que las deformaciones son constantes, por lo que el elemento es llamado un triángulo de deformación constante. Por lo tanto las deformaciones  $\varepsilon_x, \varepsilon_y, \gamma_{xy}$  pueden ser obtenidas en términos de los desplazamientos nodales  $u_1, v_1, u_2, v_2, u_3, v_3$ . Aplicando las ecuaciones (1.69) en cada nodo obtendremos:

$$\begin{aligned}
 u_1 &= \alpha_1 + \alpha_2 x_1 + \alpha_3 y_1 \\
 v_1 &= \beta_1 + \beta_2 x_1 + \beta_3 y_1 \\
 u_2 &= \alpha_1 + \alpha_2 x_2 + \alpha_3 y_2 \\
 v_2 &= \beta_1 + \beta_2 x_2 + \beta_3 y_2 \\
 u_3 &= \alpha_1 + \alpha_2 x_3 + \alpha_3 y_3 \\
 v_3 &= \beta_1 + \beta_2 x_3 + \beta_3 y_3
 \end{aligned}
 \tag{1.71}$$

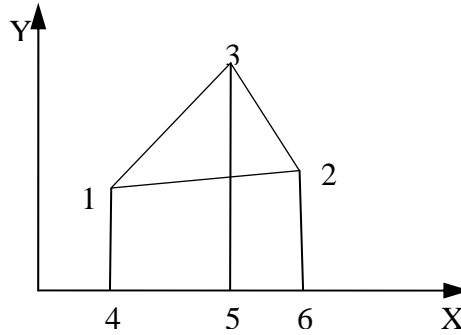
Podemos escribirlo de otra forma:

$$\{u\} = \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_1 & y_1 \\ 1 & x_2 & y_2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_2 & y_2 \\ 1 & x_3 & y_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & x_3 & y_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} = [A]\{\alpha\}
 \tag{1.72}$$

De una forma similar podemos escribir las deformaciones (1.70) como:

$$\{\varepsilon\} = \begin{bmatrix} \varepsilon_x \\ \varepsilon_y \\ \gamma_{xy} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix} \{\alpha\} = [B_\alpha]\{\alpha\}
 \tag{1.73}$$

Las constantes  $\{\alpha\}$  pueden ser eliminadas para así poder expresar la deformación en términos de los desplazamientos nodales  $\{U\}$ . Antes de realizar esto el área del elemento se expresa en términos de las coordenadas nodales como se muestra en la Figura 1-7.



**Figura 1-7** Área del elemento calculada en términos de las coordenadas nodales.

$$\text{Área del triángulo (123)} = \text{área (1354)} + \text{área (2356)} - \text{área (1462)}$$

Donde el área del triángulo (123) = A

$$A = \frac{1}{2} [(y_1 + y_3)(x_3 - x_1) + (y_3 + y_2)(x_2 - x_3) - (y_1 - y_2)(x_2 - x_1)]$$

Nótese que el área A es positiva, si el elemento es numerado en sentido horario.

Regresando a la (1.72), las constantes  $\alpha_2, \alpha_3, \beta_2, \beta_3$  pueden ser solucionadas en términos de otras variables:

$$\begin{aligned}
\alpha_2 &= \frac{1}{2} A [(y_2 - y_3)u_1 + (y_3 - y_1)u_2 + (y_1 - y_2)u_3] \\
\alpha_3 &= -\frac{1}{2} A [(x_2 - x_3)u_1 + (x_3 - x_1)u_2 + (x_1 - x_2)u_3] \\
\beta_2 &= \frac{1}{2} A [(y_2 - y_3)v_1 + (y_3 - y_1)v_2 + (y_1 - y_2)v_3] \\
\beta_3 &= -\frac{1}{2} A [(x_2 - x_3)v_1 + (x_3 - x_1)v_2 + (x_1 - x_2)v_3]
\end{aligned} \tag{1.74}$$

La sustitución de ecuaciones (1.74) en (1.70) da las deformaciones del elemento en términos de los desplazamientos nodales.

$$\begin{aligned}
\varepsilon_x &= \frac{1}{2} A [(y_2 - y_3)u_1 + (y_3 - y_1)u_2 + (y_1 - y_2)u_3] \\
\varepsilon_y &= \frac{1}{2} A [(y_2 - y_3)v_1 + (y_3 - y_1)v_2 + (y_1 - y_2)v_3]
\end{aligned} \tag{1.75}$$

$$\gamma_{xy} = \frac{1}{2} A [(x_3 - x_2)u_1 + (x_1 - x_3)u_2 + (x_2 - x_1)u_3 + (y_2 - y_3)v_1 + (y_3 - y_1)v_2 + (y_1 - y_2)v_3]$$

En términos matriciales:

$$\{\varepsilon\} = \frac{1}{2A} \begin{bmatrix} b_1 & 0 & b_2 & 0 & b_3 & 0 \\ 0 & a_1 & 0 & a_2 & 0 & a_3 \\ a_1 & b_1 & a_2 & b_2 & a_3 & b_3 \end{bmatrix} \{U\} = [B] \{U\} \tag{1.76}$$

Donde:

$$\begin{aligned}
 a_1 &= x_3 - x_2 \\
 a_2 &= x_1 - x_3 \\
 a_3 &= x_2 - x_1 \\
 b_1 &= y_2 - y_3 \\
 b_2 &= y_3 - y_1 \\
 b_3 &= y_1 - y_2
 \end{aligned}
 \tag{1.77}$$

Para determinar los esfuerzos que corresponden a estas deformaciones, es necesario indagar en relaciones específicas de esfuerzo/deformación. Las ecuaciones constitutivas son para un material isotrópico elástico y lineal con un módulo de elasticidad ( $E$ ), modulo de rigidez ( $G$ ) y la relación de Poisson ( $\nu$ ). Los estados limitantes bidimensionales son estado plano de ( $\sigma$ ) y estado plano de ( $\varepsilon$ ).

En el plano de esfuerzo, el pequeño esfuerzo sometido  $\sigma_2 = 0$  y los pequeños desplazamientos sometidos no son restricciones.

De esa manera:

$$\begin{aligned}
 \sigma_x &= \frac{E}{1-\nu^2} (\varepsilon_x + \nu \varepsilon_y) \\
 \sigma_y &= \frac{E}{1-\nu^2} (\varepsilon_y + \nu \varepsilon_x) \\
 \tau_{xy} &= G \gamma_{xy}
 \end{aligned}
 \tag{1.78}$$

En forma matricial

$$\{\sigma\} = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \tau_{xy} \end{bmatrix} = \begin{bmatrix} \frac{E}{1-\nu^2} & \frac{\nu E}{1-\nu^2} & 0 \\ \frac{\nu E}{1-\nu^2} & \frac{E}{1-\nu^2} & 0 \\ 0 & 0 & \frac{E}{2(1-\nu)} \end{bmatrix} \{\varepsilon\} = [D] \{\varepsilon\}
 \tag{1.79}$$



Sabiendo que  $E = 2G(1+\nu)$ , los coeficientes de la matriz  $D$  para un estado plano de esfuerzo son:

$$\begin{aligned}d_{11} &= d_{22} = \frac{E}{1-\nu^2} \\d_{12} &= d_{21} = \frac{\nu E}{1-\nu^2} \\d_{33} &= G = \frac{E}{2(1+\nu)}\end{aligned}\tag{1.80}$$

En deformación plana,  $\varepsilon_z = 0$  y  $\sigma_y = \nu(\sigma_x + \sigma_y)$  no hay desplazamientos a través del grosor. Ya que  $\sigma_z$  se puede derivar de  $\sigma_x, \sigma_y$ , no se incluye en el vector  $\{\sigma\}$ . Los coeficientes en la matriz  $[D]$  para un plano de deformación son:

$$\begin{aligned}d_{11} &= d_{22} = \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} \\d_{12} &= d_{21} = \frac{\nu E}{(1+\nu)(1-2\nu)} \\d_{33} &= \frac{E}{2(1+\nu)}\end{aligned}\tag{1.81}$$

Quedando de la siguiente forma:

$$[D] = \begin{bmatrix} \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & \frac{\nu E}{(1+\nu)(1-2\nu)} & 0 \\ \frac{\nu E}{(1+\nu)(1-2\nu)} & \frac{E(1-\nu)}{(1+\nu)(1-2\nu)} & 0 \\ 0 & 0 & \frac{E}{2(1+\nu)} \end{bmatrix}\tag{1.82}$$

Nótese la inclusión del término  $(1 - 2\nu)$  en los denominadores, esto quiere decir que no hay solución posible para la condición limitativa de  $\nu = 0,5$ , donde no hay cambio volumétrico matricial en el material durante la deformación.

Sustituyendo:

$$\{\varepsilon\} = [B]\{U\}$$

en

$$\{\sigma\} = [D]\{\varepsilon\}$$

Nos queda:

$$\{\sigma\} = [D]\{\varepsilon\} = [D][B]\{U\} \quad (1.83)$$

Lo que quiere decir que el estado de  $(\sigma)$  en el elemento, expresado en términos de desplazamientos nodales nos queda:

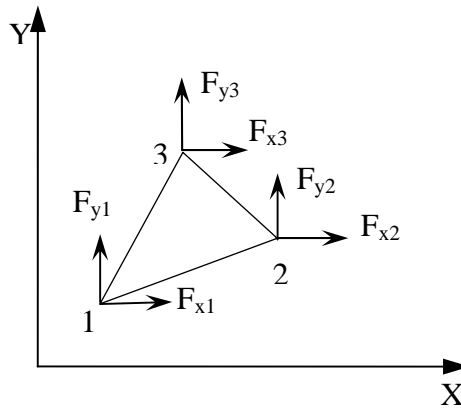
$$\begin{aligned} \sigma_x &= \frac{1}{2A} \{d_{11}(y_2 - y_3)u_1 + d_{11}(y_3 - y_1)u_2 + d_{11}(y_1 - y_2)u_3 + d_{12}(x_3 - x_2)v_1 + d_{12}(x_1 - x_3)v_2 + d_{12}(x_2 - x_1)v_3\} \\ \sigma_y &= \frac{1}{2A} \{d_{12}(y_2 - y_3)u_1 + d_{12}(y_3 - y_1)u_2 + d_{12}(y_1 - y_2)u_3 + d_{11}(x_3 - x_2)v_1 + d_{11}(x_1 - x_3)v_2 + d_{11}(x_2 - x_1)v_3\} \quad (1.84) \\ \tau_{xy} &= \frac{1}{2A} \{d_{33}(x_3 - x_2)u_1 + d_{33}(x_1 - x_3)u_2 + d_{33}(x_2 - x_1)u_3 + d_{33}(y_2 - y_3)v_1 + d_{33}(y_3 - y_1)v_2 + d_{33}(y_1 - y_2)v_3\} \end{aligned}$$

Quedando como ultimo paso, relacionar las fuerzas nodales correspondientes a los esfuerzos antes mencionados con los desplazamientos nodales, y por lo tanto determinar los coeficientes de la matriz de rigidez. Para obtener el arreglo de las

fuerzas nodales en equilibrio con los esfuerzos en el elemento, se utiliza la ecuación de *trabajo virtual*.

Para condiciones de esfuerzo plano y fuerzas puntuales únicamente en los límites, ver (Figura 1-8) la ecuación de *trabajo virtual es*:

$$\int_v (\sigma_x \varepsilon_x + \sigma_y \varepsilon_y + \tau_{xy} \gamma_{xy}) dV - F_{x1}u_1 - F_{x2}u_2 - F_{x3}u_3 - F_{y1}v_1 - F_{y2}v_2 - F_{y3}v_3 = 0 \quad (1.85)$$



**Figura 1-8** Condiciones de esfuerzo plano y fuerzas puntuales únicamente en los límites.

Visto de la forma matricial:

$$\iiint \{\varepsilon\}^T \{\sigma\} dx dy dz - \{U\}^T \{P\} = 0 \quad (1.86)$$

Donde:

$$\{P\} = \begin{bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \\ x_3 \\ y_3 \end{bmatrix}, \{\varepsilon\}^T = [\varepsilon_x \quad \varepsilon_y \quad \gamma_{xy}] \text{ y } \{U\}^T = [u_1 \quad v_1 \quad u_2 \quad v_2 \quad u_3 \quad v_3]$$

Ahora en la ecuación, el esfuerzo  $\{\sigma\}$  y las fuerzas nodales externas  $\{P\}$  forman cualquier arreglo de equilibrio y las deformaciones  $\{\varepsilon\}$  con los desplazamientos nodales  $\{U\}$  forman cualquier deformación compatible o un arreglo geométrico para este elemento en particular. La ventaja de la ecuación de *trabajo virtual* es que para un problema en particular, el arreglo de equilibrio correspondiente a una carga puede ser multiplicado por el arreglo de deformación correspondiente a otra carga, y el resultado seguirá siendo cero. Esta propiedad es usada para determinar los coeficientes de rigidez, seleccionando un arreglo de deformación  $\{U'\}$  correspondiente a algunas cargas inusuales  $\{P'\}$  donde se simplifica de una forma considerable la relación entre  $\{\sigma\}$  (escrita en términos de  $\{U\}$  (1.84)) y  $\{P\}$ .

Para  $F_{x1}$  tenemos:

$$\sigma_x, \sigma_y, \tau_{xy} \text{ con } F_{x1}, F_{y1}, F_{x2}, F_{y2}, F_{x3}, F_{y3}$$

De la geometría obtenemos:

$$u'_1 \neq 0, v'_1 = u'_2 = v'_2 = u'_3 = v'_3 = 0$$

Producido por alguna carga inusual ( $F'_{x1}, F'_{y1}, \text{ect}$ ), que de hecho no necesita estar determinado aquí. De la ecuación (1.74):

$$\begin{aligned}\varepsilon'_x &= \frac{1}{2A}(y_2 - y_3)u'_1 \\ \varepsilon'_y &= 0 \\ \tau'_{xy} &= \frac{1}{2A}[(x_3 - x_2)u'_1]\end{aligned}$$

Quedando la ecuación (1.85) de la siguiente forma:

$$\begin{aligned}\int_v [\sigma_x \varepsilon'_x + \sigma_y \varepsilon'_y + \tau_{xy} \gamma'_{xy}] dv - F_{x1} u'_1 &= 0 \\ \int_v \left[ \sigma_x \frac{1}{2A}(y_2 - y_3)u'_1 + 0 + \tau_{xy} \frac{1}{2A}(x_3 - x_2)u'_1 \right] dv - F_{x1} u'_1 &= 0 \\ F_{x1} &= \frac{1}{2A} \int_v [\sigma_x (y_2 - y_3) + \tau_{xy} (x_3 - x_2)] dv\end{aligned}$$

Como en el elemento  $\sigma_x$  y  $\tau_{xy}$  son constantes:

$$F_{x1} = \frac{V}{2A} [\sigma_x (y_2 - y_3) + \tau_{xy} (x_3 - x_2)]$$

Siendo el volumen del elemento  $V = A.h$ . Donde  $h$  es el espesor. Usando la ecuación (1.84), se puede escribir en términos de los desplazamientos nodales:

$$F_{x1} = \frac{h}{4A} \left\{ \begin{array}{l} u_1 \left[ d_{11} (y_2 - y_3)^2 + d_{33} (x_3 - x_2)^2 \right] \\ + v_1 \left[ d_{12} (x_3 - x_2)(y_2 - y_3) + d_{33} (y_2 - y_3)(x_3 - x_2) \right] \\ + u_2 \left[ d_{11} (y_3 - y_1)(y_2 - y_3) + d_{33} (x_1 - x_3)(x_3 - x_2) \right] \\ + v_2 \left[ d_{12} (x_1 - x_3)(y_2 - y_3) + d_{33} (y_3 - y_1)(x_3 - x_2) \right] \\ + u_3 \left[ d_{11} (y_1 - y_2)(y_2 - y_3) + d_{33} (x_2 - x_1)(x_3 - x_2) \right] \\ + v_3 \left[ d_{12} (x_2 - x_1)(y_2 - y_3) + d_{33} (y_1 - y_2)(x_3 - x_2) \right] \end{array} \right\} \quad (1.87)$$

De una forma similar obtendremos las  $(F_{y1}, F_{x2}, F_{y2}, F_{x3}, F_{y3})$

A continuación se presentará un ejemplo ilustrativo.

Con la finalidad de ilustrar la aplicación de los elementos finitos a un problema en particular, se considera la siguiente estructura:

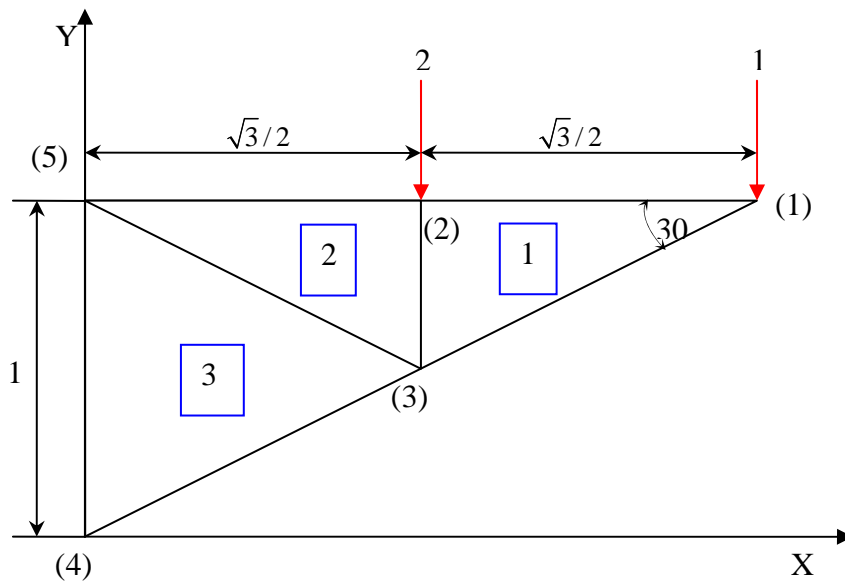
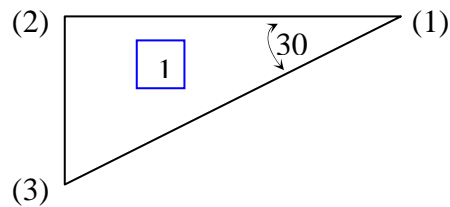


Figura 1-9 Estructura del ejemplo ilustrativo.

Se enumeran los nodos de (1) al (5), se definen los elementos como se muestra en la Figura 1-9. Los elementos son numerados en sentido antihorario. Para el elemento  $\boxed{1}$  son (1),(2),(3), para el elemento  $\boxed{2}$  son (5),(3),(2) y para el elemento  $\boxed{3}$  son (5),(4),(3). Se asume que la viga posee unidades pequeñas y está en un estado plano de esfuerzos con  $E=1$  y  $\nu = 0,3$ .

Considerar cada elemento en orden y establecer la matriz de rigidez.

Para el elemento  $\boxed{1}$ :



**Figura 1-10** Elemento 1.

$$\begin{aligned}
 x_1 &= \sqrt{3} \\
 x_2 &= \sqrt{3}/2 \\
 x_3 &= \sqrt{3}/2 \\
 y_1 &= 1 \\
 y_2 &= 1 \\
 y_3 &= 1/2
 \end{aligned}$$

$$\begin{aligned}a_1 &= x_3 - x_2 = 0 \\a_2 &= x_1 - x_3 = \sqrt{3}/2 \\a_3 &= x_2 - x_1 = -\sqrt{3}/2 \\b_1 &= y_2 - y_3 = 1/2 \\b_2 &= y_3 - y_1 = -1/2 \\b_3 &= y_1 - y_2 = 0\end{aligned}$$

Donde el valor del área es:

$$A = \frac{Base \times Altura}{2} = \frac{(\sqrt{3}/2) \times (1/2)}{2} = 0,2165$$

Y la del espesor :

$$h = 1$$

Usado las ecuaciones obtenidas en (1.87), se coloca de forma matricial. Se obtiene:





$$\begin{bmatrix}
 1^X_1 \\
 1^Y_1 \\
 1^X_2 \\
 1^Y_2 \\
 1^X_3 \\
 1^Y_3
 \end{bmatrix}
 = \frac{1}{4.0,2165}
 \begin{bmatrix}
 d_{11}(1/2)^2 + d_{33}(0)^2 & d_{12}(0)(1/2) + d_{33}(1/2)(0) & d_{12}(\sqrt{3}/2)(1/2) + d_{33}(-1/2)(0) & d_{11}(-1/2)(1/2) + d_{33}(\sqrt{3}/2)(0) & d_{12}(0)(1/2) + d_{33}(1/2)(0) & d_{12}(\sqrt{3}/2)(1/2) + d_{33}(-1/2)(0) \\
 d_{12}(0)(1/2) + d_{33}(1/2)(0) & d_{11}(0)^2 + d_{33}(1/2)^2 & d_{11}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{12}(-1/2)(0) + d_{33}(\sqrt{3}/2)(1/2) & d_{11}(0)(1/2) + d_{33}(1/2)(0) & d_{11}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) \\
 d_{12}(\sqrt{3}/2)(1/2) + d_{33}(-1/2)(0) & d_{11}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{11}(\sqrt{3}/2)(-1/2) + d_{33}(\sqrt{3}/2) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) \\
 d_{11}(-1/2)(1/2) + d_{33}(\sqrt{3}/2)(0) & d_{12}(-1/2)(0) + d_{33}(\sqrt{3}/2)(1/2) & d_{11}(-1/2)^2 + d_{33}(\sqrt{3}/2)^2 & d_{11}(-1/2) + d_{33}(\sqrt{3}/2) & d_{11}(-1/2)(-1/2) + d_{33}(\sqrt{3}/2) & d_{12}(\sqrt{3}/2)(-1/2) + d_{33}(-1/2) \\
 d_{12}(0)(1/2) + d_{33}(1/2)(0) & d_{11}(0)(1/2) + d_{33}(1/2)(0) & d_{11}(0)(1/2) + d_{33}(1/2)(0) & d_{12}(-1/2)(0) + d_{33}(\sqrt{3}/2)(1/2) & d_{11}(0)(1/2) + d_{33}(1/2)(0) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) \\
 d_{12}(\sqrt{3}/2)(1/2) + d_{33}(-1/2)(0) & d_{11}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{11}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{11}(\sqrt{3}/2)(-1/2) + d_{33}(\sqrt{3}/2) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2) & d_{12}(\sqrt{3}/2)(0) + d_{33}(-1/2)(1/2)
 \end{bmatrix}
 \begin{bmatrix}
 u_1 \\
 v_1 \\
 u_2 \\
 v_2 \\
 u_3 \\
 v_3
 \end{bmatrix}$$

Donde  ${}_1F_{x_2}$  significa que la fuerza  $F$  es aplicada por el elemento 1 en el nodo 2. Así con cada uno de ellos se usa la ecuación (1.86). Se realiza como ejemplo el cálculo de  ${}_1F_{x_1}$ ,

$${}_1F_{x_1} = \frac{1}{0,866} \left( \frac{1}{4} d_{11} u_1 + 0 \cdot v_1 - \frac{1}{4} d_{11} u_2 + \frac{\sqrt{3}}{4} d_{12} v_2 + 0 \cdot u_3 - \frac{\sqrt{3}}{4} d_{12} v_3 \right)$$

Del plano de esfuerzos tenemos:

$$d_{11} = \frac{E}{1-\nu^2} = 1,0989$$

$$d_{12} = \frac{\nu E}{1-\nu^2} = 0,3297$$

$$d_{33} = \frac{E}{2(1+\nu)} = 0,384$$

Obteniendo así la primera ecuación del sistema.

