

Solución Numérica de Problemas de Control Óptimo Asociados a Procesos de Difusión y Difusión-Advección

Tesis

que para obtener el grado de Maestro en Matemáticas presenta

David Parra Guevara

bajo la dirección del Dr. Alfredo Nicolás Carrizosa

1995

Departamento de Matemáticas
Universidad Autónoma Metropolitana-Iztapalapa

Solución Numérica de Problemas de Control Optimo
Asociados a Procesos de Difusión y Difusión-Advección

Por
David Parra Guevara

Tesis de Maestría

Deseo manifestar mi agradecimiento a mi asesor, Dr. Alfredo Nicolás Carrizosa, por sus conocimientos, apoyo, tiempo y paciencia para elaborar este trabajo, así también agradezco al Dr. Eugenio Gómez Reyes y al M. en C. Gustavo Izquierdo Buenrostro, por la revisión y sugerencias para mejorar esta tesis.

Finalmente deseo expresar mi agradecimiento a CONACyT por su apoyo económico y al Departamento de Matemáticas de la UAM-I, especialmente a la Coordinadora de la Maestría en Matemáticas Dra. Patricia Saavedra Barrera por las facilidades que me brindó.

Indice

- Introducción	1
- Capítulo I Problemas de Control Optimo.	5
1.1 Problema Inverso como un problema de Control Optimo	5
1.2 Propiedades de la Funcional de Energía	9
1.3 Existencia y Unicidad del Control Optimo	12
1.4 Método de Solución	19
- Capítulo II Control en la Frontera tipo Dirichlet para la Ecuación de Calor.	23
2.1 Cálculo de la Ecuación Adjunta	24
2.2 Aplicación del Método de Gradiente Conjugado	26
2.3 Discretización del Problema Directo	30
2.4 Aspectos de Algebra Lineal Numérica en la Implementación de Gradiente Conjugado	34
2.5 Resultados Numéricos	41
- Capítulo III Control en la Frontera tipo Dirichlet para la Ecuación de Transporte.	52
3.1 Cálculo de la Ecuación Adjunta	53
3.2 Aplicación del Método de Gradiente Conjugado	56
3.3 Discretización del Problema Directo	59
3.4 Aspectos de Algebra Lineal Numérica en la Implementación de Gradiente Conjugado	62
3.5 Resultados Numéricos	67

- Capítulo IV Control en la Frontera Tipo Neumann para la Ecuación de Transporte	77
4.1 Cálculo de la Primera Variación	78
4.2 Cálculo de la Ecuación Adjunta	79
4.3 Aplicación del Método de Gradiente Conjugado	83
4.4 Discretización y Resultados Numéricos	86
- Conclusiones	100
- Apéndice A	103
- Apéndice B	104
- Bibliografía.	105

INTRODUCCION

Una de las actividades de la ciencia es el estudio de los fenómenos naturales o sistemas físicos, a través del análisis de modelos matemáticos.

Un modelo matemático es la simplificación de un sistema, de tal forma que se consideran las variables más importantes que influyen en éste y se relacionan siguiendo leyes físicas para obtener ecuaciones ya sean algebraicas, diferenciales ordinarias o parciales, por mencionar algunas.

La información que se obtiene de un modelo es una aproximación de la realidad, ya que en principio el modelo es una simplificación del sistema.

Se cree que al elaborar modelos que involucren un mayor número de variables del sistema la solución estará más cerca de la realidad, el problema aquí es que el modelo será más complejo y darle solución puede requerir un gran trabajo teórico y/o computacional. Un ejemplo de esto son las ecuaciones no estacionarias de Navier-Stokes que modelan la velocidad y presión de un fluido viscoso e incompresible, también se puede modelar la temperatura si las consideramos en el contexto de la aproximación de Boussinesq en fluidos no isotérmicos. Para dimensión tres este problema no está teóricamente resuelto para tiempos grandes desde el punto de vista de problema bien planteado: existencia de solución única y dependencia continua respecto de los datos; además la solución computacional con simplificaciones requiere de un gran trabajo.

Una forma de clasificar los modelos es en problemas *directos* (forward) y problemas *inversos*.