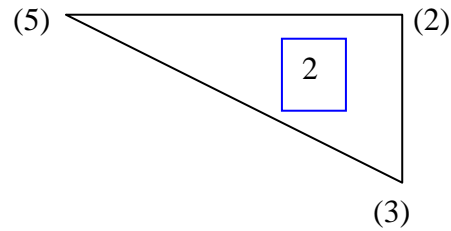
$$\therefore {}_1F_{x_1} = 0,317u_1 + 0v_1 - 0,317u_2 + 0,165v_2 + 0u_3 - 0,165u_3$$

Este proceso se repite para las otras ecuaciones restantes obteniendo el sistema total.

Como se presenta a continuación:

$$\begin{bmatrix} {}_1F_{x_1} \\ {}_1F_{y_1} \\ {}_1F_{x_2} \\ {}_1F_{y_2} \\ {}_1F_{x_3} \\ {}_1F_{y_3} \end{bmatrix} = \begin{bmatrix} 0,317 & 0 & -0,317 & 0,165 & 0 & -0,165 \\ 0 & 0,111 & 0,192 & -0,111 & -0,192 & 0 \\ -0,317 & 0,192 & 0,650 & -0,357 & -0,333 & 0,165 \\ 0,165 & -0,111 & -0,357 & 1,063 & 0,192 & -0,952 \\ 0 & -0,192 & -0,333 & 0,192 & 0,333 & 0 \\ -0,165 & 0 & 0,165 & -0,952 & 0 & 0,952 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{bmatrix} \quad (1.88)$$

Para el elemento  $\boxed{2}$ :



**Figura 1-11** Elemento 2.

$$\begin{aligned}x_5 &= x'_1 = 0 \\x_3 &= x'_2 = \sqrt{3}/2 \\x_2 &= x'_3 = \sqrt{3}/2 \\y_5 &= y'_1 = 1 \\y_3 &= y'_2 = 1/2 \\y_2 &= y'_3 = 1\end{aligned}$$

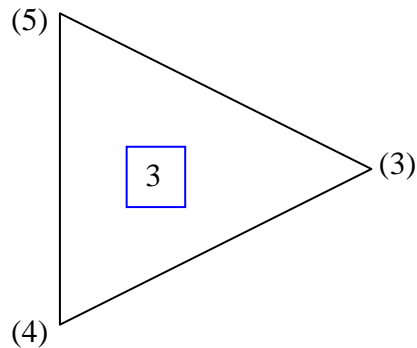
$$\begin{aligned}a_1 &= x'_3 - x'_2 = 0 \\a_2 &= x'_1 - x_3 = \sqrt{3}/2 \\a_3 &= x'_2 - x'_1 = \sqrt{3}/2 \\b_1 &= y'_2 - y'_3 = -1/2 \\b_2 &= y'_3 - y'_1 = 0 \\b_3 &= y'_1 - y'_2 = 1/2\end{aligned}$$

Realizando el mismo procedimiento que en el elemento  $\boxed{1}$ , obtenemos:

$$\begin{bmatrix} {}_2F_{x5} \\ {}_2F_{y5} \\ {}_2F_{x3} \\ {}_2F_{y3} \\ {}_2F_{x2} \\ {}_2F_{y2} \end{bmatrix} = \begin{bmatrix} 0,317 & 0 & 0 & 0,165 & -0,317 & -0,165 \\ 0 & 0,111 & 0,192 & 0 & -0,192 & -0,111 \\ 0 & 0,192 & 0,333 & 0 & -0,333 & -0,192 \\ 0,165 & 0 & 0 & 0,952 & -0,165 & -0,952 \\ -0,317 & -0,192 & -0,333 & -0,165 & 0,650 & 0,357 \\ -0,165 & -0,111 & -0,192 & -0,952 & 0,357 & 1,063 \end{bmatrix} \begin{bmatrix} u_5 \\ v_5 \\ u_3 \\ v_3 \\ u_2 \\ v_2 \end{bmatrix} \quad (1.89)$$

Nótese que en elemento  $\boxed{2}$ ,  $u_5 = v_5 = 0$  entonces, las columnas uno y dos del elemento de la matriz de rigidez podrán ser eliminadas. Además, las primeras dos ecuaciones no son usadas inicialmente, sólo se usan después de que obtengamos la solución de los desplazamientos desconocidos.

Para el elemento  $\boxed{3}$ :



**Figura 1-12** Elemento 3.

$$\begin{aligned}
 x_5 &= x'_1 = 0 \\
 x_4 &= x'_2 = 0 \\
 x_3 &= x'_3 = \sqrt{3}/2 \\
 y_5 &= y'_1 = 1 \\
 y_4 &= y'_2 = 0 \\
 y_3 &= y'_3 = 1/2
 \end{aligned}$$

$$\begin{aligned}
 a_1 &= x'_3 - x'_2 = \sqrt{3}/2 \\
 a_2 &= x'_1 - x_3 = -\sqrt{3}/2 \\
 a_3 &= x'_2 - x'_1 = 0 \\
 b_1 &= y'_2 - y'_3 = -1/2 \\
 b_2 &= y'_3 - y'_1 = -1/2 \\
 b_3 &= y'_1 - y'_2 = 1
 \end{aligned}$$

Realizando el mismo procedimiento que en el elemento [1] y [2], obtenemos:

$$\begin{bmatrix} {}_3F_{x5} \\ {}_3F_{y5} \\ {}_3F_{x4} \\ {}_3F_{y4} \\ {}_3F_{x3} \\ {}_3F_{y3} \end{bmatrix} = \begin{bmatrix} 0,325 & -0,179 & -0,008 & -0,014 & -0,317 & 0,192 \\ -0,179 & 0,531 & 0,014 & -0,420 & 0,165 & -0,111 \\ -0,008 & 0,014 & 0,325 & 0,179 & -0,317 & -0,192 \\ -0,014 & -0,420 & 0,179 & 0,531 & -0,165 & -0,111 \\ -0,317 & 0,165 & -0,317 & -0,165 & 0,634 & 0 \\ 0,192 & -0,111 & -0,192 & -0,111 & 0 & 0,222 \end{bmatrix} \begin{bmatrix} u_5 \\ v_5 \\ u_3 \\ v_3 \\ u_2 \\ v_2 \end{bmatrix} \quad (1.90)$$

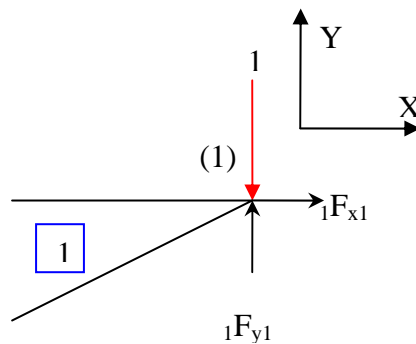
En elemento [3],  $u_4 = v_4 = u_5 = v_5 = 0$ , por lo tanto pueden ser eliminadas las columnas uno a la cuatro de la matriz de rigidez del elemento.

En este problema hay seis (6) desplazamientos nodales desconocidos, y por lo tanto se requiere seis ecuaciones, las cuales vienen de la aplicación de las condiciones de equilibrio en el nodo del desplazamiento desconocido y en la dirección

correspondiente. Ya que las incógnitas son  $u_1, v_1, u_2, v_2, u_3, v_3$ , se debe aplicar el equilibrio en ambas direcciones  $x$  y  $y$  en los nodos (1),(2) y (3).

### 1.10.2 Equilibrio en los nodos.

El equilibrio en el nodo (1) viene dado por la fuerza externa aplicada en el nodo  $F_{x1}$  y la fuerza aplicada por el elemento  ${}_1F_{x1}$ .



**Figura 1-13** Equilibrio en el nodo 1.

Donde el equilibrio está dado por:

$${}_1F_{x1} = F_{x1} = 0$$

Entonces, de (1.88):

$$0,317u_1 + 0v_1 - 0,317u_2 + 0,165v_2 + 0u_3 - 0,165v_3 = 0 \quad (1.91)$$

En la dirección de Y el equilibrio está dado por:

$${}_1F_{y1} = F_{y1} = -1$$

Y, de (1.88) resulta:

$$0u_1 + 0,111v_1 - 0,192u_2 - 0,111v_2 - 0,192u_3 + 0v_3 = -1 \quad (1.92)$$

De igual forma se realiza en el nodo (2) y (3).

Nodo (2)

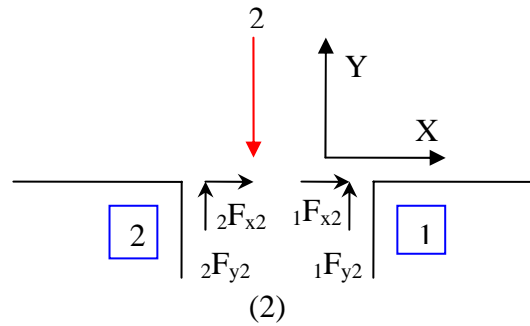


Figura 1-14 Equilibrio en el nodo 2.

En la dirección de X:

Condición de equilibrio:

$${}_1F_{x2} + {}_2F_{x2} = F_{x2} = 0$$

De (1.88) y (1.89), queda:

$$\begin{aligned} & -0,317u_1 + 0,192v_1 + (0,650 + 0,650)u_2 \\ & + (-0,357 + 0,357)v_2 + (-0,333 + 0,333)u_3 + (0,165 - 0,165)v_3 = 0 \end{aligned}$$

Resultando:

$$-0,317u_1 + 0,192v_1 + 1,300u_2 + 0v_2 + 0,666u_3 + 0v_3 = 0 \quad (1.93)$$



En la dirección de Y:

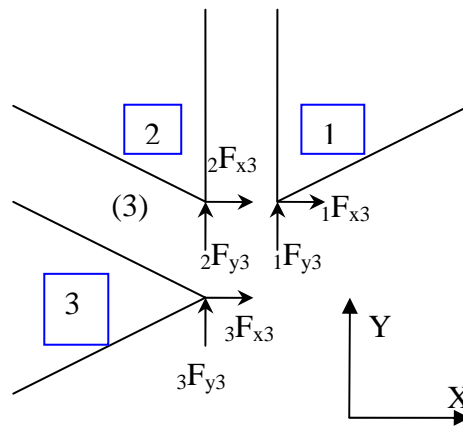
Condición de equilibrio:

$${}_1F_{y2} + {}_2F_{y2} = F_{y2} = -2$$

De (1.88) y (1.89), queda:

$$0,165u_1 - 0,111v_1 + 0u_2 + 2,125v_2 + 0u_3 - 1,903v_3 = -2 \quad (1.94)$$

Nodo (3):



**Figura 1-15** Equilibrio en el nodo 3.

En la dirección X:

Condición de equilibrio:

$${}_1F_{x3} + {}_2F_{x3} + {}_3F_{x3} = F_{x3} = 0$$

De (1.88) ,(1.89) y (1.90), queda:

$$0u_1 - 0,192v_1 - 0,666u_2 + 0v_2 + 1,301u_3 + 0v_3 = 0 \quad (1.95)$$

En la dirección Y:

Condición de equilibrio:

$${}_1F_{y3} + {}_2F_{y3} + {}_3F_{y3} = F_{y3} = 0$$

De (1.88) ,(1.89) y (1.90), queda:

$$-0,165u_1 + 0v_1 + 0u_2 - 1,903v_2 + 0u_3 + 2,125v_3 = 0 \quad (1.96)$$

Agrupando todas las ecuaciones, desde la (1.91) hasta la (1.96), obtenidas por el equilibrio de los nodos, se tiene el siguiente sistema de ecuaciones:

$$\begin{aligned} 0,317u_1 + 0v_1 - 0,317u_2 + 0,165v_2 + 0u_3 - 0,165v_3 &= 0 \\ 0u_1 + 0,111v_1 - 0,192u_2 - 0,111v_2 - 0,192u_3 + 0v_3 &= -1 \\ -0,317u_1 + 0,192v_1 + 1,300u_2 + 0v_2 + 0,666u_3 + 0v_3 &= 0 \\ 0,165u_1 - 0,111v_1 + 0u_2 + 2,125v_2 + 0u_3 - 1,903v_3 &= -2 \\ 0u_1 - 0,192v_1 - 0,666u_2 + 0v_2 + 1,301u_3 + 0v_3 &= 0 \\ -0,165u_1 + 0v_1 + 0u_2 - 1,903v_2 + 0u_3 + 2,125v_3 &= 0 \end{aligned}$$

El cual se puede expresar matricialmente de la siguiente manera:

$$\begin{bmatrix} 0,317 & 0 & -0,317 & 0,165 & 0 & -0,165 \\ 0 & 0,111 & 0,192 & -0,111 & 0,192 & 0 \\ 0,317 & 0,192 & 1,300 & 0 & -0,666 & 0 \\ 0,165 & -0,111 & 0 & 2,125 & 0 & -1,903 \\ 0 & -0,192 & -0,666 & 0 & 1,301 & 0 \\ -0,165 & 0 & 0 & -1,903 & 0 & 2,125 \end{bmatrix} \begin{bmatrix} u_1 \\ v_1 \\ u_2 \\ v_2 \\ u_3 \\ v_3 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \\ -2 \\ 0 \\ 0 \end{bmatrix}$$

Resultando las siguientes soluciones:

$$u_1 = 7,712$$

$$u_2 = 6,542$$

$$u_3 = -2,686$$

$$v_1 = -40,823$$

$$v_2 = -15,835$$

$$v_3 = -13,582$$

Con la finalidad de calcular los esfuerzos en cada elemento, se forma el vector de desplazamiento para cada elemento  $\{u_1, v_1, u_2, v_2, u_3, v_3\}$ , y los valores son sustituidos en la ecuación (1.83). La que resulta del producto  $[D][B]\{U\}$ . Dando los siguientes resultados:

	Elemento 1	Elemento 2	Elemento 3
$\sigma_x$	0	6,816	-3,408
$\sigma_y$	-4,505	-2,461	-1,022
$\tau_{xy}$	-4	0,065	-6,032

## 2. Métodos directos de solución de sistemas de ecuaciones.

En este capítulo se presentarán algunos de los métodos empleados para resolver los sistemas de ecuaciones que se generan en el análisis de problemas de ingeniería por el Método de los Elementos Finitos. Se presentan métodos directos de solución de sistemas de ecuaciones tales como, la descomposición LU, Gauss y Gauss-Jordan, los cuales son utilizados para resolver sistemas de ecuaciones lineales de la forma  $Ax = b$ .

### 2.1 Descomposición LU.

La descomposición LU es una técnica muy utilizada como paso intermedio para la resolución de sistemas lineales. La finalidad es el de descomponer la matriz  $\mathbf{K}$  en dos matrices triangulares (*una triangular inferior y otra triangular superior*):

$$\mathbf{K} = \mathbf{L}\mathbf{U} \quad (2.1)$$

Si el determinante la matriz  $\mathbf{K}$  es no nulo, existen una infinidad de matrices soluciones a esa igualdad, (*salvo casos particulares*). Estas matrices obtenidas  $\mathbf{L}$  y  $\mathbf{U}$  son también no singulares (*el determinante es diferente de cero*). Al ejecutar el producto descrito anteriormente si tiene lo siguiente:

$$\begin{aligned} K_{1j} &= u_{1j} \\ K_{1j} &= l_{i1}u_{11} \\ K_{ij} &= l_{ij}u_{jj} + \sum_{k=1}^{j-1} l_{ik}u_{kj} \\ K_{ij} &= u_{ij} + \sum_{k=1}^{i-j} l_{ik}u_{kj} \end{aligned}$$

## Matriz: Descomposición LU

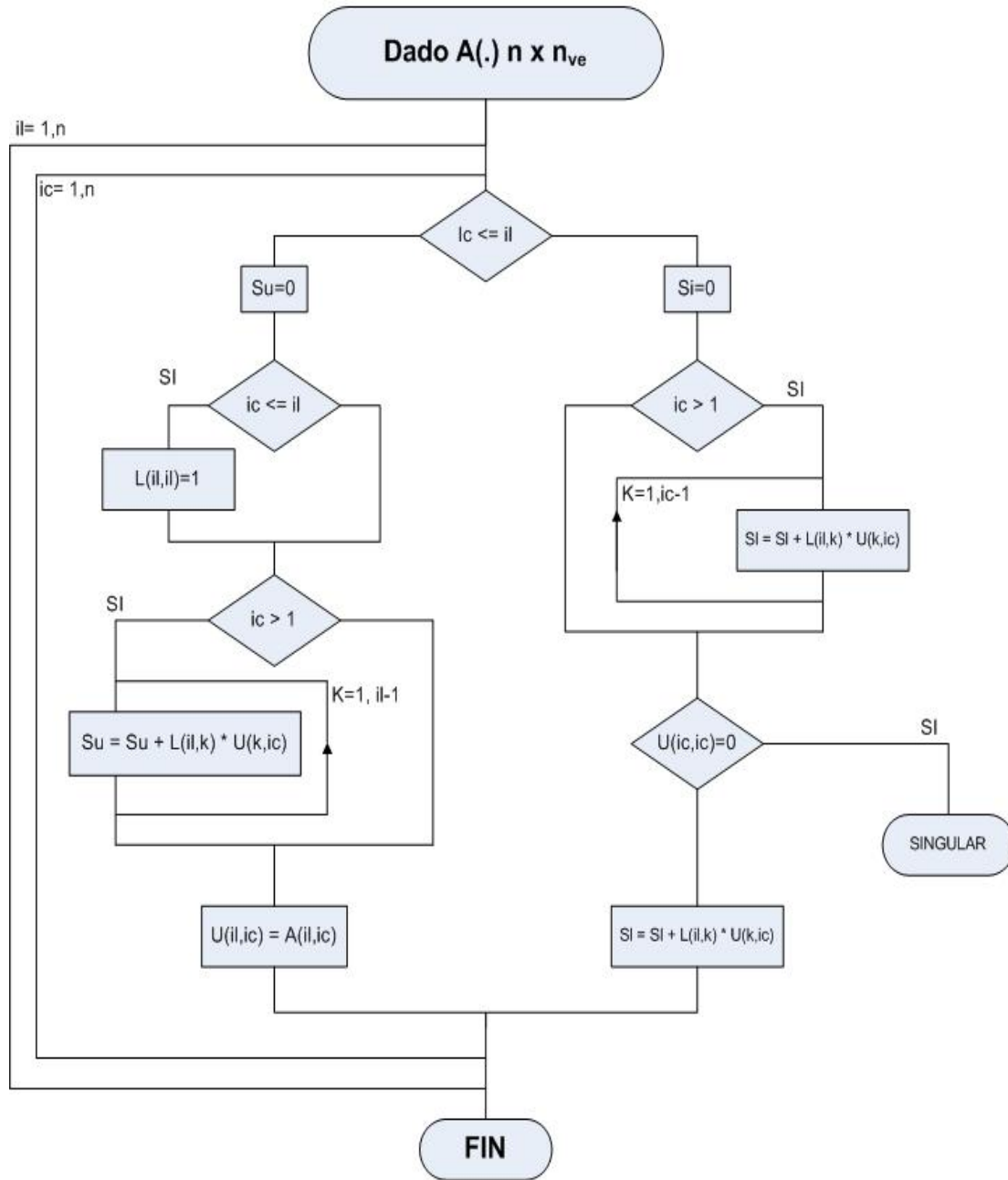


Figura 2-1 Diagrama de descomposición LU.

## 2.2 Eliminación de Gauss.

El método de eliminación de Gauss es el más eficiente de todos los métodos de solución para un sistema lineal en forma matricial ( $AX = B$ ). El método se basa en una sucesión de operaciones, con la finalidad de reemplazar todos los elementos que se encuentran por debajo de la diagonal principal por elementos nulos, esto garantiza un sistema equivalente, el cual permite determinar la solución completa por medio de una inspección.

$AX = B$ , donde  $A$  es  $(n \times n)$  y  $B$  es de  $(n \times 1)$ :

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

Para aplicar el método se siguen los siguientes pasos:

$$a_{ij}^{(1)} = a_{ij} - a_{1j} \frac{a_{i1}}{a_{11}} \quad (i \neq 1, 1 \leq j \leq n+1)$$

$$b_i = b_i - b_1 \frac{a_{i1}}{a_{11}}$$

Obteniendo la primera matriz intermedia del proceso. En donde se puede observar que la primera columna posee  $(n-1)$  valores nulos:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix}$$

Repetiendo el proceso anterior a la submatriz, la cual será de una dimensión menor cada vez que el proceso se repita. El procedimiento se puede poner de una forma más general:

$$a_{ij}^{(k-1)} = a_{ij}^{(k-1)} - a_{kj}^{(k-1)} \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

$$b_i^{(k)} = b_i^{(k-1)} - b_k^{(k-1)} \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

Una vez terminado este procedimiento se obtiene un sistema de la siguiente forma:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(n-1)} \end{bmatrix}$$

Donde la solución del sistema puede ser obtenida a partir del último vector. Para que el método de Gauss pueda ser aplicado, los valores de la diagonal principal no pueden ser nulos, ya que si existiera algún término nulo en la diagonal principal, sería necesario permutar por lo menos dos filas del sistema para obtener un valor no nulo. Si existiera un valor nulo en la diagonal principal entonces el sistema lineal no posee solución y es singular.