En un problema directo se quiere hallar la solución a un conjunto de ecuaciones con datos conocidos, es decir, se conoce la geometría, condiciones iniciales, condiciones a la frontera, valores de parámetros, etc. Por ejemplo sea $\Omega \subset R^2$ un dominio conexo, acotado y de frontera suave Γ , queremos hallar $u = u(x)$ con $x \in \Omega$, que sea solución del problema elíptico:

$$\begin{aligned} u - \Delta u &= 0 && \text{en } \Omega \\ \frac{\partial u}{\partial n} &= v && \text{sobre } \Gamma, \end{aligned}$$

donde v es conocida.

Este problema directo es un modelo normalizado para un proceso de difusión estacionario en un medio homogéneo isotrópico y absorbente, $u(x)$ representa la concentración (o temperatura) de sustancia en $x \in \Omega$.

En un problema inverso se tiene alguna información sobre la solución de las ecuaciones y lo que se busca es la condición sobre los datos para que la solución del problema directo tenga las características pedidas.

Para el ejemplo anterior un problema inverso es el siguiente, sea $\omega \subset \Omega$ un subconjunto abierto y conexo y $u_T = u_T(x)$ con $x \in \omega$. Se quiere hallar una función v definida en Γ de tal forma que la solución $u = u(x)$ del problema directo:

$$\begin{aligned} u - \Delta u &= 0 && \text{en } \Omega \\ \frac{\partial u}{\partial n} &= v && \text{sobre } \Gamma, \end{aligned}$$

satisfaga $u(x) = u_T(x)$ para cada $x \in \omega$.

Para este ejemplo la función u_T es una distribución que se desea alcanzar en un sub-dominio ω y para tal efecto se *controla* el flujo en la frontera, es decir, se busca v .

En el problema inverso la ecuación que modela el sistema recibe el nombre de *ecuación de estado*, la solución de la ecuación es la *variable de estado* y el dato buscado es el *control*. En estos términos un problema inverso consiste en hallar un control que permita que la variable de estado alcance una distribución deseada. Estos problemas inversos o de control se plantean como problemas de *control óptimo*, es decir, no solo se busca un control adecuado sino también que el control sea de mínima energía (norma).

En los últimos treinta años han cobrado gran importancia los problemas de control óptimo de sistemas gobernados por ecuaciones diferenciales parciales (sistemas distribuidos), pero es en la década de los ochentas y hasta la fecha cuando han aparecido más trabajos en relación con este tópico (solución numérica), en particular con ecuaciones parciales lineales y en especial con la ecuación de onda. Esto se debe a que el control de la ecuación de onda es una primera aproximación para estabilizar estructuras que presentan oscilaciones.

El primer trabajo relevante en Control Óptimo es el libro [1] de J.L. Lions, publicado en 1971. En éste se resuelven problemas de controlabilidad exacta de sistemas gobernados por ecuaciones elípticas, hiperbólicas y parabólicas, con condiciones de frontera Dirichlet y Neumann. Los problemas que se abordan son lineales y para estos se establecen desigualdades variacionales con la finalidad de demostrar existencia y unicidad del control óptimo en espacios

de Hilbert adecuados, sin embargo, no se abordan algoritmos numéricos para el cálculo explícito del control.

Posteriormente en 1986 J.L. Lions en su artículo: "Exact Controllability, Stabilization and Perturbations for Distributed Systems" [2], demuestra la existencia del control exacto, independientemente de si la dimensión espacial es par o impar, utiliza una técnica constructiva, la cual permite calcular el control numéricamente. El trabajo es bastante general y aborda en detalle el problema de controlabilidad exacta de la ecuación de onda con control Dirichlet y control Neumann.

Tomando como punto de partida el trabajo anterior de J.L. Lions, en 1987 R. Glowinski y Li [3], resuelven numéricamente el problema de controlabilidad exacta para la ecuación de onda en rectángulos, usando diferencias finitas. En 1988 A. Nicolás [4] resuelve el mismo problema en dominios arbitrarios usando elemento finito. Estos dos trabajos usan una regularización tipo Tichonoff, la cual consiste en la adición de un término biarmónico con el objetivo de obtener un problema discreto bien planteado. Posteriormente en 1989 R. Glowinski y Mary Wheeler resuelven sin regularización de Tichonoff la controlabilidad exacta para la ecuación de onda en rectángulos [5].

Por otra parte, aún cuando la controlabilidad exacta de la ecuación de calor también se aborda en el trabajo de Lions, a diferencia de la ecuación de onda el cálculo numérico del control no es claro, o al menos no fácil por los espacios involucrados y al hecho que la ecuación de calor, a diferencia de la ecuación de onda no modela un proceso reversible. Sin embargo, los primeros experimentos numéricos para la controlabilidad (desde el punto de vista de control óptimo) de problemas de difusión y difusión-advección (ecuación de calor y transporte), se deben a R. Glowinski y M. Berggren en 1992 con la Tesis: "Control and Simulation of Advection-Diffusion Systems" [6]. En este trabajo se busca un control óptimo aproximado de tipo Dirichlet en una sección de la frontera, el problema se plantea para rectángulos y la semidiscretización de los problemas directos se hace con un esquema de *separación de operadores*. Para el cálculo del control óptimo se aplica el método de *gradiente conjugado* y para la implementación numérica se usó el método de elemento finito.

En el trabajo que se presenta aquí, también se busca el control óptimo aproximado para procesos de difusión y difusión-advección sobre rectángulos, sólo que usamos una discretización implícita del operador espacial, mediante diferencias finitas.

Hay que considerar que para la solución de un problema inverso es necesario resolver el problema directo asociado, por lo tanto, se necesita de un método de solución *eficiente* para el problema directo. En los problemas que presentamos la geometría de los dominios es *rectangular*, por lo cual se han elegido para resolver los problemas directos métodos de *diferencias finitas* para discretizar en tiempo y espacio. La extensión a otras geometrías discretizando en espacio mediante elementos finitos, no es una complicación esencial. Además, tomando en cuenta que para datos suaves en rectángulos el problema aproximado en diferencias finitas coincide con el problema aproximado por elementos finitos para la triangulación regular de Friedrichs-Keller, entonces los resultados que se obtengan con diferencias finitas son una base fuerte para la validez con elementos finitos en regiones arbitrarias.

Por último señalamos que el artículo de 1994 de J.L. Lions y R. Glowinski: "Exact and Aproximate Controllability for Distributed Parameter Systems" [7], contiene métodos numéricos para el control puntual en problemas de difusión y en éste se anuncia la aparición en 1995 de un estudio mas completo teórico y computacional del control óptimo tipo Dirichlet y Neumann para las ecuaciones de difusión y de sistemas tipo Stokes.

En los capítulos II y III de este trabajo, buscamos un control tipo Dirichlet en una sección de la frontera y en el capítulo IV usamos un control Neumann, lo cual representa una aportación nueva a los trabajos en esta dirección.

CAPITULO I

1 Problemas de Control Optimo

1.1 Problema Inverso como un Problema de Control Optimo

Consideremos el siguiente problema directo:

Hallar $u : R^2 \times (0, T) \rightarrow R$ que satisfaga el problema parabólico

$$\frac{\partial u}{\partial t} - \Delta u = f \quad \text{en } \Omega \times (0, T) \quad (1.1)$$

con condición de frontera

$$au + \frac{\partial u}{\partial n} = ag \quad \text{en } \Gamma_1 \times (0, T) \quad (1.2)$$

$$u = v \quad \text{en } \Gamma_c \times (0, T) \quad (1.3)$$

y condición inicial

$$u(0) = u(x, y) \quad \text{en } \Omega \quad (1.4)$$

Donde Ω es un dominio acotado y conexo de R^2 con frontera Γ . La frontera se divide en dos secciones Γ_1 y Γ_c tales que $\Gamma_1 \cup \Gamma_c = \Gamma$ y estas solo se intersectan en dos puntos extremos.

La ecuación parabólica (1.1) conocida como la *ecuación de calor* es un modelo normalizado de la distribución de temperatura $u(x, y, t)$ en cada punto (x, y) de una lámina al tiempo t , dicha lámina se considera como un cuerpo homogéneo e isotrópico. La función f representa una fuente de densidad de calor sobre cada punto (x, y) de la lámina y puede variar con el tiempo.

El problema inverso asociado que nos interesa es el siguiente: Dada una distribución de temperatura $u_T(x, y)$ en la lámina, deseamos hallar la función v de temperatura sobre Γ_c adecuada para que la solución del problema directo (1.1,1.2,1.3,1.4) satisfaga al tiempo $t = T$, $u(x, y, T) = u_T(x, y)$ para cada $(x, y) \in \Omega$.

Es decir, asumimos que la evolución del sistema se da de acuerdo a la ecuación (1.1), el flujo de calor en la frontera Γ_1 está dado de acuerdo a (1.2) y además partimos de una distribución inicial conocida de temperatura

de acuerdo a (1.4). Lo que falta por determinar es la condición de frontera Dirichlet (1.3) para que el problema, así *cerrado*, tenga solución que al tiempo T se distribuya sobre la lámina Ω en la forma deseada $u_T(x, y)$.