## Solución sistema lineal: Método de Gauss

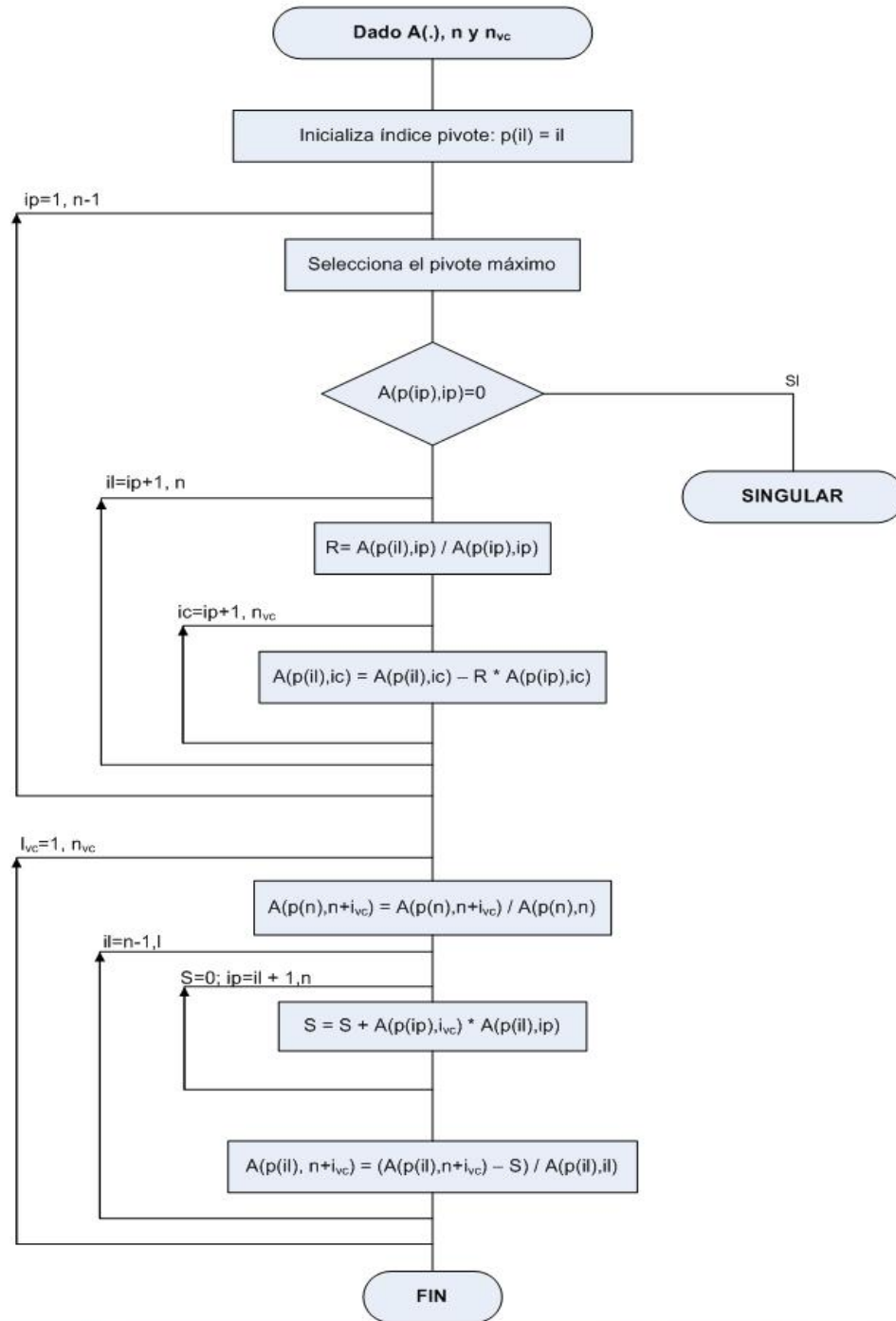


Figura 2-2 Diagrama de método de Gauss.



### 2.3 Eliminación de Gauss-Jordan.

El método de Gauss-Jordan, también conocido como el método de diagonalización, consiste en una modificación al método original (*método de Gauss*). Esta diferencia radica en que los únicos elementos no nulos que se presentarán en la matriz  $\mathbf{A}$  al final de la ejecución, serán los que se encuentren en la diagonal principal.

El proceso para su resolución, es el mismo utilizado en el método de Gauss. Para ello se aplica como primer paso:

$$a_{ij}^{(1)} = a_{ij} - a_{1j} \frac{a_{i1}}{a_{11}} \quad (i \neq 1, 1 \leq j \leq n+1)$$

$$b_i = b_i - b_1 \frac{a_{i1}}{a_{11}}$$

Obteniendo:

$$\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(1)} & \dots & a_{nn}^{(1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(1)} \end{bmatrix}$$

Y luego se procede a hacer nulo el resto de los valores, sin incluir los valores de la diagonal principal. Para lo cual se aplica:

$$a_{ij}^{(k-1)} = a_{ij}^{(k-1)} - a_{kj}^{(k-1)} \frac{a_{ik}^{(k-1)}}{a_{kk}^{(k-1)}}$$

$$b_i^{(k)} = b_i^{(k-1)} - b_k^{(k-1)} \frac{a_{i1}^{(k-1)}}{a_{kk}^{(k-1)}}$$

Realizando este paso tantas veces sea necesario, se obtiene una matriz que sólo tendrá valores no nulos en su diagonal principal. Como se puede observar:

$$\begin{bmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22}^{(1)} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_{nn}^{(n-1)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2^{(1)} \\ \vdots \\ b_n^{(n-1)} \end{bmatrix}$$

Donde la solución del sistema es obtenida de una forma rápida.

## Solución sistema lineal: método de Gauss-Jordan

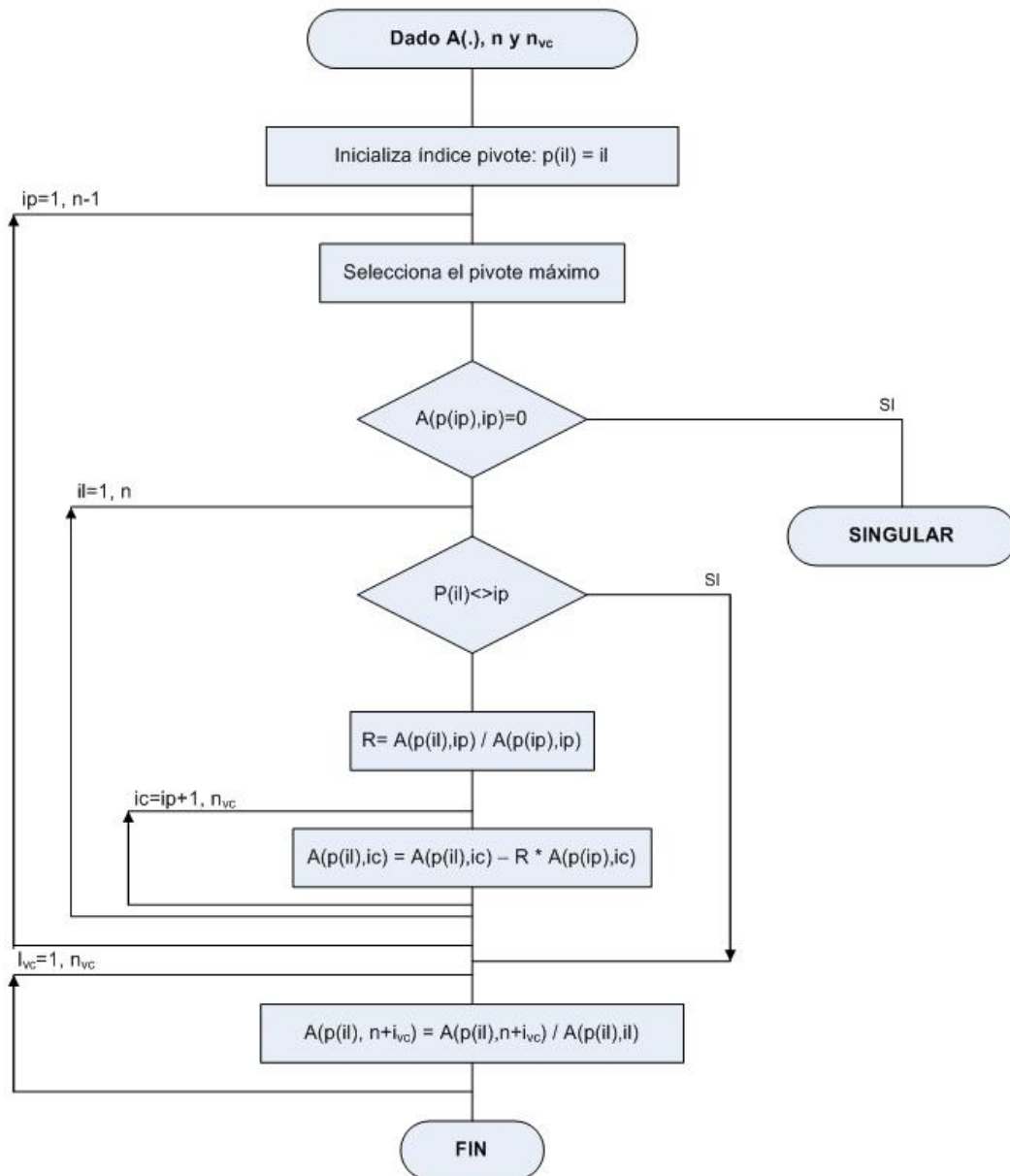


Figura 2-3 Diagrama del método de Gauss-Jordan.

### 3. Métodos iterativos de solución de sistemas de ecuaciones.

En ingeniería se puede llegar a problemas muy grandes, donde el número de términos no nulos dentro del perfil de la matriz de coeficientes es pequeño comparado con el número de elementos nulos. En consecuencia, para este tipo de problemas los métodos iterativos de solución de sistemas de ecuaciones son más eficientes que los métodos directos [43]. La ventaja de la iteración es la reducción del almacenamiento en la memoria central del ordenador y la eliminación de la descomposición triangular que es el proceso con mayor dificultad que presentan los métodos directos, pero poseen desventajas, como el desconocimiento del número de iteraciones que son necesarias para obtener una solución aceptable y este método a veces falla cuando es aplicado a ecuaciones no simétricas o no es definida positiva.

En la actualidad, el Método del Gradiente Conjugado (GC), es un método iterativo ampliamente difundido para resolver sistemas grandes de ecuaciones lineales. El GC es efectivo para resolver sistemas de la forma:

$$Ax = b \quad (3.1)$$

donde  $x$  es un vector incógnita,  $b$  es un vector conocido, y  $A$  es una matriz conocida, cuadrada, simétrica y definida positiva. Una matriz definida positiva es aquella que para cualquier vector no nulo  $x$  se cumple que:

$$x^T Ax > 0 \quad (3.2)$$

Los métodos iterativos como el GC son adecuados para utilizarse con matrices dispersas. Si  $A$  es densa, la mejor forma de resolver el sistema es factorizando  $A$  y

aplicando sustitución hacia atrás. El tiempo consumido factorizando una matriz  $A$  densa es aproximadamente equivalente al tiempo consumido resolviendo el sistema por métodos iterativos; y una vez que  $A$  está factorizada, el sistema se puede resolver rápidamente por sustitución hacia atrás para múltiples valores de  $b$ .

Comparando una matriz densa con una matriz dispersa (sparse) de mayor tamaño que ocupa la misma cantidad de memoria. Los factores triangulares de una matriz  $A$  dispersa generalmente tienen mayor cantidad de elementos no nulos que  $A$  por sí misma. Factorizar puede ser imposible debido a que la memoria es limitada, y puede además consumir demasiado tiempo; hasta el proceso de sustitución hacia atrás puede consumir más tiempo que la solución iterativa. Por otra parte, la mayoría de los métodos iterativos son eficientes en el manejo de la memoria y corren rápidamente con matrices dispersas [27].

### ***3.1 Formas cuadráticas.***

La introducción de las formas cuadráticas se debe a que el campo de desplazamiento desconocido fue interpolado por funciones de forma lineal dentro de cada elemento. Sin embargo, en algunos problemas, el uso de la interpolación cuadrática nos conduce a resultados más exactos que los obtenidos por las funciones de forma lineal.

Una *forma cuadrática* es simplemente una función cuadrática escalar de un vector que tiene la siguiente forma:

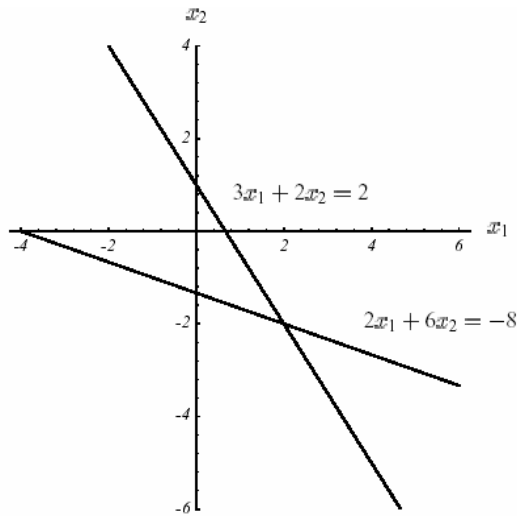
$$f(x) = \frac{1}{2} x^T A x - b^T x + c \quad (3.3)$$

donde  $A$  es una matriz,  $x$  y  $b$  son vectores, y  $c$  es un escalar constante. Si  $A$  es simétrica y definida positiva, la solución del sistema  $Ax=b$  corresponde a un mínimo de  $f(x)$ .

A manera de ejemplo tomemos los siguientes valores:

$$A = \begin{bmatrix} 3 & 2 \\ 2 & 6 \end{bmatrix} \quad b = \begin{bmatrix} 2 \\ -8 \end{bmatrix} \quad c = 0 \quad (3.4)$$

El sistema  $Ax=b$  se ilustra a continuación:



**Figura 3-1**<sup>3</sup> Sistema  $Ax=b$ .

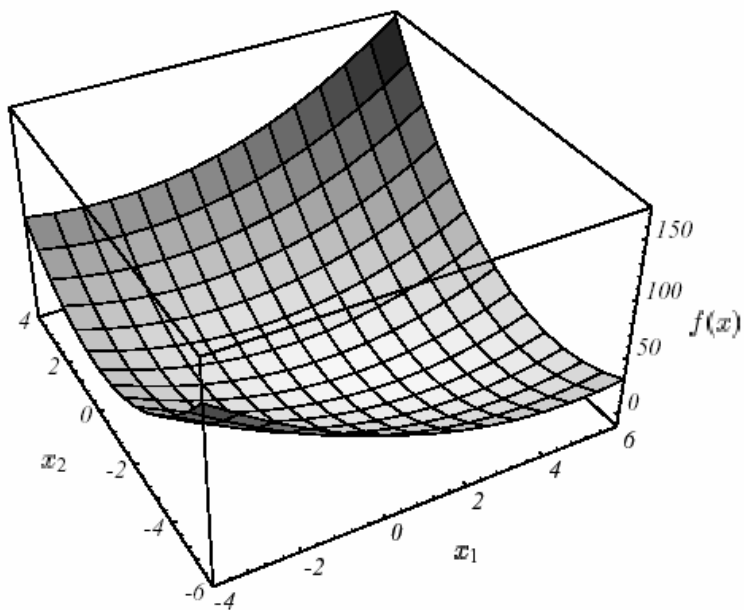
En forma general, la solución  $x$  se encuentra en el punto de intersección de los  $n$  hiperplanos existentes, donde cada hiperplano tiene una dimensión  $n-1$ . Para este

caso la solución es  $x = \begin{bmatrix} 2 \\ -2 \end{bmatrix}$ .

<sup>3</sup> Tomado de [37]

Su forma cuadrática correspondiente se muestra en la Figura 3-2. El punto más bajo de la gráfica es la solución del sistema  $Ax=b$ . La gráfica tiene la forma de un paraboloides debido a que  $A$  definida positiva.

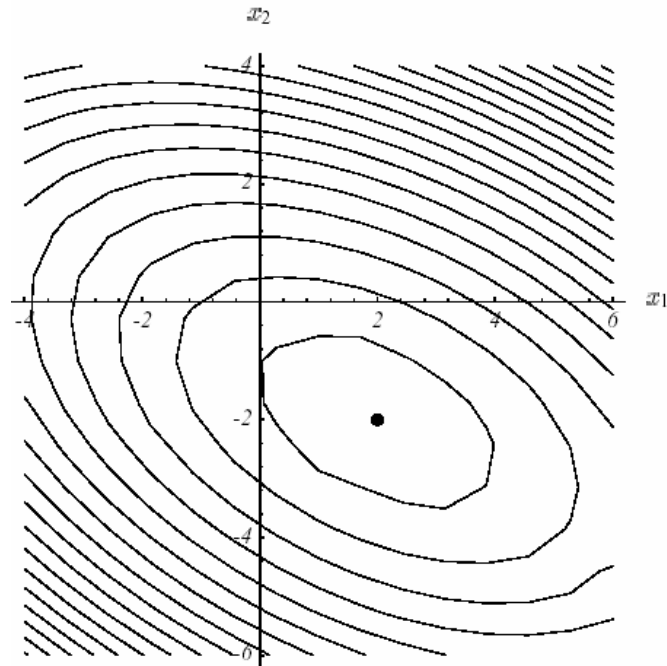
En la Figura 3-3 se grafican las formas cuadráticas pero en forma de curvas de nivel. Se puede observar el punto de solución, el cual corresponde al punto más bajo. Cada curva elipsoidal corresponde a un  $f(x)$  constante.



**Figura 3-2<sup>4</sup>** Formas cuadráticas del sistema  $Ax=b$ .

---

<sup>4</sup> Tomado de [37]



**Figura 3-3**<sup>5</sup> Formas cuadráticas en forma de curvas de nivel.

El gradiente es un campo vectorial que, para un punto dado  $x$ , apunta en la dirección del máximo incremento de  $f(x)$ .

El gradiente de una forma cuadrática se define como:

$$f'(x) = \begin{bmatrix} \frac{\delta}{\delta x_1} f(x) \\ \frac{\delta}{\delta x_2} f(x) \\ \vdots \\ \frac{\delta}{\delta x_n} f(x) \end{bmatrix} \quad (3.5)$$

<sup>5</sup> Tomado de [37]



si aplicamos la ecuación (3.5) a la ecuación (3.3) resulta en:

$$f'(x) = \frac{1}{2} A^T x + \frac{1}{2} Ax - b \quad (3.6)$$

Si  $A$  es simétrica, la ecuación anterior se reduce a:

$$f'(x) = Ax - b \quad (3.7)$$

Si igualamos el gradiente a cero obtenemos la ecuación (3.1), que es el sistema lineal a resolver. En consecuencia la solución de  $Ax = b$  es el punto crítico de  $f(x)$ . Si  $A$  es definida positiva y simétrica, entonces esta solución es el mínimo de  $f(x)$ . En consecuencia se puede resolver  $Ax = b$  encontrando un valor de  $x$  que minimice a  $f(x)$ .

### ***3.2 El método del descenso más rápido.***

En éste método se empieza desde un punto arbitrario  $x_{(0)}$  y nos deslizamos hasta el fondo del paraboloide. Hacemos una serie de pasos  $x_{(1)}, x_{(2)}, \dots$  hasta que nos encontremos suficientemente cerca de la solución  $x$ .

Cuando hacemos uno de los pasos, escogemos la dirección en la que la función desciende más rápidamente, la cual es la dirección contraria a  $f'(x_{(i)})$ . Según la ecuación (3.7), esta dirección es  $-f'(x_{(i)}) = b - Ax_{(i)}$ .

A continuación vamos a introducir algunas definiciones.

El error  $e_{(i)} = x_{(i)} - x$  es un vector que indica cuán lejos nos encontramos de la solución  $x$ . El residuo  $r_{(i)} = b - Ax_{(i)}$  indica cuán lejos nos encontramos del valor correcto de  $b$ . Es fácil darse cuenta de que  $r_{(i)} = -Ae_{(i)}$ . Se puede considerar el residual como el error que está siendo transformado por  $A$  hacia el mismo espacio como  $b$ . Más importante aún, se puede hacer  $r_{(i)} = -f'(x_{(i)})$  y considerar al residual como la dirección de la pendiente más pronunciada.

Como dijimos anteriormente empezábamos en un punto arbitrario  $x_{(0)}$  y hacíamos una serie de pasos hacia la solución. El paso lo definimos como:

$$x_{(1)} = x_{(0)} + \alpha r_{(0)} \quad (3.8)$$

Una *búsqueda de línea* es un procedimiento para elegir a un  $\alpha$  para minimizar  $f$  a lo largo de una línea. Esto se ilustra en la siguiente figura.

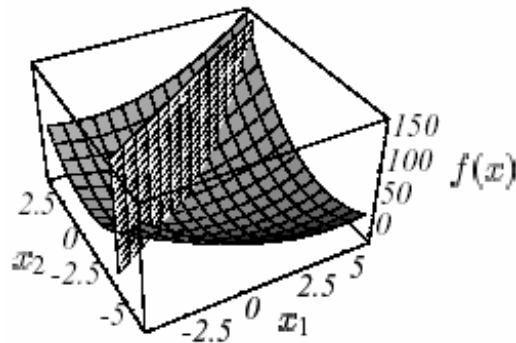


Figura 3-4<sup>6</sup> Búsqueda de línea.

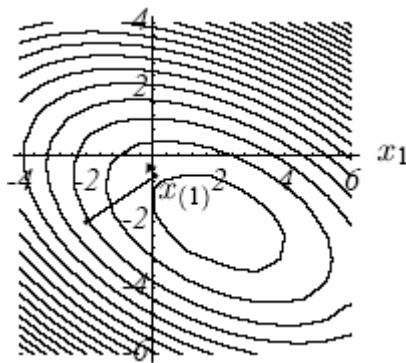
<sup>6</sup> Tomado de [37]

Por cálculo básico, sabemos que  $\alpha$  minimiza a  $f$  cuando la derivada direccional

$\frac{d}{d\alpha} f(x_{(1)})$  es igual a cero. Según la regla de la cadena,

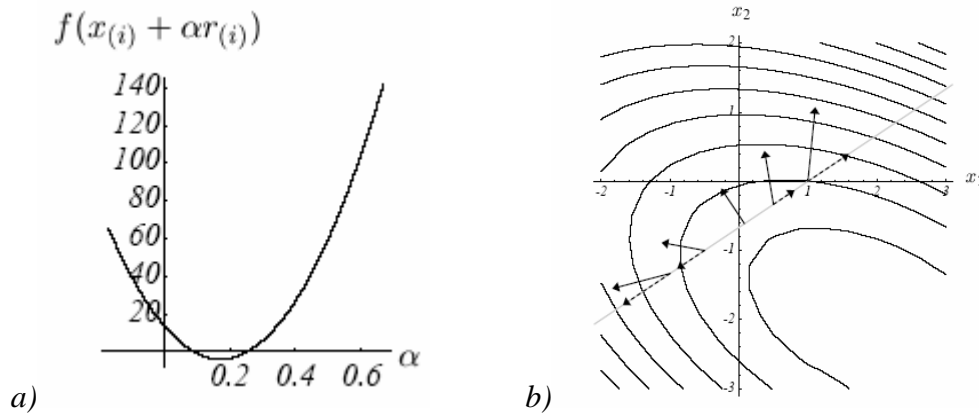
$$\frac{d}{d\alpha} f(x_{(1)}) = f'(x_{(1)})^T \frac{d}{d\alpha} x_{(1)} = f'(x_{(1)})^T r_{(0)}$$

Igualando esto a cero, nos damos cuenta de que  $\alpha$  se escoge de manera de que  $r_{(0)}$  y  $f'(x_{(1)})$  sean ortogonales.



**Figura 3-5<sup>7</sup>** El gradiente en el punto más bajo es ortogonal al gradiente del paso anterior.

La pendiente de la parábola en cualquier punto es igual a la magnitud de la proyección del gradiente en la línea. Esto se puede observar en la Figura 3-6.



**Figura 3-6<sup>8</sup>** a) Esta parábola es la intersección de las superficies. b) Pendiente de la parábola y magnitud de la proyección del gradiente en la línea.

<sup>7</sup> Tomado de [37]

Estas proyecciones representan la tasa de incremento de la función  $f$  a través de la línea de búsqueda.  $f$  es minimizada donde la proyección es cero, donde el gradiente es ortogonal a la línea de búsqueda.

Para determinar  $\alpha$ , hacemos  $f'(x_1) = -r_{(1)}$  y forzamos los vectores a ser ortogonales:

$$\begin{aligned}
 r_{(1)}^T r_{(0)} &= 0 \\
 (b - Ax_{(1)})^T r_{(0)} &= 0 \\
 (b - A(x_{(0)} + \alpha r_{(0)}))^T r_{(0)} &= 0 \\
 (b - Ax_{(0)})^T r_{(0)} - \alpha (Ar_{(0)})^T r_{(0)} &= 0 \\
 (b - Ax_{(1)})^T r_{(0)} &= \alpha (Ar_{(0)})^T r_{(0)} \\
 r_{(0)}^T r_{(0)} &= \alpha r_{(0)}^T (Ar_{(0)}) \\
 \alpha &= \frac{r_{(0)}^T r_{(0)}}{r_{(0)}^T Ar_{(0)}} \tag{3.9}
 \end{aligned}$$

Uniendo todo lo anterior nos queda, como método del descenso más rápido, lo siguiente:

$$r_{(i)} = b - Ax_{(i)} \tag{3.10}$$

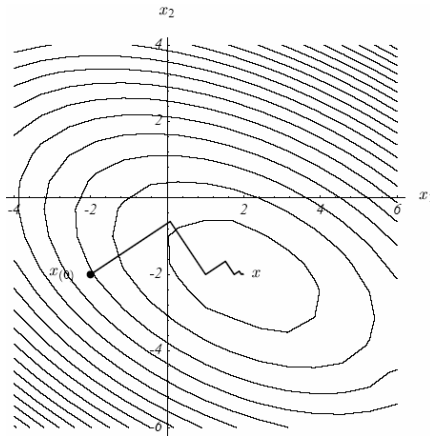
$$\alpha = \frac{r_{(i)}^T r_{(i)}}{r_{(i)}^T Ar_{(i)}} \tag{3.11}$$

---

<sup>8</sup> Tomado de [37]

$$x_{(i+1)} = x_{(i)} - \alpha_{(i)} r_{(i)} \quad (3.12)$$

Considerando el ejemplo anterior, utilizamos el método y nos converge para la solución mostrada en la Figura 3-7. Notemos el patrón en forma de Zig-Zag, que aparece debido a que cada gradiente es ortogonal al gradiente anterior.



**Figura 3-7<sup>9</sup>** Solución del ejemplo. Nótese el patrón en forma de Zig-Zag.

El algoritmo anterior requiere dos operaciones de multiplicación entre una matriz y un vector por cada iteración. El costo computacional del método de la pendiente más pronunciada está dominado por productos entre matrices y vectores. Una de estas multiplicaciones se puede eliminar multiplicando ambos lados de la ecuación (3.12) por  $-A$  e introduciendo  $b$ , nos queda:

$$r_{(i+1)} = r_{(i)} - \alpha_{(i)} A r_{(i)} \quad (3.13)$$

La ecuación (3.10) sigue siendo necesaria para calcular  $r_{(0)}$ , luego utilizaremos la ecuación (3.13) para cada iteración subsiguiente. El producto  $Ar$  sólo hace falta

<sup>9</sup> Tomado de [37]

calcularlo una vez. La desventaja de utilizar esta recurrencia es que la secuencia definida por la ecuación (3.13) es generada sin retroalimentación del valor de  $x_{(i)}$ , de forma que la acumulación del error de redondeo de punto flotante puede causar que  $x_{(i)}$  tenga una convergencia hacia algún punto cercano a  $x$ . Este efecto se puede anular utilizando periódicamente la ecuación (3.10) para recalcular el valor correcto del residual.

### 3.3 El método de las direcciones conjugadas.

El método del descenso más rápido a veces realiza pasos en la misma dirección que en pasos anteriores como se muestra en la Figura 3-7. Para evitar que esto suceda, y realizar así el menor número de pasos, tomamos un grupo de *direcciones de búsqueda* ortogonales  $d_{(0)}, d_{(1)}, \dots, d_{(n-1)}$ . En cada dirección, daremos solamente un paso, y ese paso va a ser la longitud justa para alinearse eventualmente con  $x$ . Luego de  $n$  pasos, tendremos el resultado. Esta idea se ilustra en la Figura 3-8.

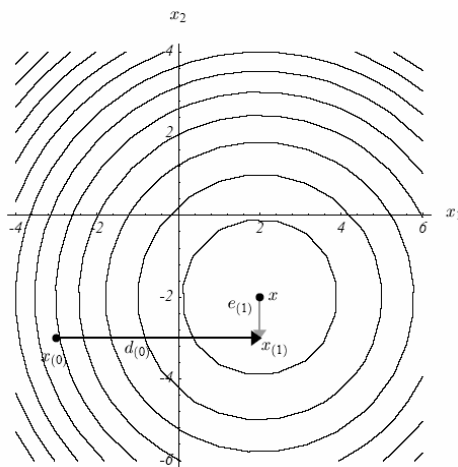


Figura 3-8<sup>10</sup> Direcciones de búsqueda ortogonales.

<sup>10</sup> Tomado de [37]

En este caso tomamos los ejes coordenados como las direcciones de búsqueda. El primer paso (horizontal) nos lleva a la coordenada correcta  $x_1$ . El segundo paso (vertical) nos lleva a la respuesta. Nótese que  $e_{(1)}$  es ortogonal a  $d_{(0)}$ . En general, para cada paso elegimos un punto:

$$x_{(i+1)} = x_{(i)} + \alpha_{(i)} d_{(i)} \quad (3.14)$$

Para hallar el valor de  $\alpha_{(i)}$ , obligamos  $e_{(i-1)}$  a ser ortogonal a  $d_{(i)}$ , de forma de no volver a caer en la dirección de  $d_{(i)}$ . Usando esta condición nos queda:

$$\begin{aligned} d_{(i)}^T e_{(i-1)} &= 0 \\ d_{(i)}^T (e_{(i)} + \alpha_{(i)} d_{(i)}) &= 0 \\ \alpha_{(i)} &= -\frac{d_{(i)}^T e_{(i)}}{d_{(i)}^T d_{(i)}} \end{aligned} \quad (3.15)$$

Desafortunadamente, no podemos calcular  $\alpha_{(i)}$  sin conocer  $e_{(i)}$ ; pero si conociéramos  $e_{(i)}$  ya el problema estaría resuelto.

La solución es hacer las direcciones de búsqueda conjugadas en vez de ortogonales. Dos vectores  $d_{(i)}$  y  $d_{(j)}$  son conjugados si:

$$d_{(i)}^T A d_{(j)} = 0 \quad (3.16)$$

Nuestros nuevos requerimientos son que  $e_{(i+1)}$  sea conjugado con  $d_{(i)}$ .

Esta condición de ortogonalidad es equivalente a encontrar el punto mínimo en la dirección de búsqueda  $d_{(i)}$ , como hacíamos en el método de la pendiente más pronunciada.

Para poder observar esto hagamos la derivada direccional igual a cero:

$$\begin{aligned}\frac{d}{d\alpha} f(x_{(i-1)}) &= 0 \\ f'(x_{(i+1)})^T \frac{d}{d\alpha} x_{(i+1)} &= 0 \\ -r_{(i-1)}^T d_{(i)} &= 0 \\ d_{(i)}^T A e_{(i-1)} &= 0\end{aligned}$$

Siguiendo la deducción de la ecuación (3.15), obtenemos la expresión para  $\alpha_{(i)}$  cuando las direcciones de búsqueda son conjugadas:

$$\alpha_{(i)} = -\frac{d_{(i)}^T A e_{(i)}}{d_{(i)}^T A d_{(i)}} \quad (3.17)$$

$$\alpha_{(i)} = \frac{d_{(i)}^T r_{(i)}}{d_{(i)}^T A d_{(i)}} \quad (3.18)$$

A diferencia de la expresión (3.15) la ecuación (3.18) sí se puede calcular. Nótese que si el vector de búsqueda es el residuo, esta fórmula va a ser idéntica a la fórmula usada en el método de la pendiente más pronunciada.

Para demostrar que este procedimiento en realidad calcula  $x$  en  $n$  pasos, expresaremos el error como una combinación lineal de direcciones de búsqueda, de forma que:



$$e_{(0)} = \sum_{j=0}^{n-1} \delta_j d_{(j)} \quad (3.19)$$

Los valores de  $\delta_j$  pueden ser encontrados utilizando el siguiente procedimiento. Como las direcciones de búsqueda son conjugadas, es posible eliminar todos los valores de  $\delta_j$  excepto uno. Esto se hace premultiplicando la expresión (3.19) por  $d_{(k)}^T A$ . De forma que:

$$\begin{aligned} d_{(k)}^T A e_{(0)} &= \sum_j \delta_j d_{(k)}^T A d_{(j)} \\ d_{(k)}^T A e_{(0)} &= \delta_{(k)} d_{(k)}^T A d_{(k)} \quad (\text{Por los } d \text{ vectores conjugados}) \\ \delta_{(k)} &= \frac{d_{(k)}^T A e_{(0)}}{d_{(k)}^T A d_{(k)}} \\ &= \frac{d_{(k)}^T A \left( e_{(0)} + \sum_{i=0}^{k-1} \alpha_{(i)} d_{(i)} \right)}{d_{(k)}^T A d_{(k)}} \quad (\text{Por los } d \text{ vectores conjugados}) \\ &= \frac{d_{(k)}^T A e_{(k)}}{d_{(k)}^T A d_{(k)}} \quad (\text{Por (3.14)}) \end{aligned} \quad (3.20)$$

Igualando las ecuaciones (3.17) y (3.20), encontramos que  $\alpha_{(i)} = -\delta_{(i)}$ . Esto nos da una nueva visión del error. Según las siguientes ecuaciones, el proceso de encontrar  $x$  componente por componente puede ser considerado como un proceso de eliminar el término del error componente por componente.

$$\begin{aligned}
e_{(i)} &= e_{(0)} + \sum_{j=0}^{i-1} \alpha_{(j)} d_{(j)} \\
&= \sum_{j=0}^{n-1} \delta_{(j)} d_{(j)} - \sum_{j=0}^{i-1} \delta_{(j)} d_{(j)} \\
&= \sum_{j=i}^{n-1} \delta_{(j)} d_{(j)}
\end{aligned} \tag{3.21}$$

Luego de  $n$  iteraciones, cada componente es eliminado y  $e_{(n)} = 0$  ; completando así la demostración.

El método de las Direcciones Conjugadas escoge el valor  $e_{(i)}$  de  $e_{(0)} + D_i$  que minimiza a  $\|e_{(i)}\|_A$  (figura 3.12). Donde  $D_i$  es el subespacio  $i$ -dimensional subespacio  $\{d_{(0)}, d_{(1)}, \dots, d_{(i-1)}\}$ , y  $\|e_{(i)}\|_A$  es la norma de energía. De la misma forma que el término del error se puede expresar como una combinación lineal de direcciones de búsqueda, su norma de energía se puede expresar como una sumatoria. La norma de energía se expresa como:

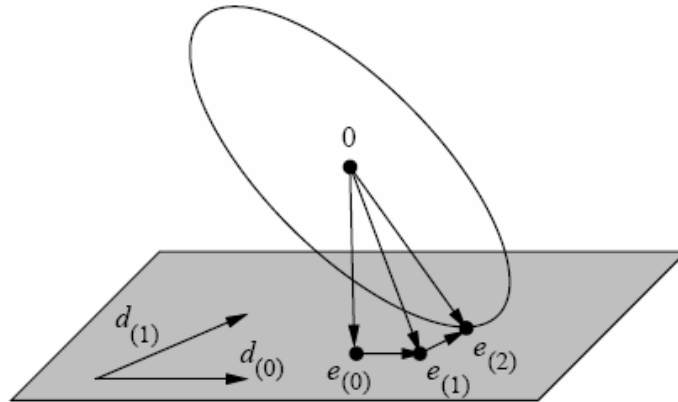
$$\|e_{(i)}\|_A = \sum_{j=i}^{n-1} \sum_{k=i}^{n-1} \delta_{(j)} \delta_{(k)} d_{(j)}^T A d_{(k)} \quad (\text{De la ec (3.21)}) \tag{3.22}$$

$$\|e_{(i)}\|_A = \sum_{j=i}^{n-1} \delta_{(j)}^2 d_{(j)}^T A d_{(j)} \quad (\text{Por los } d \text{ vectores conjugados}) \tag{3.23}$$

Cada término en ésta sumatoria está asociada con una dirección de búsqueda que no ha sido recorrida. Cualquier otro vector  $e$  escogido de  $e_{(0)} + D_i$  debe tener éstos

mismos términos en su expansión, lo cual demuestra que  $e_{(i)}$  debe tener la mínima norma de energía.

Otra propiedad importante del método de las Direcciones Conjugadas es que en cada paso, el hiperplano  $X_{(0)} + D_i$  es tangente al elipsoide donde se encuentra  $x_{(i)}$ .



**Figura 3-9<sup>11</sup>** Minimización de  $\|e\|_A$ .

En esta figura, el área sombreada es  $\epsilon_{(0)} + D_2 = \epsilon_{(0)} + \text{subespacio}\{d_{(0)}, d_{(1)}\}$ . El elipsoide es un contorno en el cual la norma de energía es constante. Después de dos pasos, el método de las direcciones conjugadas consigue  $e_{(2)}$ , en ese punto  $e_{(0)} + D_2$ , eso minimiza  $\|e\|_A$ .

Recordemos que el residuo en cualquier punto es ortogonal a la superficie elipsoidal en ese punto. En consecuencia  $r_{(i)}$  es ortogonal a  $D_i$  también. Mostremos esta propiedad matemáticamente, multiplicando la ecuación (3.21) por  $-d_{(i)}^T A$  nos queda:

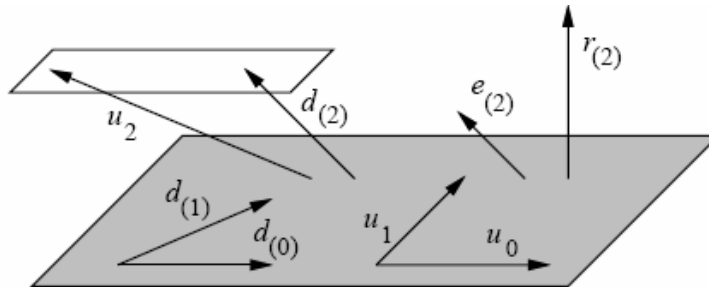
<sup>11</sup> Tomado de [37]

$$-d_{(i)}^T A e_{(j)} = -\sum_{j=i}^{n-1} \delta_{(j)} d_{(i)}^T A d_{(j)} \tag{3.24}$$

$$d_{(i)}^T r_{(j)} = 0, \quad i < j \quad (\text{Por los d-vectores conjugados}) \tag{3.25}$$

Como las direcciones de búsqueda son construidas a partir de  $u$  vectores, el subespacio constituido por  $u_0, \dots, u_{i-1}$  es  $D_i$ , y el residual  $r_{(i)}$  es ortogonal a estos vectores  $u$  también. Esto se ilustra en la Figura 3-10.

Para poder demostrar esto se deben generar un grupo de direcciones de búsqueda conjugados  $\{d_i\}$ . Supongamos que tenemos un grupo de  $n$  vectores linealmente independientes  $u_0, u_1, \dots, u_{n-1}$ .



**Figura 3-10**<sup>12</sup> Debido a que las direcciones de búsqueda  $d_{(0)}, d_{(1)}$  se construyen a partir de los vectores  $u_0, u_1$ , éstos crean el subespacio  $D_2$  (el plano coloreado en gris). El término del error  $e_{(2)}$  es conjugado con  $D_2$ , el residuo  $r_{(2)}$  es ortogonal a  $D_2$ , y una nueva dirección de búsqueda  $d_{(2)}$  es construida de  $u_2$  para ser conjugado con  $D_2$ . Los puntos finales de  $u_2$  y  $d_{(2)}$  se encuentran en un plano paralelo a  $D_2$ , debido a que  $d_{(2)}$  se construye a partir de  $u_2$  por la conjugación de Gram-Schmidt.

<sup>12</sup> Tomado de [37]

Para construir  $d_i$ , tomamos  $u_i$  y le restamos cualquier componente que no sea conjugado con los vectores previos  $d$  (Figura 3-11). Hacemos  $d_{(0)} = u_{(0)}$  y para  $i > 0$  nos queda:

$$d_{(i)} = u_i + \sum_{k=0}^{i-1} \beta_{ik} d_{(k)} \tag{3.26}$$

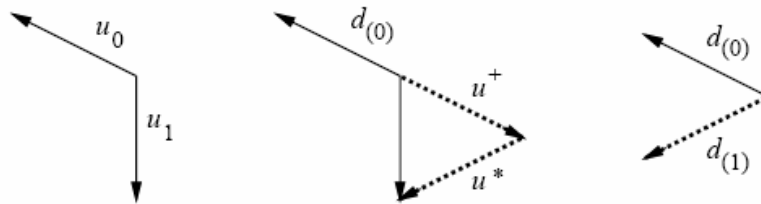
Donde  $\beta_{ik}$  está definido para  $i > k$ . Para hallar los valores hacemos lo siguiente:

$$d_{(i)}^T A d_{(i)} = u_i^T A d_{(j)} + \sum_{k=0}^{i-1} \beta_{ik} d_{(k)}^T A d_{(j)}$$

$$0 = u_i^T A d_{(j)} + \beta_{ij} d_{(j)}^T A d_{(j)}, \quad i > j \quad (\text{Por los } d \text{ vectores conjugados})$$

$$\beta_{ij} = -\frac{u_i^T A d_{(j)}}{d_{(j)}^T A d_{(j)}} \tag{3.27}$$

Este procedimiento se llama Conjugación de Gram-Schmidt.



**Figura 3-11**<sup>13</sup> Conjugación de Gram-Schmidt de dos vectores. Se empieza con dos vectores linealmente independientes  $u_0$  y  $u_1$ . Se hace  $d_{(0)} = u_0$ . El vector  $u_1$  posee dos componentes, que es conjugado con  $d_{(0)}$ , y  $u^+$ , que es paralelo a  $d_{(0)}$ .

<sup>13</sup> Tomado de [37]

Si a la ecuación (3.26) la multiplicamos por  $r_{(j)}$ :

$$d_{(i)}^T r_{(j)} = u_i^T r_{(j)} + \sum_{k=0}^{i-1} \beta_{ik} d_{(k)}^T r_{(j)} \quad (3.28)$$

Y por la ecuación (3.25) nos queda:

$$\therefore 0 = u_i^T r_{(j)}, \quad i < j \quad (3.29)$$

Completando así la demostración.

Hay otra propiedad que se genera a partir de la ecuación (3.28) y la Figura 3-10:

$$d_{(i)}^T r_{(i)} = u_i^T r_{(i)} \quad (3.30)$$

Igual que en el método del descenso más rápido, se puede utilizar una fórmula de recurrencia para hallar el residuo:

$$\begin{aligned} r_{(i+1)} &= -Ae_{(i+1)} \\ &= -A(e_{(i)} + \alpha_{(i)} d_{(i)}) \\ &= r_{(i)} - \alpha_{(i)} A d_{(i)} \end{aligned} \quad (3.31)$$

### ***3.4 El Método del Gradiente Conjugado.***

Este método fue introducido por Hestenes y Stiefel [15] en el año 1952, y tiene la ventaja teórica de que converge a la solución en  $N$  iteraciones. Siendo  $N$  el

número de ecuaciones e incógnitas. En la realidad, el redondeo de punto flotante no permite que esta propiedad se cumpla.

El método del Gradiente Conjugado se basa en el hecho de que su resolución es equivalente a minimizar el funcional cuadrático Ec. (3.3). Este sistema de resolución es análogo a la aproximación variacional del método de los elementos finitos.

### 3.4.1 Método del Gradiente Conjugado.

El método del Gradiente Conjugado es simplemente el método de las Direcciones Conjugadas, donde las *direcciones de búsqueda* son construidas conjugando los residuos, esto se logra haciendo  $u_i = r_{(i)}$ . Los residuos tienen la propiedad de que son ortogonales a las direcciones de búsqueda previas. En consecuencia siempre van a producir una dirección de búsqueda nueva y linealmente independiente, excepto si el residuo es cero, para el cual la solución ya fue encontrada. Como los vectores de búsqueda se construyen a partir de los residuos, el subespacio  $\{r_{(0)}, r_{(1)}, \dots, r_{(i-1)}\}$  es igual a  $D_i$ . Como cada residuo es ortogonal a las direcciones de búsqueda anteriores, también es ortogonal a los residuos anteriores (Figura 3-12). Entonces, la ecuación (3.29) se convierte en:

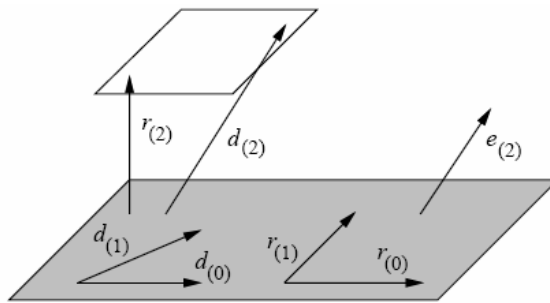
$$r_{(i)}^T r_{(j)} = 0, \quad i \neq j \quad (3.32)$$

La ecuación (3.31) nos muestra que cada residuo  $r_{(i)}$  es una combinación lineal de los residuos previos y  $Ad_{(i-1)}$ . Recordemos que  $d_{(i-1)} \in D_i$ . Este hecho implica que cada

subespacio nuevo  $D_{i+1}$  se forma a partir de la unión del subespacio previo  $D_i$  y el subespacio  $AD_i$ . En consecuencia:

$$D_i = \text{subespacio} \{d_{(0)}, Ad_{(0)}, A^2d_{(0)}, \dots, A^{i-1}d_{(0)}\}$$

$$= \text{subespacio} \{r_{(0)}, Ar_{(0)}, A^2r_{(0)}, \dots, A^{i-1}r_{(0)}\}$$



**Figura 3-12**<sup>14</sup> En el método del Gradiente Conjugado, cada nuevo residuo es ortogonal a todos los residuos anteriores y direcciones de búsqueda; y cada nueva dirección de búsqueda es construida para ser conjugado a todos los residuos y direcciones de búsqueda previos. Los puntos finales de  $r_{(2)}$  y  $d_{(2)}$  se encuentran en un plano paralelo a  $D_2$  (el espacio gris). En este método,  $d_{(2)}$  es una combinación lineal de  $r_{(2)}$  y  $d_{(1)}$ .