Nos interesa controlar el sistema de forma óptima en el sentido que el control v calculado sea el que represente mínima energía (mínimo costo), es decir queremos que v minimice la funcional que calcula su energía

$$J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt \quad ;$$

esto es, desde el punto de vista matemático nos interesa el control de mínima norma en $L^2(\Gamma_c \times (0, T))$.

El problema así planteado es un problema clásico de teoría de control óptimo denominado *controlabilidad exacta*. El problema como está planteado puede no tener solución, es decir, es posible que para una distribución $u_T(x, y)$ deseada no exista un control v que haga que la solución del problema directo cumpla con $u(x, y, T) = u_T(x, y)$ y, por lo tanto, no hay un conjunto de controles admisibles sobre los cuales se pueda minimizar la funcional $J[v]$ que mide la energía, tal función u_T se denomina *distribución no-alcanzable* del sistema.

Para que el problema no sea tan estricto se plantea como de *control aproximado*, es decir, buscamos el control v que genere una solución u del problema directo que al tiempo T esté lo más próxima a la distribución deseada u_T . En este caso queremos minimizar

$$\int_{\Omega} (u_T - u(T))^2 dx dy$$

que es la norma en $L^2(\Omega)$, y mide el error de aproximación.

Como ambas funcionales a minimizar son no-negativas podemos minimizar la suma, es decir

$$J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u_T - u(T))^2 dx dy \quad \epsilon > 0 \quad (1.5)$$

donde ϵ es un parámetro de penalización. Cuando ϵ es pequeño, damos mayor peso a minimizar el error de aproximación en $L^2(\Omega)$, es decir, esperamos que cuando $\epsilon \rightarrow 0$ $u(T) \rightarrow u_T$.

La funcional (1.5) la llamaremos en lo sucesivo *funcional de energía*.

En conclusión el problema de control óptimo aproximado es

$$\text{Min } J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u_T - u(T))^2 dx dy, \quad \epsilon > 0 \quad (1.5)$$

sujeto a:

$$\frac{\partial u}{\partial t} - \Delta u = f \quad \text{en } \Omega \times (0, T) \quad (1.1)$$

$$au + \frac{\partial u}{\partial n} = ag \quad \text{en } \Gamma_1 \times (0, T) \quad (1.2)$$

$$u = v \quad \text{en } \Gamma_c \times (0, T) \quad (1.3)$$

$$u(0) = u_0(x, y) \quad \text{en } \Omega \quad (1.4)$$

Otro problema a tratar es el de control óptimo en un proceso de difusión-advección (el fenómeno de difusión está dado en términos matemáticos por Δu y el de advección por $\mathbf{w} \cdot \nabla u$), en este caso la ecuación de estado es conocida como la *ecuación de transporte*

$$\frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u = f, \quad \text{div}(\mathbf{w}) = 0$$

y es el modelo de la distribución de temperatura o concentración de contaminantes $u(x, y, t)$ en cada punto de un fluido incompresible ($\text{div}(\mathbf{w}) = 0$) de velocidad conocida \mathbf{w} , en el tiempo t .

También en este caso nos restringimos a un dominio acotado conexo Ω de R^2 con frontera Γ , la frontera se divide también en dos secciones Γ_1 y Γ_c tales que $\Gamma_1 \cup \Gamma_c = \Gamma$ y estas solo se tocan en dos puntos extremos.

Se desea controlar la temperatura del fluido a un tiempo $T > 0$ mediante el control de la temperatura en la sección de frontera Γ_c (control tipo Dirichlet) a mínima energía (costo) de forma aproximada.

El problema de control óptimo en este caso es

$$\text{Min } J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u_T - u(T))^2 dx dy, \quad \epsilon > 0$$

sujeto a:

$$\frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u = f, \quad \text{div}(\mathbf{w}) = 0 \quad \text{en } \Omega \times (0, T)$$

$$au + \beta \frac{\partial u}{\partial n} = ag \quad \text{en } \Gamma_1 \times (0, T)$$

$$u = v \quad \text{en } \Gamma_c \times (0, T)$$

$$u(0) = u_0(x, y) \quad \text{en } \Omega$$

Con la finalidad de desarrollar un algoritmo para resolver numéricamente ambos problemas de control óptimo (difusión y difusión-advección) necesitamos asumir las siguientes características. Los datos f, g y $u_0(x, y)$ se consideran funciones cuadrado integrables en su dominio, de hecho para los experimentos numéricos se consideran funciones de clase C^2 . La velocidad del fluido \mathbf{w} la consideramos constante, es decir $\mathbf{w} = (w_1, w_2)$ con $w_1, w_2 \in \mathbb{R}$, por último las constantes a y β son reales positivos.