Este subespacio es llamado *subespacio de Krylov*, que es un subespacio creado aplicando repetidamente una matriz a un vector. Tiene una propiedad muy particular: como  $AD_i$  está incluido en  $D_{i+1}$ , el hecho de que el próximo residual  $r_{(i+1)}$  es ortogonal a  $D_{i-1}$  (ecuación (3.25)) implica que  $r_{(i+1)}$  está conjugado a  $D_i$ . La conjugación de Gram-Schmidt se realiza muy fácilmente debido a que  $r_{i+1}$  ya se encuentra conjugado con todas las direcciones de búsqueda previas exceptuando  $d_{(i)}$ .

<sup>14</sup> Tomado de [37]



Recordemos que de la ecuación (3.27) las constantes de Gram-Schmidt son

$$\beta_{ij} = -\frac{u_i^T Ad_{(j)}}{d_{(j)}^T Ad_{(j)}}, \text{ simplifiquemos esta expresión. Tomando el producto interno de } r_{(i)}$$

y la ecuación (3.31),

$$\begin{aligned} r_{(i)}^T r_{(j+1)} &= r_{(i)}^T r_{(j)} - \alpha_{(j)} r_{(i)}^T Ad_{(j)} \\ \alpha_{(j)} r_{(i)}^T Ad_{(j)} &= r_{(i)}^T r_{(j)} - r_{(i)}^T r_{(j-1)} \\ r_{(i)}^T Ad_{(j)} &= \begin{cases} \frac{1}{\alpha_{(i)}} r_{(i)}^T r_{(i)} & i = j \\ -\frac{1}{\alpha_{(i-1)}} r_{(i)}^T r_{(i)} & i = j+1 \\ 0 & \text{En cualquier otro caso} \end{cases} \\ \therefore \beta_{ij} &= \begin{cases} \frac{1}{\alpha_{(i-1)}} \frac{r_{(i)}^T r_{(i)}}{d_{(i-1)}^T Ad_{(i-1)}} & i = j+1 \\ 0 & i > j+1 \end{cases} \end{aligned}$$

Nótese que la mayoría de los términos de  $\beta_{ij}$  desaparecieron. No sigue siendo necesario guardar los anteriores vectores de búsqueda para asegurar que los nuevos vectores de búsqueda sean conjugados. En esto radica la importancia del algoritmo del GC, debido a que la complejidad del espacio y la complejidad del tiempo por cada iteración son reducidas de  $O(n^2)$  a  $O(m)$ , donde  $m$  es el número de términos no nulos de  $A$  y  $n \times n$  es el tamaño de la matriz. En adelante utilizaremos la abreviación  $\beta_{(i)} = \beta_{i,i-1}$ . Simplificando aún más:

$$\begin{aligned}\beta_{(i)} &= \frac{r_{(i)}^T r_{(i)}}{d_{(i-1)}^T r_{(i-1)}} \\ &= \frac{r_{(i)}^T r_{(i)}}{r_{(i-1)}^T r_{(i-1)}}\end{aligned}$$

Agrupando todo, tenemos el Método de Gradiente Conjugado:

$$d_{(0)} = r_{(0)} = b - Ax_{(0)} \quad (3.33)$$

$$\alpha_{(i)} = \frac{r_{(i)}^T r_{(i)}}{d_{(i)}^T A d_{(i)}} \quad (3.34)$$

$$x_{(i+1)} = x_{(i)} + \alpha_{(i)} d_{(i)} \quad (3.35)$$

$$r_{(i+1)} = r_{(i)} - \alpha_{(i)} A d_{(i)} \quad (3.36)$$

$$\beta_{(i-1)} = \frac{r_{(i+1)}^T r_{(i+1)}}{r_{(i)}^T r_{(i)}} \quad (3.37)$$

$$d_{(i+1)} = r_{(i+1)} + \beta_{(i-1)} d_{(i)} \quad (3.38)$$

### 3.5 Precondicionamiento.

En la práctica, el Método del Gradiente Conjugado se aplica de una manera más eficiente a un sistema lineal precondicionado:

$$R^{-1} A (R^{-1})^T (R^T x) = (R^{-1} b) \quad (3.39)$$

Donde  $R$  es una matriz no singular. El objetivo del preconditionamiento es que cuando se escoge  $R$  de manera apropiada, el algoritmo del Gradiente Conjugado presentará una convergencia mucho más rápida para la matriz  $R^{-1}A(R^{-1})^T$  que para  $A$ . Debido a que  $R$  no es singular,  $M = RR^T$  es una matriz definida positiva, y la ecuación (3.39) se transforma en:

$$M^{-1}Ax = M^{-1}b \quad (3.40)$$

El preconditionamiento es una técnica para mejorar el *número de condición*<sup>15</sup> de una matriz. Supongamos que  $M$  es simétrica y definida positiva y que aproxima  $A$ , pero que es más fácil de invertir. Podemos resolver  $Ax = b$  indirectamente resolviendo la ecuación (3.40) si  $k(M^{-1}A) \ll k(A)$ , o si los autovalores de  $M^{-1}A$  están mejor agrupados que los de  $A$ , podemos resolver la ecuación (3.40) iterativamente más rápido que el problema original. El artificio es que  $M^{-1}A$  generalmente no es simétrica y definida positiva, aún si  $M$  y  $A$  lo son.

La efectividad del preconditionador  $M$ , se determina por medio del número de condición de  $M^{-1}A$ . El objetivo es hacer que el número de condición esté lo más cerca posible de 1.

El preconditionamiento es un intento de estirar la forma cuadrática para hacerla más esférica, de forma que los autovalores estén más cerca los unos de los otros.

---

<sup>15</sup> Refiérase al anexo A.

### 3.6 Algoritmos en pseudo-código.

- El método del descenso más rápido.

```

 $i \leftarrow 0$ 
 $r \leftarrow b - Ax$ 
 $\delta \leftarrow r^T r$ 
 $\delta_0 \leftarrow \delta$ 
Mientras  $i < i_{\max}$  y  $\delta > \varepsilon^2 \delta_0$  hacer
     $q \leftarrow Ar$ 
     $\alpha \leftarrow \frac{\delta}{r^T q}$ 
     $x \leftarrow x + \alpha r$ 
    Si  $i$  es divisible por 50
         $r \leftarrow b - Ax$ 
    De otra manera
         $r \leftarrow r - \alpha q$ 
     $\delta \leftarrow r^T r$ 
     $i \leftarrow i + 1$ 

```

Dados los términos  $A$ ,  $b$ , un valor inicial  $x$ , un número máximo de iteraciones  $i_{\max}$ , y una tolerancia de error  $\varepsilon < 1$ . Este algoritmo termina cuando el número máximo de iteraciones  $i_{\max}$  ha sido excedido, o cuando  $\|r_{(i)}\| \leq \varepsilon \|r_{(0)}\|$ .

- El método del Gradiente Conjugado.

```

i ← 0
r ← b − Ax
d ← r
 $\delta_{new} \leftarrow r^T r$ 
 $\delta_0 \leftarrow \delta_{nuevo}$ 
Mientras i < imax y  $\delta_{nuevo} > \varepsilon^2 \delta_0$  Hacer
    q ← Ad
     $\alpha \leftarrow \frac{\delta_{nuevo}}{d^T q}$ 
    x ← x +  $\alpha d$ 
    Si i es divisible por 50
        r ← b − Ax
    De otra manera
        r ← r −  $\alpha q$ 
     $\delta_{viejo} \leftarrow \delta_{nuevo}$ 
     $\delta_{nuevo} \leftarrow r^T r$ 
     $\beta \leftarrow \frac{\delta_{nuevo}}{\delta_{viejo}}$ 
    d ← r +  $\beta d$ 
    i ← i + 1
  
```

Dados los términos  $A$ ,  $b$ , un valor inicial  $x$ , un número máximo de iteraciones  $i_{\max}$ , y una tolerancia de error  $\varepsilon < 1$ .

- El método del Gradiente Conjugado Precondicionado.

```

i ← 0
r ← b − Ax
d ← M−1r
δnuevo ← rTd
δ0 ← δnuevo
Mientras i < imax y δnuevo > ε2δ0 Hacer
    q ← Ad
    α ←  $\frac{\delta_{nuevo}}{d^T q}$ 
    x ← x + αd
    Si i es divisible por 50
        r ← b − Ax
    De otra manera
        r ← r − αq
    s ← M−1r
    δviejo ← δnuevo
    δnuevo ← rTs
    β ←  $\frac{\delta_{nuevo}}{\delta_{viejo}}$ 
    d ← s + βd
    i ← i + 1
  
```

Dados los términos  $A$ ,  $b$ , un valor inicial  $x$ , un preconditionador  $M$  (tal vez implícitamente definido), un número máximo de iteraciones  $i_{\max}$ , y una tolerancia de error  $\varepsilon < 1$ . La declaración  $s \leftarrow M^{-1}r$  implica que se debe aplicar el preconditionador, el cual puede no estar en forma de matriz.

## 4. El Precondicionador de Elemento a Elemento.

El concepto que es utilizado en una estrategia de precondicionamiento con almacenamiento de Elemento a Elemento y una técnica de solución iterativa, es que se almacena solamente la información que es única, o no redundante, tanto para la discretización de elementos finitos como para el precondicionador de Elemento a Elemento. El objetivo de esta estrategia de almacenamiento elemental es el de minimizar la información de la matriz que se debe almacenar. Así se permite el procesamiento de problemas de mayor escala, sin degradar el rendimiento computacional del precondicionador de Elemento a Elemento o la técnica de solución iterativa.

Consideramos la solución del sistema  $n$  por  $n$  de ecuaciones lineales,  $Ax = b$ , donde la matriz  $A$  es dispersa y estructurada de forma que se puede expresar como:

$$A = \sum_{i=1}^p A_i \quad (4.1)$$

Las matrices de los elementos de  $A_i$  son de bajo rango<sup>16</sup> cuando se expresa en el sistema de coordenadas global, y son usualmente dispersas, de manera que sus términos no nulos se representan como un bloque pequeño y denso, de la siguiente forma:

---

<sup>16</sup> Rango: número de vectores linealmente independientes que constituyen la matriz.

$$A_i = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & k_i & -k_i & 0 & 0 \\ 0 & -k_i & k_i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Una función  $f$  se puede decir que es parcialmente separable si  $f(x) = \sum_{i=1}^p f_{i(x)}$ , donde cada función  $f_i$  tiene un subespacio grande e invariante.

Asumimos que  $A$  es una matriz grande, simétrica y definida positiva. La técnica de solución considerada es el Método del Gradiente Conjugado utilizando el preconditionador de Elemento a Elemento, el cual fué introducido por [16] y [29].

Este preconditionador tiene unas características importantes. Los cálculos se pueden realizar centrándose en el elemento y la mayoría no necesita de ensamblaje [9]. El objetivo del preconditionador es aproximar  $A$  a un bajo costo computacional.

Como estamos asumiendo que la matriz de elemento  $A_i$  tiene elementos no nulos en solamente  $n_i$  filas, podemos expresarla como:

$$A_i = C_i^T A_i^\varepsilon C_i \quad (4.2)$$

Donde las filas de la matriz  $n_i$  por  $n$  de conectividad  $C_i$ , son simplemente las filas de la matriz identidad  $n$  por  $n$  correspondientes a las variables usadas en el elemento, y  $A_i^\varepsilon$  es una matriz  $n_i$  por  $n_i$  simétrica y densa.



Esto es:

$$A_i^\varepsilon = \begin{bmatrix} k_i & -k_i \\ -k_i & k_i \end{bmatrix} \quad A_i = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & k_i & -k_i & 0 & 0 \\ 0 & -k_i & k_i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A_i = C_i^T A_i^\varepsilon C_i$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & k_i & -k_i & 0 & 0 \\ 0 & -k_i & k_i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} k_i & -k_i \\ -k_i & k_i \end{bmatrix} \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

Por lo tanto la matriz de conectividad es:

$$C_i = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

#### ***4.1 Precondicionador de Elemento a Elemento.***

Asumiendo que  $A$  es definida positiva, podemos expresar  $A$  como:

$$A = \sum_{i=1}^p M_i + \sum_{i=1}^p (A_i - M_i) = M + \sum_{i=1}^p (A_i - M_i)$$

Donde:

$$M_i = \text{diag}(A_i) \text{ y } M = \sum_{i=1}^p M_i.$$

Ahora, hagamos  $M = L_M L_M^T$  mediante la *factorización de Cholesky*<sup>17</sup>. Nos queda  $L_m$  como una matriz diagonal. Entonces:

$$A = L_M \left( I + \sum_{i=1}^p L_M^{-1} (A_i - M_i) L_M^{-T} \right) L_M^T = L_M \left( I + \sum_{i=1}^p E_i \right) L_M^T \quad (4.3)$$

Donde  $E_i = L_M^{-1} (A_i - M_i) L_M^{-T}$ . Utilizando la aproximación  $I + \sum_{i=1}^p E_i \approx \prod_{i=1}^n (I + E_i)$ ,

obtenemos:

$$A \approx L_M \prod_{i=1}^p (I + E_i) L_M^T \quad (4.4)$$

Finalmente ensamblamos el Precondicionador de Elemento a Elemento:

$$P_{EBE} = L_M \left( \prod_{i=1}^p L_i \right) \left( \prod_{i=1}^p D_i \right) \left( \prod_{i=p}^1 L_i^T \right) L_M^T \quad (4.5)$$

Donde los factores  $L_i$  y  $D_i$  vienen de la factorización  $LDL^T$  de las matrices  $I + E_i$ .

La eficiencia de este preconditionador depende de la partición de la matriz inicial y de la magnitud de los elementos fuera de la diagonal en las matrices pertenecientes al elemento. La descomposición de  $A$  no es única, así que las diferentes descomposiciones pueden alterar significativamente el rendimiento del preconditionador.

<sup>17</sup> Refiérase al anexo 2.

En función de que queremos resolver eficientemente el sistema de ecuaciones  $P_{EBE}x = y$ , sacamos provecho de la ecuación (4.5). Podemos ordenar los elementos de cualquier forma queelijamos.

## 5. Implementación.

A continuación se describe la implementación del algoritmo. En 1<sup>er</sup> lugar se presentan las fases del mismo, seguida de su versión en lenguaje C.

### 5.1 Pseudo-código.

Para la solución del sistema lineal algebraico:

$$Ax = b$$

Se siguen los siguientes pasos:

$$p^0 = r^0 = b - Ax^0$$

Donde  $r^0$  es el residuo o error para el primer intento  $x^0$ . Se procede a realizar los siguientes  $k$  pasos del proceso:

$$\begin{aligned} u^k &= Ap^k \\ \alpha^k &= \frac{(r^k)^T r^k}{(p^k)^T u^k} \\ x^{k+1} &= x^k + \alpha^k p^k \\ r^{k+1} &= r^k - \alpha^k u^k \\ \beta^k &= \frac{(r^{k+1})^T r^{k+1}}{(r^k)^T r^k} \\ p^{k+1} &= r^{k+1} + \beta^k r^k \end{aligned}$$

Hasta que la diferencia entre  $x^{k+1}$  y  $x^k$  es suficientemente pequeño. Donde  $u$ ,  $p$  y  $r$  son vectores y  $\alpha$  y  $\beta$  son escalares.

## 5.2 Implementación en lenguaje C.

En el capítulo 3 se explicó de forma detallada los métodos iterativos de solución de sistemas de ecuaciones. Uno de estos métodos fue llevado a lenguaje C (Método del Gradiente Conjugado con Precondicionador de Elemento a Elemento) y se implementó al programa existente, el cual proviene del trabajo especial de grado [13].

Para desarrollo del Método del Gradiente Conjugado con Precondicionador de Elemento a Elemento en lenguaje C, se utilizó el pseudo-código anteriormente descrito, quedando de la siguiente forma:

```
#define TOLCG 1.0E-5
#define MAXITER 1000

int ChekConv( real *oldx, real *newx, ent num );

/*-----*/

void CGItera( real *b, real *Diago, real *s, real *gd,
             real *gp, real *xnew, real *pmul, real *gx,
             real *u, real *utemp, real *lm, ent nec,
             ent nevab, ent NumEle, ent igln, ent nne )
{
    int iters, converged;
    ent i;
    real up, down, alfa, beta, ome;

    cero( xnew, nec, CERO ); /* Llena el arreglo xnew con ceros */
    cero( gd , nec, CERO );
    cero( gp , nec, CERO );
    cero( gx , nec, CERO );
}
```

```

/*----- invierta el preconditionador -----*/
for ( i = 1; i <= nec; i++ )
{
  Diago(i) = 1./ Diago(i);
  gd(i)  = Diago(i) * B(i);
  gp(i)  = gd(i);
}

/*----- Ciclo iterativo principal-----*/
iters = converged = 0;
do
{
  iters++;  printf("  Iteration No %3d\n", iters );

  EnsamU( u, gp, pmul, utemp, s, lm, nec, NumEle, nne, igln, nevad );
  /* Ensamble vector "u" */

  /*----- solución del de ecuación -----*/

  up  = dotpro( b, gd, nec );
  down = dotpro( gp, u, nec );
  alfa = up / down;
  for ( i = 1; i <= nec; i++ )
  {
    xnew(i) = gx(i) + gp(i) * alfa;
    b(i)  = b(i) - u(i) * alfa;
    gd(i)  = Diago(i) * b(i);
  }
  ome = dotpro( b, gd, nec );
  beta = ome/up;
  vecaddscal( gp, gd, beta, nec );
  converged = ChekConv( gx, xnew, nec );
  veccopy( gx, xnew, nec );
}
while( (converged == 0) && (iters < MAXITER) );

veccopy( b, xnew, nec ); /*Almacena xnew en b*/

}

/*-----*/

int ChekConv( real *oldx, real *newx, ent num ) /*Verifica si hay convergencia*/

```

```

{
  ent i;
  int res = 1;
  real big = CERO;

  for ( i = 0; i < num; i++ )
    if ( fabs(newx[i]) > big ) big = fabs(newx[i]);

  for ( i = 0; i < num; i++ )
    if ( fabs( newx[i] - oldx[i] )/big > TOLCG ) res = 0;

  return( res );
}
/*-----*/

void EnsamU( real *u, real *gp, real *pmul, real *utemp, real *s, ent *lm,
            ent nec, ent NumEle, ent nne, ent igln, ent nevab )
{
  ent iel, i, j, k, indx;
  real de[6];

  cero( u, nec, 0.0 );
  for ( iel = 1; iel <= NumEle; iel++ )
  {
    MatrizE( iel, de );
    ElemE( iel, de, TRUE );
#ifdef LMIJ
    indx = 0;
    for ( i = 1; i <= nne; i++ )
    {
      for ( j = 1; j <= igln; j++ )
      {
        indx++;
        k = lm(i,j);
#ifdef TESTIDCERO
        if ( k != ICERO )
#endif
          pmul(indx) = gp(k);
      }
    }
#else
    for ( i = 1; i <= nevab; i++ )
    {

```

```

        k = lm(i)
        pmul(i) = gp(k);
    }
#endif

                                /* Calcula utemp = s * pmul */
    for ( i = 1; i <= nevab; i++ )
        utemp(i) = dotpro( pmul, &s(i,1), nevab );
#ifdef LMIJ
    indx = 0;
    for ( i = 1; i <= nne; i++ )
    {
        for ( j = 1; j <= igln; j++ )
        {
            indx++;
            k = lm(i,j);
#ifdef TESTIDCERO
            if ( k != ICERO )
#endif
            u(k) = u(k) + utemp(indx);
        }
    }
#else
    for ( i = 1; i <= nevab; i++ )
    {
        k = lm(i)
        u(k) = u(k) + utemp(i);
    }
#endif
}
}

/*****
/

```

Al finalizar el algoritmo, el arreglo “b” contiene a la solución del sistema.



## 6. Parte Experimental.

### 6.1 Factor de concentración de esfuerzos.

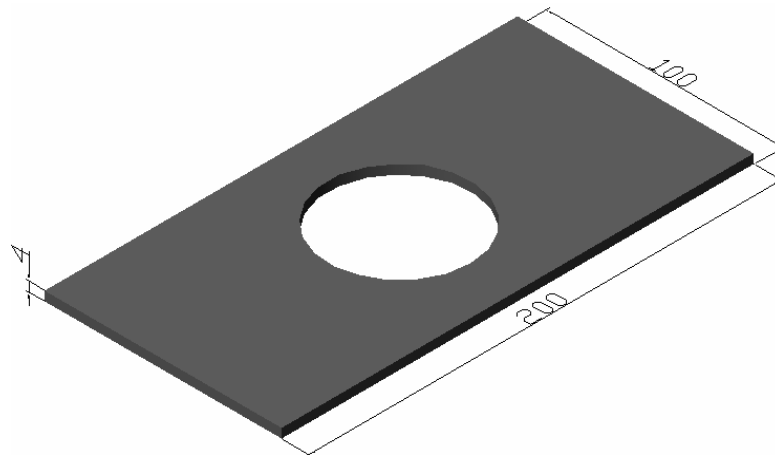
Se tiene una placa delgada finita con agujero circular sometida a esfuerzos de tracción. El objetivo es el de comparar el factor de concentración de esfuerzos teórico para este tipo de disposición experimental investigado y calculado por [30], con aquellos factores resultantes del análisis de elementos finitos. Aprovechando las condiciones de simetría se procede a dividir la probeta en ocho elementos. Esto se hace con la intención de realizar un análisis con un mayor número de elementos manteniendo el mismo costo computacional de realizar un análisis a la probeta completa. Este proceso lo llamamos idealización de la probeta. En la Figura 6-1 se muestra la disposición de la probeta que se va a analizar. Los datos son:

$$P = 9,5 * 10^8 \frac{\text{N}}{\text{mm}^2} * 4 \text{ mm} * 100 \text{ mm} = \underline{3,8 * 10^{11} \text{ N}}$$

$$W = 100 \text{ mm}$$

$$d = 60 \text{ mm}$$

$$h = 4 \text{ mm}$$



**Figura 6-1** Disposición de la probeta.

Y las fórmulas de los factores de concentración de esfuerzos:

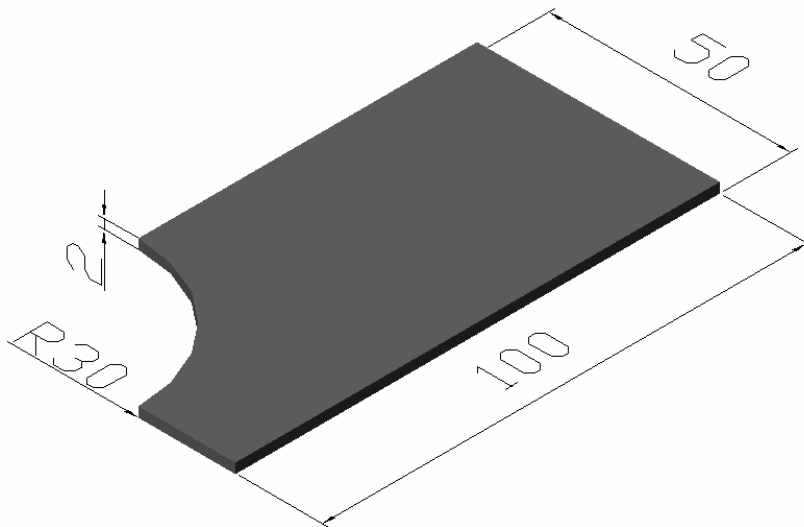
$$K = \frac{\sigma_{\max}}{\sigma_0}, \text{ Factor de concentración de esfuerzos.}$$

$$\sigma_0 = \frac{F}{A_0}, \text{ Esfuerzo nominal.}; A_0 = (w-d)t$$

Introduciendo los datos, nos queda:

$$\sigma_0 = \frac{3,8 \cdot 10^{11}}{(100 - 60) \cdot 4} = 2,375 \cdot 10^9 \text{ MPa}$$

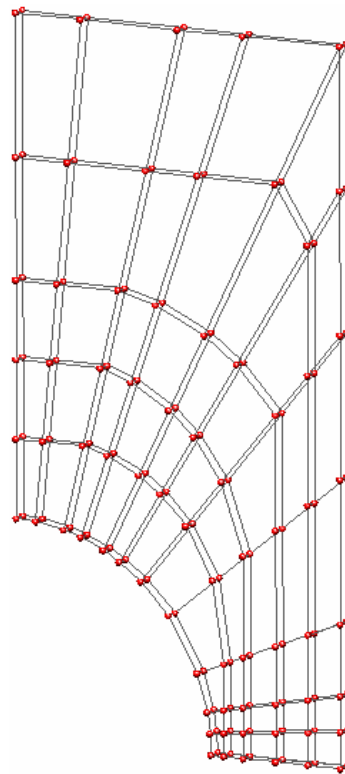
En la Figura 6-2 se muestra la probeta idealizada, que se va a simular.



**Figura 6-2** Probeta idealizada.

Debido a que el programa *MefetGC* carece de un módulo de pre-proceso, se decidió realizar un mallado de forma manual. Regiones tridimensionales pueden ser

efectivamente rellenos por elementos tetraédricos, pero la preparación manual de los datos se convierte en una tarea compleja. Para evitar esto, para regiones simples, se hace más fácil dividir las regiones en bloques hexaédricos de ocho nodos [5]. Debido a la complejidad que representa realizar un mallado manual de muchos elementos se decidió utilizar elementos hexaédricos de ocho nodos. El mallado realizado se presenta en la Figura 6-3. Este mallado es de 144 nodos y resulta en 55 elementos. Este mallado es el utilizado en el *MefetGC* y *Hexafron*, que son los programas que carecen de módulo de pre-proceso. Los programas *Mechanical Desktop* y *Visual Nastran 4D* generan su propia malla de elementos tetraédricos.



**Figura 6-3** Mallado realizado en forma manual de 144 nodos y 55 elementos.

## 6.2 Procedimiento de validación (Metodología).

La metodología fundamental consistió en resolver un problema cuya solución se ha obtenido por medios experimentales, y comparar los resultados obtenidos con el programa elaborado. Adicionalmente se realizaron comparaciones con otros programas:

- *Mechanical Desktop* (<http://www.autodesk.com>)
- *Visual Nastran 4D* (<http://www.msc.com>)
- *Hexafro* [5].

El proceso de validación constó de varios pasos generales, los cuales se expondrán detalladamente a continuación.

**Paso 1** (Elección de la probeta): Se realiza una escogencia de la probeta a usar, esto consta del material a utilizar, el módulo de elasticidad (E), el módulo de Poisson ( $\nu$ ), las dimensiones de la pieza en milímetros y la carga que se ha de utilizar en el problema.

**Paso 2** (Cálculo del esfuerzo nominal): Utilizando los datos del *paso 1* se calcula el esfuerzo nominal ( $\sigma_0$ ) que se presenta en la probeta.

**Paso 3** (Cálculo del concentrador de esfuerzo): Los datos del *paso 1* se introducen en el programa *Engineers Toolbox v2.0* obtenido de <http://www.engineerstoobox.com> dando el valor del concentrador de esfuerzo teórico  $K_t$ .

**Paso 4** (Idealización de la probeta): A la probeta se le determina los planos de simetría y si posee alguno, se idealiza la probeta eliminando las partes simétricas

repetidas y quedándonos con sólo un trozo de la probeta, con esto se pretende no saturar la memoria del computador, pudiendo procesar un mayor número de elementos al mismo costo computacional.

**Paso 5** (Mallado de la probeta idealizada en 3D): La probeta idealizada es llevada al programa *Autocad 2002*, en el cual se dibujó un mallado de elementos cuadriláteros en 3D realizado por nosotros, enumerando los nodos y los elementos para determinar cuantos hay en la probeta.

**Paso 6** (Coordenada de los nodos): Cada nodo se le busca su valor en (X,Y,Z) y son colocados en un archivo de texto.

**Paso 7** (Realización de un archivo de datos): Se coloca en la 1<sup>era</sup> línea una variable de entrada (1 ó 0). Si es un 1 (uno), cuando se ejecute el programa aparecerán los datos del modelo en la pantalla, en la impresora o en el disco según como se haga la selección de la unidad de salida de resultados. Si es un 0 (cero), aparecerán estos valores cuando se esté corriendo el programa. Esta variable sirve para verificar si los datos de entrada están correctos, en la 2<sup>da</sup> línea se identifica el nombre del archivo, en la 3<sup>era</sup> línea se le asigna un título al archivo, en la 4<sup>ta</sup> línea se identifica el número de dimensiones, número de grados de libertad por nodos y el número de propiedades donde para elasticidad lineal se le coloca 2 (dos) , para elasticidad no lineal el número 3 (tres) y para transferencia de calor el numero 5 (cinco), todos estos datos deben tener un espacio entre ellos. En la 5<sup>ta</sup> línea de forma similar que en la línea anterior se debe identificar los valores de los siguientes datos, el número de nodo, el número de elementos, el número de nodos cargados, el número de nodo restringidos, el número del material, el estado de orden de integración, la escala, el número de nodos por elementos, en la 6<sup>ta</sup> línea se le colocan las propiedades del los materiales, número del material, el módulo de elasticidad (E), el módulo de Poisson ( $\nu$ ), seguidamente se

colocan las coordenadas de los nodos (un nodo por línea) en la cual se coloca el número del nodo, las coordenadas del nodo, coordenada en X coordenada Y coordenada Z (solo para caso tridimensional), después se coloca las conectividades de los elementos la cual se identifica con un número entero, los nodos son colocados en sentido anti-horario y después al material al que pertenece, a continuación en las siguientes líneas se colocan los nodos que posean alguna carga y por último se colocan los nodos que tienen restricciones siendo el código el siguiente: 0 (cero) dirección restringida y 1 (uno) dirección no restringida.

**Paso 8** (Ejecución del programa): Para la ejecución de programa *MefetGC* se utiliza el archivo de datos realizado en el *paso 14* y se ejecuta el programa, el cual nos dará los esfuerzos que sufre la probeta debido a la carga que actúa en ella, dentro de estos esfuerzos de encuentra el esfuerzo máximo ( $\sigma_{\max}$ ).

**Paso 9** (Cálculo del factor de concentrador de esfuerzo utilizando el esfuerzo máximo obtenido a través del programa *MefetGC*): Con el valor el esfuerzo máximo ( $\sigma_{\max}$ ) obtenido en el *paso 15* y utilizando el valor del esfuerzo nominal obtenido en el *paso 2* se realiza una pequeña operación matemática y se obtiene el factor del concentrador de esfuerzo ( $K_{Mefetgc}$ ).

**Paso 10** (Dibujo de la probeta en el computador): Se dibuja la probeta idealizada en el *Mechanical Desktop* utilizando los datos del *paso 1*.

**Paso 11** (Utilización del *Mechanical Desktop*): Con el dibujo obtenido en el *paso 4*, la carga o cargas que se han de utilizar en la probeta, el módulo de elasticidad (E) y el módulo de Poisson ( $\nu$ ) del *paso 1*, se cargan estos datos en el módulo de análisis de elementos finitos del *Mechanical Desktop*, se restringe los movimientos que la probeta no debe realizar y se ejecuta el programa.

**Paso 12** (Resultado del *Mechanical Desktop*): Cuando el programa termina la tarea encomendada muestra los resultados en la pantalla del computador, donde de éstos obtendremos el esfuerzo máximo ( $\sigma_{\max}$ ) que existe en la probeta.

**Paso 13** (Cálculo del factor del concentrador de esfuerzos utilizando el esfuerzo máximo obtenido a través del *Mechanical Desktop*): Con el valor del esfuerzo máximo ( $\sigma_{\max}$ ) obtenido en el *paso 7* y utilizando el valor del esfuerzo nominal ( $\sigma_0$ ) del *paso 2*, se realiza una pequeña operación matemática y de esta forma se obtiene el valor del factor de concentrador de esfuerzo ( $K_{\text{Mechanical Desktop}}$ )

**Paso 14** (Utilización del *Visual Nastran 4D*): Se importa al *Visual Nastran 4D* el dibujo realizado en el *paso 5*, se introducen las cargas a utilizar, el módulo de elasticidad (E), el módulo de Poisson ( $\nu$ ) del *paso 1*, se restringen los movimientos que la probeta no debe realizar y se ejecuta el programa.

**Paso 15** (Resultado del *Visual Nastran 4D*): Cuando el programa termina la tarea encomendada, muestra los resultados en la pantalla del computador, donde de éstos obtendremos el esfuerzo máximo ( $\sigma_{\max}$ ) que existe en la probeta.

**Paso 16** (Cálculo del factor del concentrador de esfuerzos utilizando el esfuerzo máximo obtenido a través del *Visual Nastran 4D*). Con el valor del esfuerzo máximo ( $\sigma_{\max}$ ) obtenido en el *paso 10* y utilizando el valor del esfuerzo nominal ( $\sigma_0$ ) del *paso 2*, se realiza una pequeña operación matemática y de esta forma se obtiene el valor del factor de concentrador de esfuerzo ( $K_{\text{Visual Nastran}}$ ).

**Paso 17** (Ejecución del programa *Hexafro* [5]): Para ejecutar este programa se modifica el archivo de datos realizado en el *paso 14* para cumplir con la estructura de datos del *Hexafro* [5]. Se corre el programa, el cual dará los esfuerzos que sufre la probeta, dentro de los cuales se encuentra el esfuerzo máximo ( $\sigma_{\max}$ ).

**Paso 18** (Cálculo de factor de concentrador de esfuerzo utilizando el esfuerzo máximo obtenido del programa *Hexafro* [5]): Con el valor del esfuerzo máximo ( $\sigma_{\max}$ ) obtenido en el *paso 17* y utilizando el valor del esfuerzo nominal obtenido en el *paso 2* se realiza una pequeña operación matemática y se obtiene el factor del concentrador de esfuerzo ( $K_{Hexafro}$ ).

**Paso 19** (Análisis de los resultados obtenidos por los programas utilizados anteriormente): Se analizarán los resultados obtenidos por los diferentes programas usados anteriormente, comparando los tipos de mallados obtenidos, su cantidad de nodos, su cantidad de elementos y los resultados de los factores de concentración de esfuerzos.

### ***6.3 Comparación de los programas.***

#### MefetGC:

El programa MefetGC fue desarrollado por [13] en lenguaje C. Tiene una estructura de entrada y salida de datos por archivos de texto. Su utilización se limita a la solución de problemas estáticos en dos y tres dimensiones de elasticidad lineal, no lineal y transferencia de calor. Utiliza el Método del Gradiente Conjugado con preconditionador de Elemento a Elemento. Utiliza elementos tetraédricos y



hexaédricos de cuatro, seis y ocho nodos. El pre-proceso (mallado) se debe realizar de forma manual o utilizando un mallador externo y la entrada de datos se realiza en un archivo de texto. La salida de datos es realizada por el programa también en un archivo de texto. El post-proceso (visualización) se realiza también de forma manual o con un visualizador externo. No hace referencia al número máximo de elementos que maneja.

#### Mechanical Desktop:

El Autodesk Mechanical Desktop 6.0 es un programa de diseño asistido por computadora con una plataforma C.A.D. proveniente del Autodesk Autocad 2002. Este programa posee un módulo de análisis por elementos finitos que se limita al análisis estático en dos y tres dimensiones de elasticidad lineal. No contempla cargas dinámicas ni efectos por temperatura. Posee incorporado un módulo de pre-proceso (mallador) y post-proceso (visualizador). En el proceso divide el área en triángulos y luego aproxima la solución utilizando interpolación numérica polinómica. Se limita al uso de elementos tipo triángulo (triángulos de deformación constante) de seis nodos. El método de solución utilizado es el propuesto por Robert D. Cook, *Concept and Applications of Finite Element Analysis*, 1974, pág. 81-84. No hace referencia al número máximo de elementos que maneja.

#### Visual Nastran 4D:

El MSC Visual Nastran 4D es un programa de simulación de fenómenos físicos que tiene una plataforma propia similar a la plataforma C.A.D. Su módulo de análisis de elementos finitos abarca elasticidad lineal, h-adaptabilidad, pandeo, vibración y transferencia de calor en tres dimensiones. Contempla cargas estáticas y dinámicas.

Posee incorporado un módulo de pre-proceso (mallador) y post-proceso (visualizador). No hace referencia al tipo de elementos que utiliza. El cálculo de los esfuerzos lo realiza en cada nodo elemento a elemento. No hace referencia al número máximo de elementos que maneja.

#### Hexafrofron [5]:

Es un programa realizado en lenguaje C. Tiene una estructura de entrada y salida de datos muy similar al programa MefetGC. Su utilización se limita únicamente a la solución de problemas estáticos en tres dimensiones de elasticidad lineal. Utiliza el método de solución frontal, al cual se hace referencia en [5]. Se limita exclusivamente a la utilización de elementos hexaédricos de ocho nodos. El pre-proceso (mallado) se debe realizar de forma manual o utilizando un mallador externo y la entrada de datos se realiza en un archivo de texto. La salida de datos es realizada por el programa también en un archivo de texto. El post-proceso (visualización) se realiza también de forma manual o con un visualizador externo. Fue creado por [5] y lo utilizan como ejemplo práctico en su libro. No hace referencia al número máximo de elementos que maneja.

### **6.4 Resultados.**

#### MefetGC:

Se realizó una corrida con 144 nodos y 55 elementos hexaédricos.

$$k_{MefetGC} = \frac{\sigma_{\max}}{\sigma_0} = \frac{4,94806 * 10^9}{2,375 * 10^9} = \underline{2,083}$$

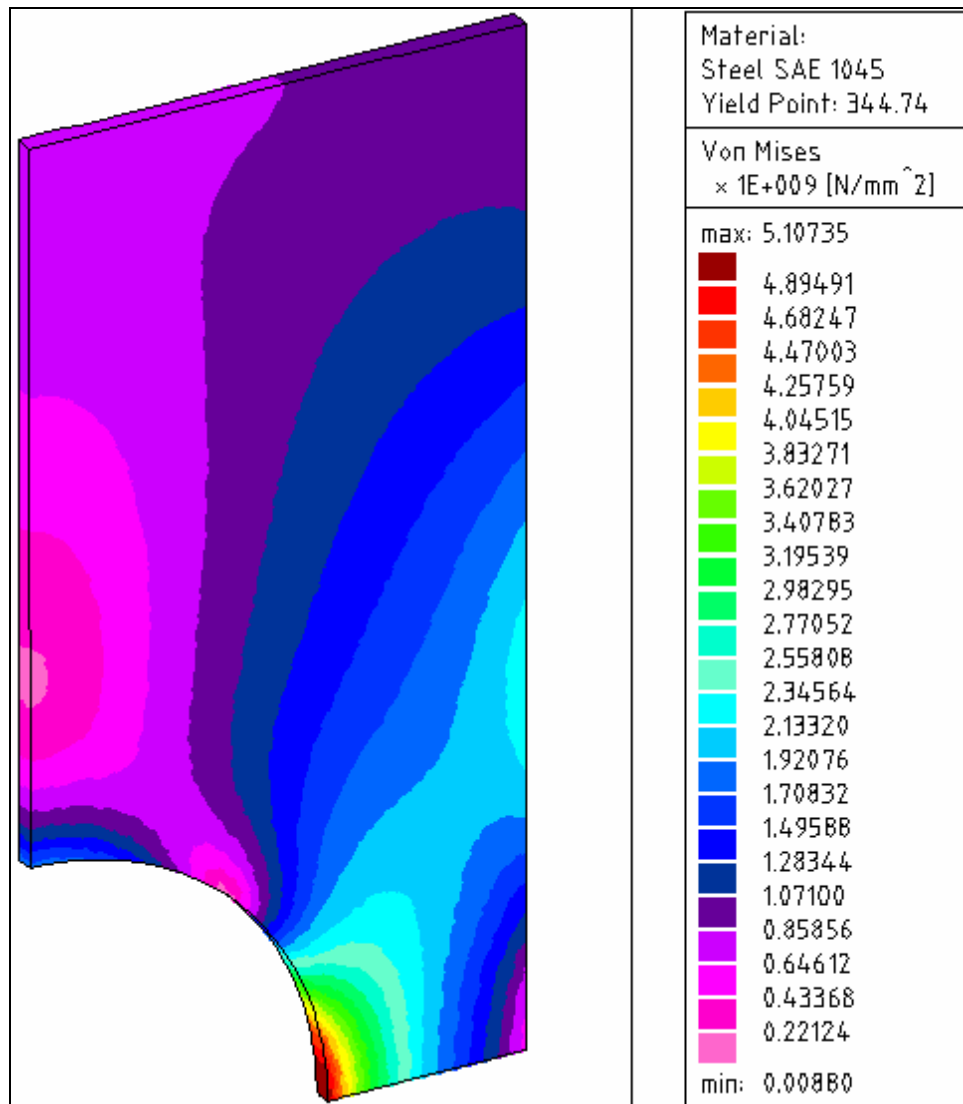
Elemento 1
4,90472E+09
4,94806E+09
3,79375E+09
3,76507E+09
4,88937E+09
4,93388E+09
3,79889E+09
3,76933E+09
4,92551E+09
4,36908E+09
3,77831E+09
4,33303E+09
4,91094E+09
4,36501E+09
3,78325E+09
4,32795E+09
4,87852E+09
4,92205E+09
3,77170E+09
3,74312E+09
4,34572E+09
4,30955E+09
4,35008E+09
4,34565E+09
4,89950E+09
3,75643E+09
4,32674E+09
Máx
4,94806E+09

**Tabla 1** Resultados de los esfuerzos en el elemento 1 para el programa MefetGC. Se presenta el esfuerzo máximo de Von Misses.

Mechanical Desktop:

Se realizó una corrida con 6398 nodos y 24142 elementos tetraédricos.

$$k_{\text{MechanicalDesktop}} = \frac{\sigma_{\text{max}}}{\sigma_0} = \frac{5,10735 * 10^9}{2,375 * 10^9} = \underline{2,1504}$$

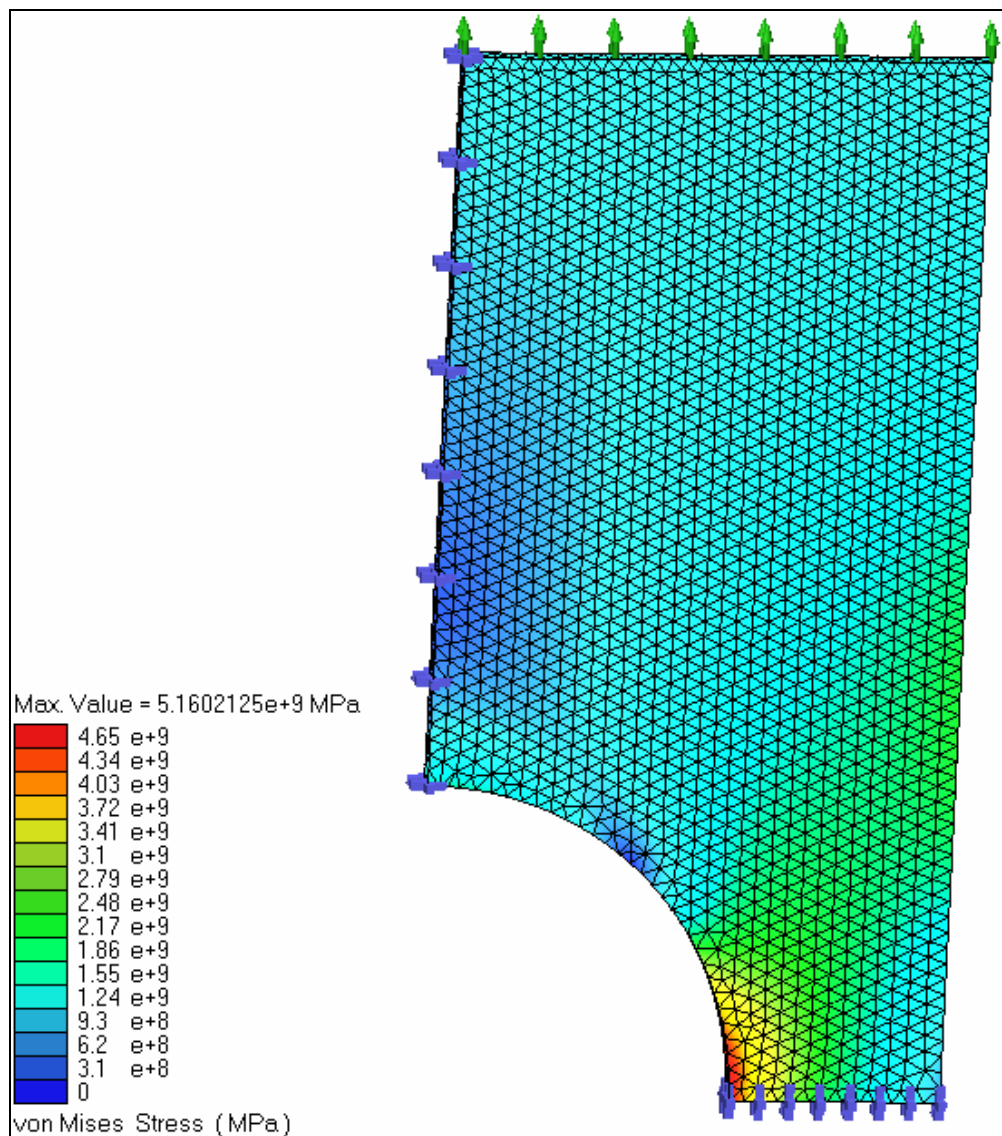


**Figura 6-4** Resultados del programa Mechanical Desktop.

### Visual Nastran:

Se realizó una corrida con 37649 nodos y 22112 elementos. El programa no hace referencia al tipo de elementos utilizados.

$$k_{\text{VisualNastran4D}} = \frac{\sigma_{\text{max}}}{\sigma_0} = \frac{5,1602125 * 10^9}{2,375 * 10^9} = \underline{2,1727}$$



**Figura 6-5 Resultados del programa Visual Nastran 4D.**

Hexafron [5]:

Se realizó una corrida con 144 nodos y 55 elementos hexaédricos.

$$k_{Hexafron} = \frac{\sigma_{\max}}{\sigma_0} = \frac{6,02130 * 10^9}{2,375 * 10^9} = \underline{2,535}$$

Elemento 1
6,02130E+09
3,20570E+09
4,59420E+09
1,44310E+09
3,23690E+09
1,61510E+09
6,86830E+08
9,73290E+08
Máx
6,02130E+09

**Tabla 2** Resultados de los esfuerzos en el elemento 1 para el programa Hexafron [5]. Se presenta el esfuerzo máximo de Von Mises.