Hacemos dos hipótesis sobre los problemas directos, la primera es que para cada $v \in L^2(\Gamma_c \times (0, T))$ existe $u = u(x, y, t)$, $(x, y) \in \Omega$ y $t > 0$ que es solución de la ecuación de difusión (difusión-advección) con las correspondientes condiciones de frontera e iniciales, esto es indispensable para que el espacio de controles admisibles en el problema de control óptimo sea no vacío. La segunda es la continuidad respecto del control, es decir si v_1 converge a v_2 entonces u_1 converge a u_2 , esto significa que para una perturbación pequeña δv de un control v se tiene una perturbación δu pequeña de la solución u del problema directo, es decir para una constante C positiva consideramos:

$$\|\delta u\|_{L^2(\Omega)} \leq C \|\delta v\|_{L^2(\Gamma_c \times (0, T))}.$$

El significado de la funcional $J[v]$ en ambos problemas (difusión y difusión-advección) es el mismo. Se puede observar que la ecuación de estado del problema de difusión se obtiene de la ecuación de difusión-advección con $\mathbf{w}=\mathbf{0}$; esto es, en el caso de que el fluido está en reposo.

1.2 Propiedades de la Funcional de Energía

La funcional de energía (1.5) es la funcional que se desea minimizar a fin de resolver los problemas de control óptimo. El espacio donde se desea minimizar es el espacio de Hilbert $L^2(\Gamma_c \times (0, T))$, esta elección es natural ya que para que la primera integral exista el control v debe pertenecer a este espacio.

Teorema 1.1 La funcional de energía es estrictamente convexa.

Antes de demostrar el teorema hay que observar que dado que las ecuaciones de estado en ambos casos (calor y transporte) son lineales y las condiciones a la frontera e iniciales también, entonces para dos controles diferentes v_1 y $v_2 \in L^2(\Gamma_c \times (0, T))$ con correspondientes soluciones u_1 y u_2 del problema directo, se tiene que al control $\lambda v_1 + \mu v_2$ le corresponde la solución del problema directo $\lambda u_1 + \mu u_2$.

Demostración

Tomamos v_1 y $v_2 \in L^2(\Gamma_c \times (0, T))$ con correspondientes soluciones del problema directo u_1 y u_2 , y $\lambda \in (0, 1)$. Considerando que $\lambda^2 < \lambda$, se establecen las desigualdades:

$$(\lambda v_1 + (1 - \lambda)v_2)^2 < \lambda v_1^2 + (1 - \lambda)v_2^2,$$

$$(\lambda u_1(T) + (1 - \lambda)u_2(T) - u_T)^2 < \lambda(u_1(T) - u_T)^2 + (1 - \lambda)(u_2(T) - u_T)^2$$

Integrando ambas desigualdades obtenemos:

$$\begin{aligned} \frac{1}{2} \int_0^T \int_{\Gamma_c} (\lambda v_1 + (1 - \lambda)v_2)^2 d\gamma dt &< \frac{\lambda}{2} \int_0^T \int_{\Gamma_c} v_1^2 d\gamma dt + \\ &\frac{(1 - \lambda)}{2} \int_0^T \int_{\Gamma_c} v_2^2 d\gamma dt \\ \frac{1}{2\epsilon} \int_{\Omega} (\lambda u_1(T) + (1 - \lambda)u_2(T) - u_T)^2 dx dy &< \frac{\lambda}{2\epsilon} \int_{\Omega} (u_1(T) - u_T)^2 dx dy + \\ &\frac{(1 - \lambda)}{2\epsilon} \int_{\Omega} (u_2(T) - u_T)^2 dx dy \end{aligned}$$

Sumando las dos ultimas desigualdades y agrupando terminos, obtenemos:

$$J[\lambda v_1 + (1 - \lambda)v_2] < \lambda J[v_1] + (1 - \lambda)J[v_2]$$

q.e.d.