$K_r$  Teórico:

Se calculó a partir del siguiente programa:

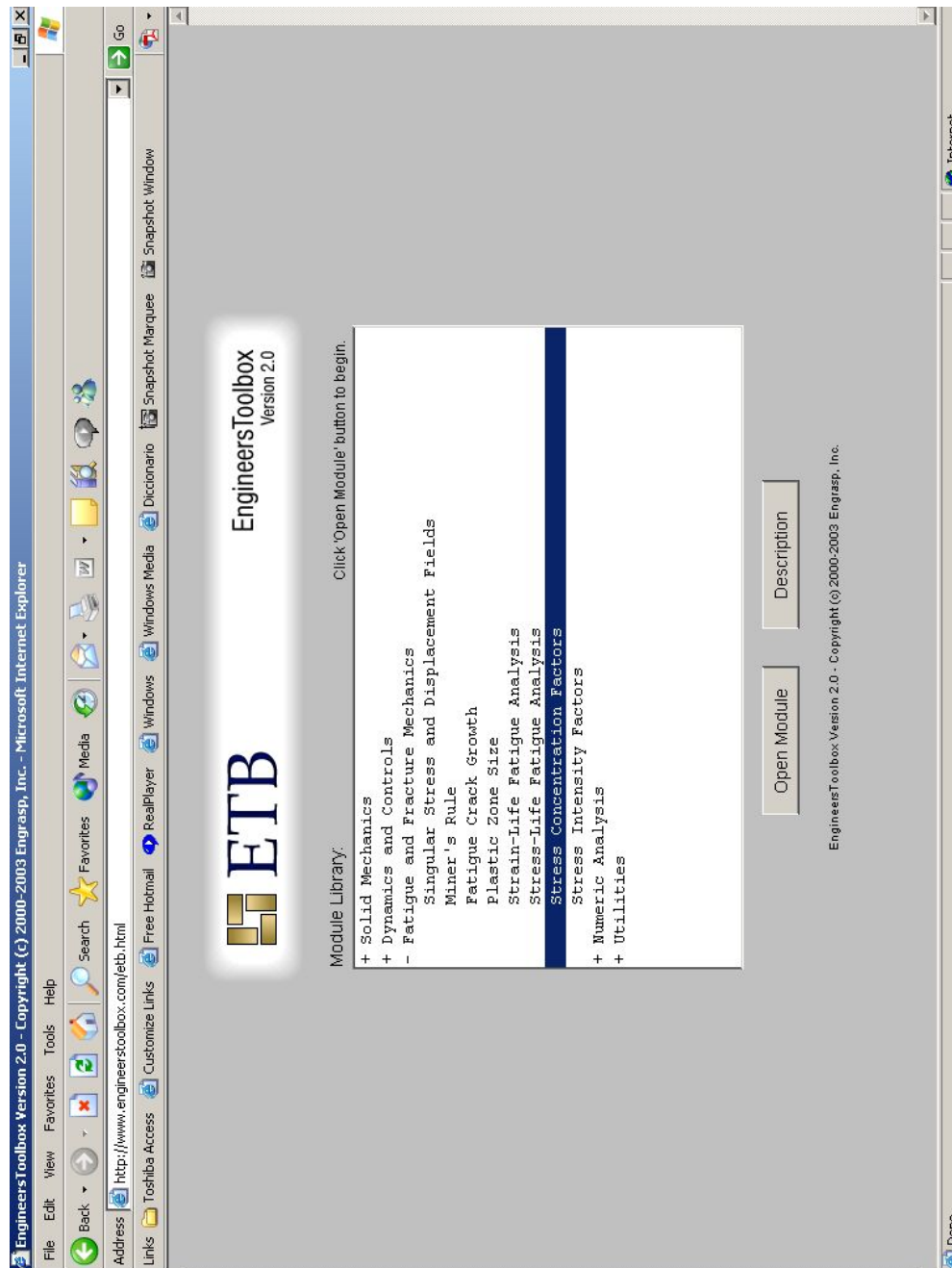


Figura 6-6<sup>18</sup> Programa Engineers Toolbox v2.0.

<sup>18</sup> Tomado de <http://www.engineerstoolbox.com>

Resultando en un valor de:  $K_t = 2,1091$

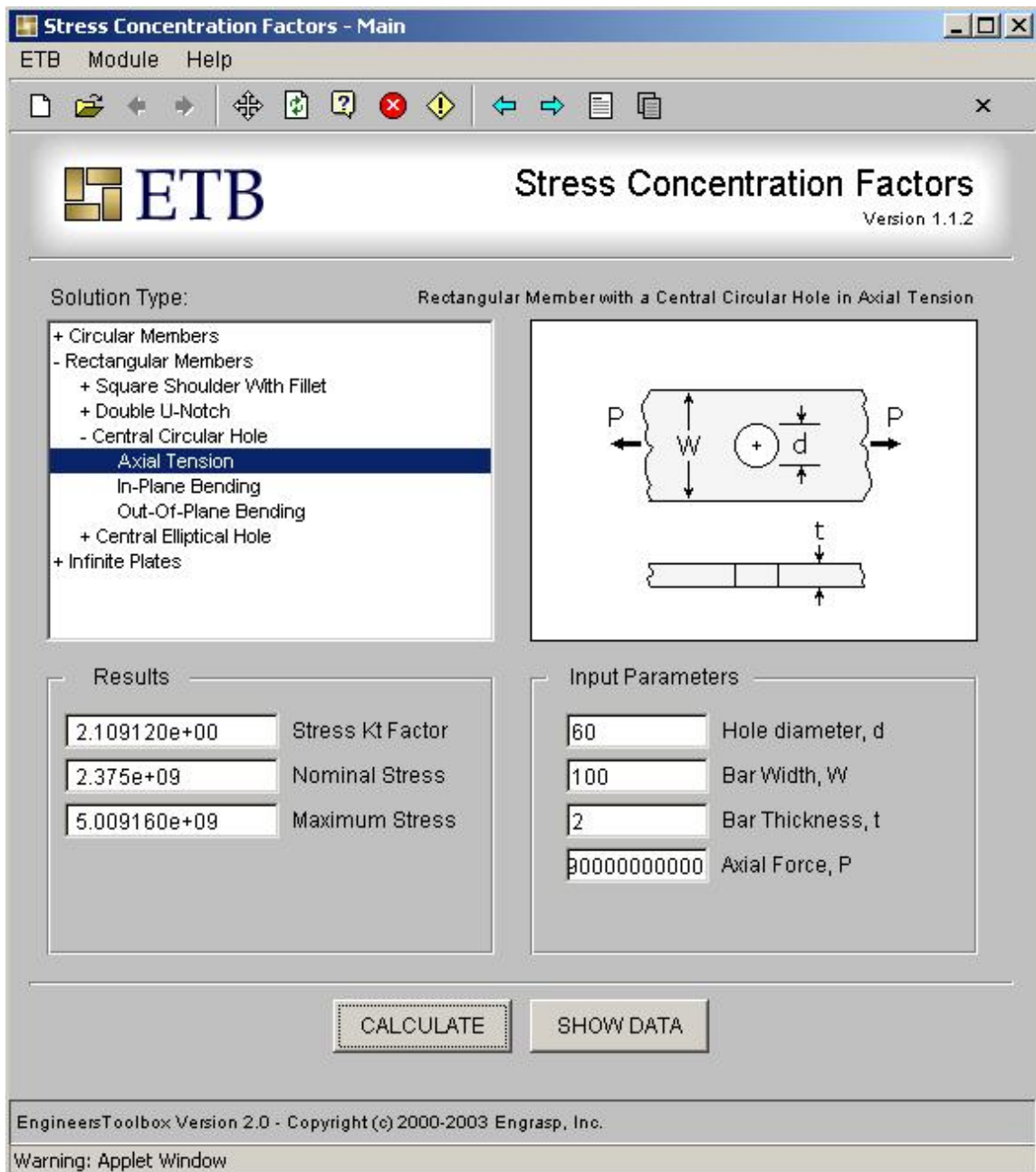


Figura 6-7<sup>19</sup> Resultado del  $K_t$ , teórico.

<sup>19</sup> Tomado de <http://www.engineerstoobox.com>



Teniendo su justificación en las siguientes ecuaciones:

- Placa plana con agujero circular sometido a tensión axial.

$$\frac{d}{W} \leq 0.65: \quad K_t = 3.0039 - 3.753 \left( \frac{d}{W} \right) + 7.9735 \left( \frac{d}{W} \right)^2 - 9.2659 \left( \frac{d}{W} \right)^3 + 1.8145 \left( \frac{d}{W} \right)^4 + 2.9684 \left( \frac{d}{W} \right)^5$$

$$\frac{d}{W} > 0.65: \quad K_t = 2.0 + \left( 1.0 - \frac{d}{W} \right)^3$$

Las cuales fueron mejoradas a partir de la ecuación de concentración de esfuerzos para placa plana finita con agujero, presentada por [32].

Estas ecuaciones son un ajuste de la curva mostrada en la Figura 6-8, tomada de [30].

El error se estimó de la siguiente forma:

$$\% \text{Error} = \frac{|K_t - K_x|}{K_t} * 100$$

Donde  $K_t$  es el valor teórico y  $K_x$  es el valor calculado por medio de los programas.

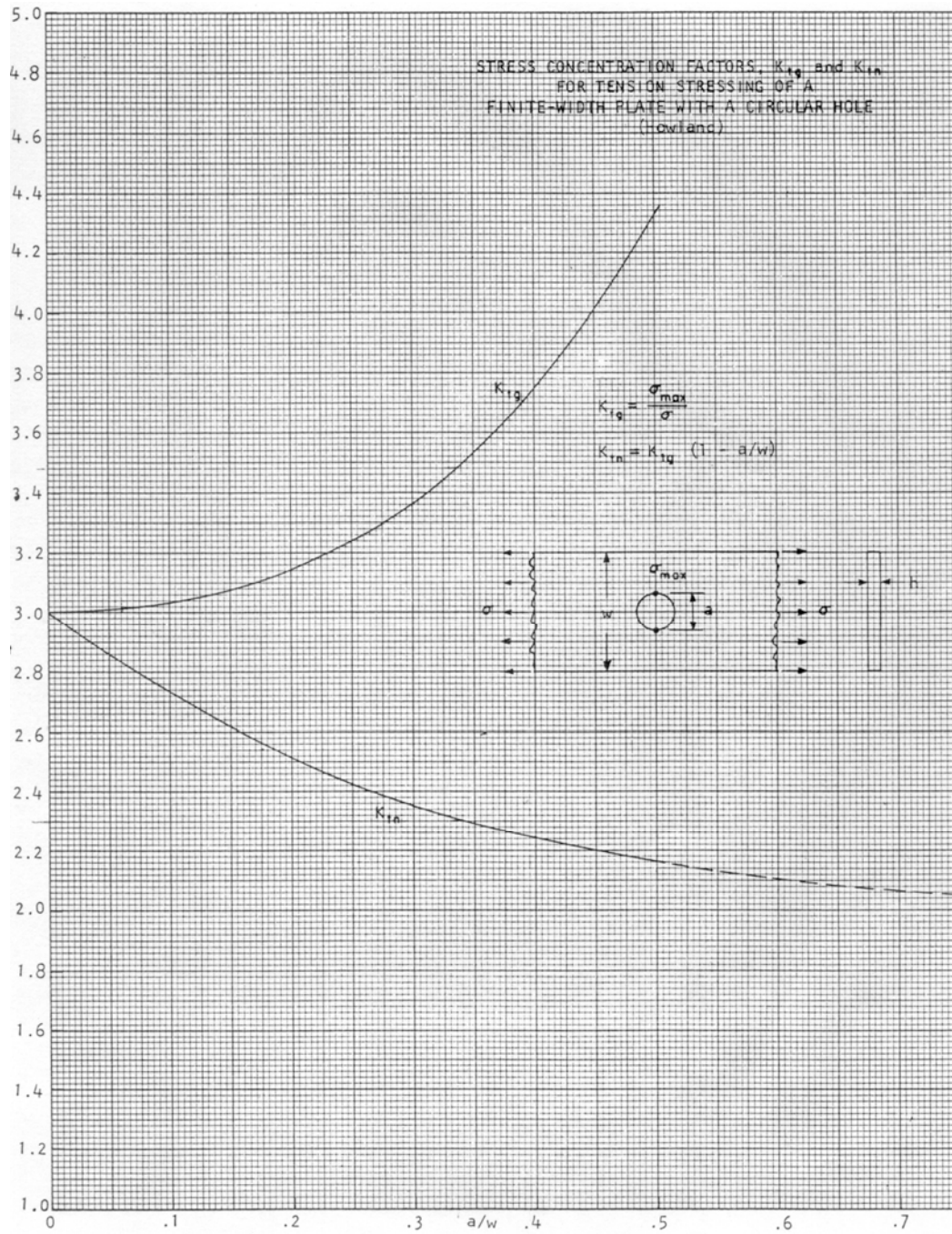
Para cada uno de los programas resultó un porcentaje de error. Se muestran a continuación:

$$\% \text{Error}_{\text{MefetGC}} = \frac{|2,083 - 2,1091|}{2,1091} * 100 = \underline{1,23\%}$$

$$\% \text{Error}_{\text{Hexafon}} = \frac{|2,535 - 2,1091|}{2,1091} * 100 = \underline{20,19\%}$$

$$\% \text{Error}_{\text{MechanicalDesktop}} = \frac{|2,1504 - 2,1091|}{2,1091} * 100 = \underline{1,95\%}$$

$$\% \text{Error}_{\text{VisualNastran4D}} = \frac{|2,1727 - 2,1091|}{2,1091} * 100 = \underline{3,01\%}$$



**Figura 6-8<sup>20</sup>** Curva del factor de concentración de esfuerzos para placa finita con agujero circular sometida a tensión.

<sup>20</sup> Tomado de [30]

## 7. Análisis de resultados.

En la actualidad, la preferencia de la industria es que los resultados son aceptables sólo cuando estos han sido resueltos usando ciertos programas estándares comerciales. Los programas comerciales manejan plataformas de entradas de datos amigables al usuario, sin embargo no proveen información sobre las formulaciones y métodos de solución [5]. En nuestro caso pudimos corroborar esta afirmación al realizar el análisis de los resultados. Los programas comerciales que utilizamos fueron el *Mechanical Desktop* y *Visual Nastran 4D*, los cuales proporcionan plataformas amigables al usuario, pero no dan información sobre las formulaciones y métodos de ensamblaje y solución.

El Método del Gradiente Conjugado con preconditionador de elemento a elemento fue implementado en el programa *Mefet*, transformándolo en *MefetGC*. Este programa fue el que obtuvo el menor error (1,23%). Este resultado se debe a una combinación de varios factores, como el uso del Método del Gradiente Conjugado, el uso de elementos hexaédricos de ocho nodos, el uso de valores de doble precisión para la realización de todos los cálculos, la reducción del error acumulado al procesar grandes filas y/o columnas y la utilización del preconditionador de elemento a elemento. Esto optimiza el método de solución haciendo más preciso el programa.

El programa comercial *Mechanical Desktop* obtuvo un mejor resultado con respecto a *Visual Nastran 4D*. Obtuvo un resultado del 1,95% de error con respecto al valor teórico que posee la pieza. El programa provee información referente al libro del cual implementaron el sistema de solución (Robert D. Cook, *Concept and Applications of Finite Element Analysis*, 1974), pero no provee información acerca del método de resolución implementado, ya que no se tiene a disposición el código fuente. El

programa informa que utiliza un mallado de elementos triangulares de deformación constante, con el cual logra optimizar el rendimiento del computador.

El programa comercial *Visual Nastran 4D* obtuvo un buen desempeño en los resultados, dando un error del 3,01% con respecto al valor teórico. En este programa tampoco se pudo acceder a los códigos fuente del programa, y no se provee información acerca del método de solución ni del tipo de elementos utilizados.

Se puede observar que el error obtenido en el programa *Hexafro* [5] es del 20,19% con respecto al valor teórico. Este resultado elevado nos indica que el mallado utilizado en la discretización del dominio no fue el óptimo. Debido a que el programa no posee un generador de malla como los paquetes comerciales usados, se empleó la misma malla que se usó en el programa *MefetGC*, la cual fue realizada manualmente. También se debe tener en cuenta que este programa utiliza el método de solución frontal, el cual es una variación del método de Gauss. Este método es menos preciso que el Método del Gradiente Conjugado con el preconditionador de elemento a elemento.

El esfuerzo de Von Misses es usado como un criterio para determinar la aparición de una falla en materiales dúctiles. Este criterio de falla establece que el esfuerzo de Von Misses ( $\sigma_{vm}$ ) debe ser menor que el esfuerzo de fluencia ( $\sigma_y$ ) del material, por lo tanto, es el mayor esfuerzo que aparece en el material. Se utilizó este esfuerzo para acercarnos lo más posible al esfuerzo teórico.

El pre-proceso y el pos-proceso son parte fundamental para el análisis por elemento finito. El preprocesamiento implica la preparación de datos, los cuales incluye las coordenadas nodales, la conectividad, la información de las cargas que actúan sobre la pieza, el material y las condiciones existentes. Por otro lado el posprocesamiento

consiste en la presentación de los resultados obtenidos por el procesamiento de los datos.

Para problemas grandes el número de términos no nulos dentro del perfil de la matriz de coeficientes es pequeño comparado con el número de elementos nulos. Esto sucede en mayor grado en problemas tridimensionales donde la altura de las columnas puede ser de varios miles de términos entre las cuales menos de cien sean no nulos antes de construir los factores triangulares (que rellenarán los elementos nulos). En consecuencia, para este tipo de problemas serán, en general, los métodos iterativos más eficientes que la solución directa. Las principales ventajas de la iteración son la reducción del almacenamiento de la memoria central y la eliminación de la descomposición triangular, que es la parte más costosa de la solución directa. Por lo tanto en los métodos iterativos, se puede utilizar un mayor número de elementos al mismo costo computacional.

## 8. Conclusiones.

La incorporación del sistema de solución de ecuaciones de alta capacidad al programa *Mefet*, fue exitoso. Se aumentó la capacidad del programa para resolver problemas en los cuales el tamaño de la malla es grande.

Se recopiló la información sobre los algoritmos de solución del sistema de ecuaciones generado en la formulación del Método de los Elementos Finitos y los algoritmos de solución basados en el Método del Gradiente Conjugado, en problemas de Elasticidad Lineal.

Se implementó la biblioteca de funciones, en la cual se aplicaron las técnicas mencionadas, y se acopló la biblioteca al programa ya existente.

Se realizó la validación del programa *MefetGC* utilizando un problema de solución conocida y se realizó la comparación con programas comerciales como el *Mechanical Desktop* y el *Visual Nastran 4D*.

El método de solución incorporada al programa *MefetGC* posee un alto rendimiento y precisión. Tan sólo produjo un 1,23% de error con respecto al valor teórico, manejando un número relativamente bajo de elementos (55 elementos) con respecto a los programas comerciales con los cuales se realizó la comparación. Los programas comerciales produjeron un mayor porcentaje de error utilizando un mayor número de elementos.

## 9. Recomendaciones.

- Creación e implementación de un mallador para el programa de *MefetGC*.
- Creación e implementación de un visualizador de resultados.
- Reestructurar el programa *MefetGC* en su módulo de 2D.
- Creación de un laboratorio de simulación para el desarrollo de prácticas de laboratorio de las materias Mecánica de Sólidos, Diseño II y Diseño III empleando el programa *MefetGC*.
- Realizar más trabajos de investigación y desarrollo de tecnologías propias en la Escuela de Ingeniería Mecánica de la Universidad Central de Venezuela.
- Se recomienda para futuros trabajos con el programa *MefetGC*, el estudio de modelos constituidos por materiales anisotrópicos.
- Para realizar el análisis de modelos muchos más complejos se recomienda implementar técnicas de paralelización en el programa *MefetGC*.
- Realizar un trabajo dedicado básicamente a probar y medir el rendimiento del programa *MefetGC*, con un banco de pruebas mucho más extenso. Esta medición del rendimiento escapó a los alcances de nuestro trabajo.

# ANEXOS



## A. Matrices Dispersas (Tipo Sparse).

Las discretizaciones de las Ecuaciones Diferenciales en Derivadas Parciales típicamente nos llevan a matrices grandes y dispersas. Se define como matriz dispersa, de una manera informal, como una matriz que tiene muy pocas entradas no nulas. Pero en forma práctica, una matriz puede ser catalogada como dispersa cuando técnicas especiales pueden ser utilizadas para tomar ventaja de su gran cantidad de elementos nulos y su ubicación. Estas técnicas para matrices dispersas provienen de la idea de que los elementos nulos no necesitan ser almacenados. Uno de los puntos claves es definir estructuras de datos para estas matrices que sean adecuadas para la implementación eficiente de métodos de solución, tanto iterativos como directos.

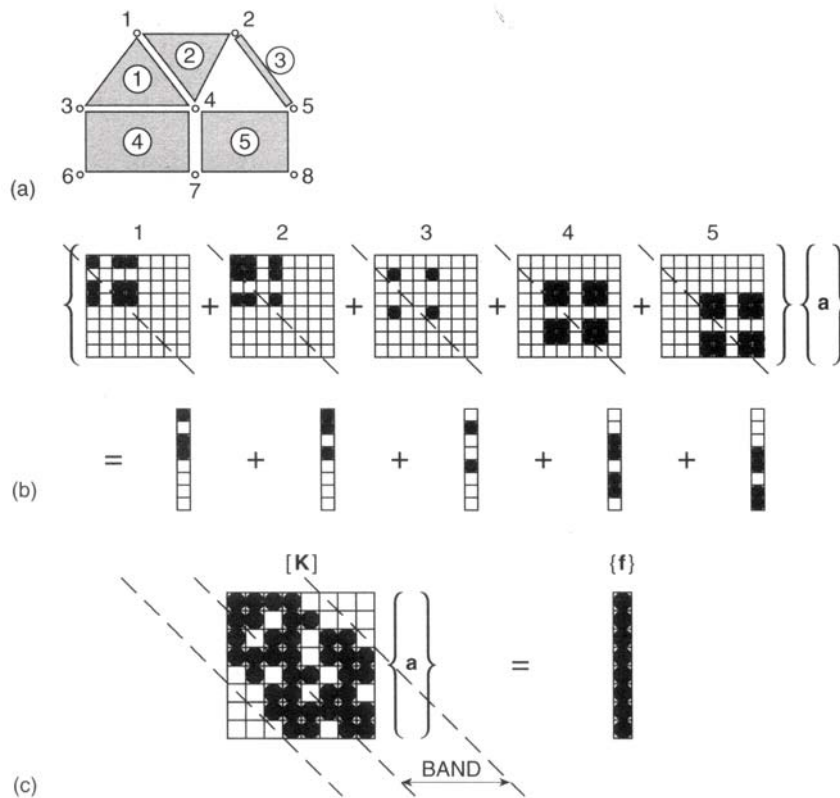
### A.1 *Matrices Dispersas.*

Existen dos tipos esenciales de matrices dispersas, las *estructuradas* y las *no estructuradas*.

Una matriz *estructurada* es una en la que sus entradas no nulas forman un patrón regular, regularmente sucede en un pequeño número de diagonales. Alternativamente, los elementos no nulos se encuentran en forma de bloques (submatrices densas) del mismo tamaño, los cuales forman un patrón regular, típicamente a lo largo de un pequeño número de diagonales (en forma de bloque).

Una matriz *no estructurada* es una en la que sus entradas no nulas son localizadas en forma irregular.

La diferencia entre estos dos tipos de matrices no afecta significativamente los métodos de solución directos. Pero esta distinción si es importante en métodos iterativos. En estos métodos, una de las operaciones esenciales son los productos entre matrices y vectores. El rendimiento de estas operaciones puede diferir significativamente en computadores, dependiendo de si son estructuradas o no estructuradas.



**Figura A-0-1<sup>21</sup>** Patrón general. Cada cuadro oscurecido representa un coeficiente o submatriz si más de una cantidad está siendo considerada en los nodos.

<sup>21</sup> Tomado de [43]

## B. Factorización de Cholesky.

La Factorización de Cholesky es un método que se basa en otro tipo de descomposición de la matriz del sistema. Sabiendo que una matriz Hermítica<sup>22</sup> y definida positiva tiene sus autovalores reales positivos, y que en la Factorización  $LU$  todos los pivotes son reales positivos.

Por ende toda matriz  $A$  hermítica y definida positiva puede ser descompuesta de la forma  $A = BB^*$ , siendo  $B$  una matriz triangular inferior.

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & u_{12} & \dots & u_{1n} \\ 0 & u_{22} & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 & \dots & 0 \\ l_{21} & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & \dots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & 0 & \dots & 0 \\ 0 & u_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & \dots & u_{1n} \\ 0 & 1 & \dots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$$

$$A = L \begin{bmatrix} u_{11} & 0 & \dots & 0 \\ 0 & u_{22} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & u_{nn} \end{bmatrix} R$$

$$A = L \begin{bmatrix} \sqrt{u_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{u_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{u_{nn}} \end{bmatrix} \begin{bmatrix} \sqrt{u_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{u_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{u_{nn}} \end{bmatrix} R$$

<sup>22</sup> Refiérase al Anexo D.

Donde  $B$  es una matriz triangular inferior

$$B = L \begin{bmatrix} \sqrt{u_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{u_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{u_{mm}} \end{bmatrix}$$

y  $C$  es una matriz triangular superior.

$$C = \begin{bmatrix} \sqrt{u_{11}} & 0 & \dots & 0 \\ 0 & \sqrt{u_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sqrt{u_{mm}} \end{bmatrix} R$$

Como la matriz  $A$  es Hermítica,  $BC = A = A^* = B^*C^*$  por lo que  $(C^*)^{-1}B = B^*C^{-1}$  y sabiendo que  $(C^*)^{-1}B$  es la matriz triangular inferior y  $B^*C^{-1}$  es la matriz triangular superior, ambas deben ser diagonales.

Por otra parte:

$$B = LD$$

$$C = DR$$

$$C^* = R^*D^* = R^*D$$

$$(C^*)^{-1} = D^{-1}(R^*)^{-1}$$

$$(C^*)^{-1}B = D^{-1}(R^*)^{-1}LD$$

Como las matrices lineales conmutan:

$$(C^*)^{-1} B = D^{-1} D (R^*)^{-1} L = (R^*)^{-1} L$$

Siendo  $(R^*)^{-1} L$  una matriz triangular inferior, la cual posee unos en su diagonal principal y  $(C^*)^{-1} B$  una matriz diagonal, podemos asegurar que  $(R^*)^{-1} L = I$ , por tanto  $R^* = L$ . Además,  $B^* C^{-1} = I$ , por lo que  $C = B^*$ , por ende  $A = BB^*$  donde  $B = LD$ .

La unicidad de las matrices  $L$  y  $U$  implica la unicidad de la matriz  $[B]$  y, por lo tanto, ésta puede ser calculada por un método directo.

### **Teorema.**

*“Una matriz hermítica y regular  $A$  es definida positiva si, y solo si admite la factorización de Cholesky.”*

Si  $A$  admite la factorización de Cholesky entonces  $A = BB^*$ . Por lo tanto:

$$A = (BB^*)^* = BB^* = A$$

Donde  $A$  es Hermítica y para cualquier vector  $X$  no nulo es:

$$x^* A x = x^* B B^* x = (B^* x)^* (B^* x) = \|B^* x\|^2 \geq 0$$

Siendo  $B^*x = 0$  pero al ser  $B^*$  regular (si no lo fuese tampoco lo sería  $A$  en contra de la hipótesis)  $B^*x = 0 \Rightarrow x = 0$  en contra de la hipótesis que  $x$  no es el vector nulo. Se tiene pues que  $x^*Ax > 0$  para cualquier vector  $x$  no nulo, es decir,  $A$  es definida positiva.

## C. Número de Condición.

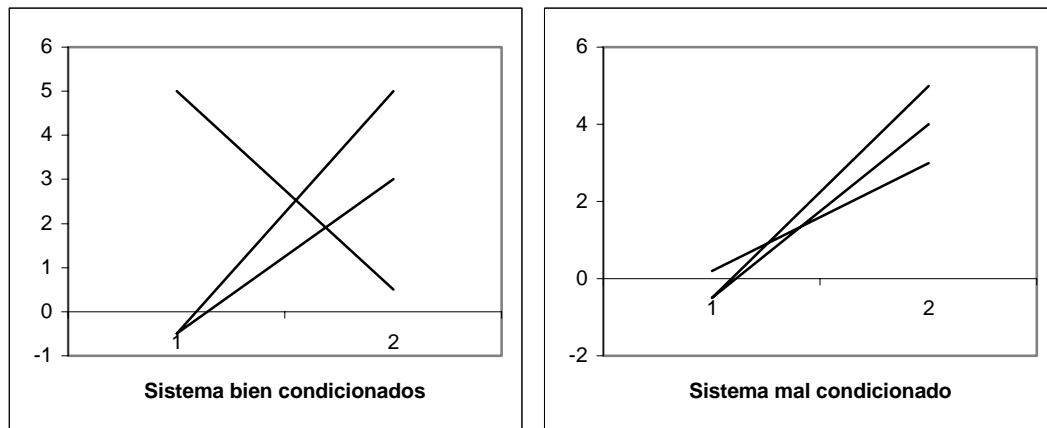
### *C.1 Condicionamiento de un sistema.*

Un sistema de ecuaciones lineales  $Ax=b$  se dice que se encuentra bien condicionado cuando los errores cometidos en la matriz  $A$  y del vector  $b$  produce una solución con un error del mismo orden. En cambio diremos que el sistema está mal condicionado cuando produce un error en la solución de orden superior a la de los datos de entrada.

Podremos decir si:

$$\begin{aligned}\|A - \bar{A}\| &< \varepsilon \\ \|b - \bar{b}\| &< \varepsilon\end{aligned}$$

El sistema se encuentra bien condicionado cuando  $\|x - \bar{x}\| \approx \varepsilon$  y estará mal condicionado cuando  $\|x - \bar{x}\| \gg \varepsilon$ . Esto se puede ver claramente en un sistema de orden dos, en el cual la solución es representada por el punto de intersección de ambas rectas en el plano. Un pequeño error en la pendiente de una de ellas puede hacer que dicho punto de corte se desplace solo un poco o una distancia considerable (ver Figura C-0-1), lo cual puede determinar si el sistema está bien o mal condicionado.



**Figura C-0-1** Sistema bien condicionado y sistema mal condicionado.

Siguiendo con el ejemplo anterior, podemos ver que el sistema se encuentra mal condicionado cuando las pendientes de las rectas son similares, mientras que el sistema se encuentra mejor condicionado cuando las rectas son más ortogonales.

### Ejemplo.

Se tiene el siguiente sistema:

$$\begin{aligned} 3x + 4y &= 7 \\ 3y + 4,00001y &= 7,00001 \end{aligned}$$

Posee la siguiente solución:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$



Suponiendo que se comete un pequeño error en los datos de entrada, obteniendo el siguiente sistema.

$$\begin{aligned}3x + 4y &= 7 \\3x + 3,99999y &= 7,00004\end{aligned}$$

Obteniendo la siguiente solución:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7,6 \\ -4 \end{pmatrix}$$

Con lo que podemos ver que el sistema está mal condicionado, pero si a la segunda ecuación la sustituimos por la que resulta de sumarle la primera multiplicada por (-1,0000016) y la ecuación resultante se multiplica por  $(1 \cdot 10^6)$  y se divide por (-1,2) el sistema nos queda de la siguiente forma:

$$\begin{aligned}3x + 4y &= 7 \\4x + 3y &= 1\end{aligned}$$

El cual tiene la siguiente solución:

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Siendo éste un sistema bien condicionado.

## C.2 Número de condición de una matriz.

El estudio del condicionamiento de un sistema se realiza a través del *número de condición*, el cual se denota por  $k(A)$ . Donde  $A$  es una matriz cuadrada y regular.

$$k(A) = \|A\| \cdot \|A^{-1}\|$$

Ya que en la práctica el cálculo de la matriz inversa  $A^{-1}$  es un proceso largo, lo que se hace es buscar la cota del número de condición.

$$k(A) = \|A\| \cdot \|A^{-1}\| < \|A\| \cdot k$$

Donde  $k$  es una cota de la norma de la matriz inversa.

Si  $\|I - A\| < 1$ , entonces  $\|A^{-1}\| \leq \frac{\|I\|}{1 - \|I - A\|}$ . En efecto:

$$\begin{aligned} A \cdot A^{-1} &= I \\ [I - (I - A)] A^{-1} &= I \\ A^{-1} - (I - A) A^{-1} &= I \\ A^{-1} &= I(I - A) A^{-1} \\ \|A^{-1}\| &= \|I + (I - A) A^{-1}\| \leq \|I\| + \|(I - A) A^{-1}\| \leq \|I\| + \|I - A\| \|A^{-1}\| \\ \|A^{-1}\| - \|I - A\| \|A^{-1}\| &\leq \|I\| \\ (1 - \|I - A\|) \|A^{-1}\| &\leq \|I\| \\ \|A^{-1}\| &\leq \frac{\|I\|}{1 - \|I - A\|} \end{aligned}$$

Es decir:

$$k(A) \leq \|A\| \cdot k \quad \text{con} \quad k = \frac{\|I\|}{1 - \|I - A\|}$$

Se debe tener cuidado con esta acotación ya que si tenemos una matriz regular, que está muy cercana de ser singular, es decir con  $\det(A) \simeq 0$ , entonces  $A$  tiene un autovalor próximo a cero. Por lo que la matriz  $I - A$  tiene un autovalor próximo a 1 y será el mayor de todos. En este caso  $\|I - A\| \simeq 1$ , por lo que  $k \rightarrow \infty$  y daría lugar a un falso condicionamiento.

### Ejemplo.

Dado el siguiente sistema:

$$\begin{aligned} 3x + 4y &= 7 \\ 3x + 4,00001y &= 7,00001 \end{aligned}$$

Se tiene que:

$$A = \begin{pmatrix} 3 & 4 \\ 3 & 4,0001 \end{pmatrix} \Rightarrow \det(A) = 0,00003$$

$$A^{-1} = \frac{1}{0,00003} \begin{pmatrix} 4,00001 & -4 \\ -3 & 3 \end{pmatrix}$$

Utilizando la norma infinito  $\|A\|_{\infty} = \max \sum_{j=1}^n |a_{ij}|$  se tiene que:

$$\left. \begin{array}{l} \|A\|_{\infty} = 7,00001 \\ \|A^{-1}\|_{\infty} = \frac{8,00001}{0,00003} \end{array} \right\} \Rightarrow k_{\infty}(A) \approx \frac{56}{3} \cdot 10^5 > 1,8 \cdot 10^6$$

Se puede ver que el sistema esta mal condicionado.

## D. Matriz Hermítica.

Por definición una matriz se llama normal si se puede conmutar con su transpuesta conjugada. Esto es si se satisface la relación:

$$A'' A = A A''$$

Recordemos del capítulo 4 la ecuación (3.16), que nos definía que dos vectores  $d_{(i)}$  y  $d_{(j)}$  son conjugados si:

$$d_{(i)}^T A d_{(j)} = 0$$

Una propiedad resultante de esta relación es que si una matriz normal es triangular, entonces es una matriz diagonal. En consecuencia, una matriz es normal si, y solo si, es unitariamente similar a la matriz diagonal.

Una matriz unitaria se define como  $Q'' Q = I$ . Una matriz unitaria  $Q$  es una matriz cuya inversa es su transpuesta conjugada  $Q''$ , de forma que:

$$Q'' Q = 1 \quad \rightarrow \quad Q^{-1} = Q''$$

Una matriz  $Q$  que cumple que  $Q'' Q$  sea diagonal se le llama matriz ortogonal.

Se puede decir inclusive que una matriz  $A$  es normal si, y sólo si, cada uno de sus autovectores son también autovectores de  $A''$ .

Las matrices *Hermíticas* son aquellas matrices normales cuyos autovectores son reales  $\sigma_{(A)} \in \mathbb{R}$ . Se define como  $A^H = A$ . Cualquier matriz *Hermítica* es similar unitariamente a una matriz diagonal real.

## Referencias/Bibliografías.

- [1] Alonzo G. Axel. y Oramas M. Maribel. *Elaboración e implementación de un sistema de representación gráfica de entidades geométricas bi y tridimensionales en computadores personales con aplicación al diseño mecánico y la bioingeniería*. Trabajo Especial. U.C.V. Facultad de Ingeniería. Escuela de Ingeniería Mecánica. 1991.
- [2] Brown, D. K., *An introduction to the Finite Element Method using BASIC programs*, Surrey University Press, New York, 1984.
- [3] Carey, G. F. & Oden, T., *Finite elements*, Prentice Hall, New Jersey, 1984.
- [4] Carey, G. F. & Jiang, B. *Element-by-element linear and nonlinear solution schemes*, Communications in applied numerical methods, 2, 145-153, 1986.
- [5] Chandrupatla, T. R., Belegundu, A. D., *Introduction to finite elements in engineering*, 3<sup>ra</sup> Edición, Prentice Hall of India, New Delhi, 2002.
- [6] Ciçek, K., Sefa, K., *Stress analysis of a truck chassis with riveted joints*, Finite Elements in Analysis and Design, 38, 1115-1130, 2002.
- [7] Chapman, R. T. & Cox, D. L., *A unique element storage implementation of the vectorized element by element preconditioned conjugate gradient algorithm*, Iterative Equation Solvers for Structural Mechanics Problems, 4, 57-65, 1991.
- [8] Cobos, J., *Apuntes de álgebra numérica*, ETS de Ingeniería Informática, 2003.

- [9] Dayde, M. J., L'Excellent, J. Y. & Gould, N. I. M., *Solution of structured systems of linear equations using element-by-element preconditioners*. Proceeding of the 2<sup>nd</sup> IMACS International Symposium Iterative methods in linear algebra, 181-190, 1995.
- [10] Daydé, M. J., Decamps, J. P., L'Excellent, J. Y. & Gould, N. I. M., *On the solution of scale partially separable unconstrained optimization problems using element-by-element preconditioners*, ENSEEIHT-IRIT Technical Report, 96, 3, 1996.
- [11] Dickinson, J. K. & Forsyth, P. A., *Preconditioned conjugate gradient methods for three-dimensional linear elasticity*, International Journal for Numerical Methods in Engineering, 37, 2211-2234, 1994.
- [12] Eles, R. J. & Utku, S., *Verification tests for computer-aided structural analysis*, Microcomputers in Civil Engineering, 3, 289-297, 1988.
- [13] Fernández S. Arturo y Gómez C. Ganímedes. *Desarrollo e implementación de un sistema computacional unificado para la resolución de problemas de elasticidad y transferencia de calor, en base al método de los elementos finitos*. Trabajo Especial. U.C.V. Facultad de Ingeniería. Escuela de Ingeniería Mecánica. 1993.
- [14] Greenbaum, A., *Iterative methods for solving linear systems*, SIAM, philadelphia, 1997.



- [15] Hestenes, M. R. & Stiefel, E., *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards, 49, 409-436, 1952.
- [16] Hughes, T. J. R, Levit, I. and Windget, J., *An element-by-element solution algorithm for problems of structural and solid mechanics*, Computational Methods in Applied Mechanics and Engineering, 36, 241-254, 1983.
- [17] *Introduction to finite element methods (ASEN 5007)-Fall 2002*, Course Material, Department of Aerospace Engineering Sciences University of Colorado at Boulder, Disponible: <http://titan.colorado.edu/courses.d/IFEM.d/>, 2003,05,28.
- [18] Jennings, A. & Malik, G. M., *The solution of sparse linear equations by the conjugate gradient method*, International Journal for Numerical Methods in Engineering, 12, 141-158, 1978.
- [19] Kernighan, B. W. & Ritchie, D. M., *El lenguaje de programación C*, 5<sup>ta</sup> Edición, Prentice Hall, México, 1988.
- [20] Klisinski, M. & Runesson, K., *Improved symmetric and non-symmetric solvers for FE calculations*, Advances in Engineering Software, 18, 41-51, 1993.
- [21] Lanczos, C., *The Variational Principles of Mechanics*, Univ. of Toronto Press, 1949.
- [22] Langhaar, H. L. *Energy Methods in Applied Mechanics*, John Wiley & Sons, New York, 1962.

- [23] Lu, M., Connell, M. & Yullberg, O., *Firsts steps towards applying vr in interactive Finite Element analysis*, Lecture Notes, Department of Structural Mechanics Chalmers University of Technology, Gothenburg, 10.
- [24] Moaveni, S., *Finite element analysis theory and application with ANSYS*, Prentice Hall, New Jersey, 1999.
- [25] Montero, G., Montenegro, R., Escobar, J. M. & Rodríguez, E., *Resolución de sistemas de ecuaciones lineales tipo Sparse: La estrategia RPK*, Instituto de Sistemas Inteligentes y Aplicaciones Numéricas en Ingeniería, 2003.
- [26] Morales E. Esther E. Y Sánchez V. Rafael J. *Implementación de un modelo de cálculo para fenómenos de transferencia de calor mediante el método de los elementos finitos y aplicación al cálculo de esfuerzos térmicos a un órgano dentario*. Trabajo Especial. U.C.V. Facultad de Ingeniería. Escuela de Ingeniería Mecánica. 1989.
- [27] Nikishkov, G. P., *Introduction to the Finite Element Method*, Lecture Notes, University of Aizu, 32, 1998.
- [28] Ordaz G. Elizabeth del Valle y Pulido R. Olga María. *Resolución de problemas geométricos asociados a la generación automática de mallas tridimensionales de elementos finitos en piezas de forma libre*. Trabajo Especial. U.C.V. Facultad de Ingeniería. Escuela de Ingeniería Mecánica. 1992.
- [29] Ortiz, M., Pinsky, P.M. y Taylor, R.L., *Unconditionally stable element by element algorithms for dynamic problems*. Computational Methods in Applied Mechanics and Engineering. 36:223-239, 1983.

- [30] Peterson, R. E., *Stress concentration factors, charts and relations useful in making strength calculations for machine parts and structural elements*, John Wiley ,New York, 1974.
- [31] Richards, T. H., *Energy methods in stress analysis*, John Wiley & Sons, New York, 1977.
- [32] Roark, R. J. & Young, W. C., *Formulas for stress and strain*, Fifth Edition, McGraw-Hill, London, 1975.
- [33] Rosales S. Luis A. *Desarrollo e implementación de un módulo de análisis no lineal por el método de los elementos finitos*. Trabajo Especial. U.C.V. Facultad de Ingeniería. Escuela de Ingeniería Mecánica. 1987.
- [34] Saad, Y., *Iterative Methods for sparse linear systems*, 2<sup>da</sup> Edición, PWS publishing, New York, 2000.
- [35] Sánchez, G., Vial, C. & Moraga, N., *Estudio de fluidos no newtonianos con los métodos de volúmenes y elementos finitos*, Revista Facultad de ingeniería, U.T.A., Chile, 10, 23-34, 2002.
- [36] Schwarz, H. R., *Numerical analysis of symmetric matrices*, Prentice Hall, New Jersey, 1973.
- [37] Shewchuk, J. R., *An introduction to the conjugate gradient method without the agonizing pain*, Lecture Notes, School of Computer Science Carnegie Mellon University, Pittsburgh, 58, 1994.

- [38] Smith, M. I., Griffiths, D. V., *Programming the finite element method*, 3<sup>ra</sup> Edición, John Wiley & Sons, New York, 1998.
- [39] Stokwell, A. E., *A verification procedure for MSC/NASTRAN finite elements models*, NASA Contractor Report 4675, 1995.
- [40] Taghavi, R., *Automatic, parallel and fault tolerant mesh generation from CAD*, Engineering with Computers, 12, 178-185, 1996.
- [41] Vignjevic, R., Morris, A. J. & Belagundu, A. D., *Towards high fidelity finite element analysis*, Advances in Engineering Software, 29, 655-665, 1998.
- [42] Wathen, A. J., *An analysis of some element-by-element techniques*, Computational Methods in Applied Mechanics and Engineering, 74, 271-287, 1989.
- [43] Zienkiewicz, O. C., Taylor, R. L., *El método de los elementos finitos*, 4<sup>ta</sup> Edición, MacGraw-Hill, Barcelona, 1994.
- [44] Zörner, T. H. & Plasmeijer, R., *Solving linear systems with functional programming languages*, Lecture Notes, Computing Science institute Catholic University of Nijmegen, The Netherlands, 20.