Antes de probar la existencia y unicidad del mínimo de la funcional de energía, enunciamos algunos resultados generales sobre diferenciabilidad de funcionales (Vease por ejemplo [8]).

Definición. Sea H un espacio de Hilbert y $J : H \rightarrow R$ una funcional, se dice que la funcional es fuertemente diferenciable si:

$$J[v + \delta v] - J[v] = L[v; \delta v] + w(v; \delta v) \forall \delta v$$

donde δv es un incremento funcional, $L[v; \delta v]$ es lineal y acotada en δv y además

$$\frac{|w(v; \delta v)|}{\|\delta v\|} \rightarrow 0 \quad \text{si} \quad \|\delta v\| \rightarrow 0$$

L se llama la *primera variación ó diferencial* de $J[v]$ y se denota

$$\delta J[v; \delta v] = L[v; \delta v].$$

Como una consecuencia directa de la definición tenemos el siguiente teorema.

Teorema 1.2 Si la funcional $J[v]$ es fuertemente diferenciable entonces es continua y además la primera variación es única.

Definición. Sea H un espacio de Hilbert y $J : H \rightarrow R$ una funcional, se dice que la funcional es débilmente diferenciable si el límite:

$$\lim_{\lambda \rightarrow 0} \frac{J[v + \lambda \delta v] - J[v]}{\lambda}$$

existe como una funcional lineal acotada en δv . En caso de existir se denota como $DJ[v; \delta v]$.

Los siguientes teoremas relacionan los dos conceptos de diferenciabilidad definidos.

Teorema 1.3 Si $J[v]$ es fuertemente diferenciable entonces, es débilmente diferenciable y además las diferenciales son iguales, es decir:

$$\delta J[v; \delta v] = DJ[v; \delta v].$$

Teorema 1.4 Si $J[v]$ es débilmente diferenciable y además $DJ[v; \delta v]$ es continua en v entonces es fuertemente diferenciable.

El siguiente teorema es necesario para plantear un método numérico eficiente en los problemas de control óptimo definidos en la sección anterior.

Teorema 1.5 La funcional de energía (1.5) es fuertemente diferenciable con primera variación continua:

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} v \delta v \, d\gamma \, dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) \, dx \, dy \quad (1.6)$$

Demostración

Para demostrar esta propiedad hacemos uso del teorema (1.4), por lo tanto, calculamos primero la derivada débil, el calculo lo hacemos para la ecuación de transporte ya que es más general que la ecuación de calor ($\mathbf{w}=\mathbf{0}$). A una variación δv del control v le corresponde una variación δu de la solución del problema directo. Si δu se define como la solución del problema directo:

$$\frac{\partial \delta u}{\partial t} - \beta \Delta \delta u + \mathbf{w} \cdot \nabla \delta u = 0 \quad , \quad \text{div}(\mathbf{w}) = 0 \quad \text{en } \Omega \times (0, T)$$

$$a \delta u + \beta \frac{\partial \delta u}{\partial n} = 0 \quad \text{en } \Gamma_1 \times (0, T)$$

$$\delta u = \delta v \quad \text{en } \Gamma_c \times (0, T)$$

$$\delta u(0) = 0 \quad \text{en } \Omega$$

entonces por la linealidad de la ecuación de estado y condiciones de frontera e inicial al control $v + \lambda \delta v$ le corresponde la solución $u + \lambda \delta u$ del problema directo, evaluando en este control la funcional de energía, tenemos

$$J[v + \lambda \delta v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} (v + \lambda \delta v)^2 \, d\gamma \, dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) + \lambda \delta u(T) - u_T)^2 \, dx \, dy$$

si desarrollamos los términos al cuadrado y restamos $J[v]$ obtenemos:

$$\frac{J[v + \lambda \delta v] - J[v]}{\lambda} = \int_0^T \int_{\Gamma_c} v \delta v \, d\gamma \, dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) \, dx \, dy +$$

$$\frac{\lambda}{2} \int_0^T \int_{\Gamma_c} \delta^2 v \, d\gamma \, dt + \frac{\lambda}{2\epsilon} \int_{\Omega} \delta^2 u(T) \, dx \, dy$$

tomando en ambos lados de la igualdad $\lim_{\lambda \rightarrow 0}$, tenemos

$$DJ[v; \delta v] = \int_0^T \int_{\Gamma_c} v \delta v \, d\gamma \, dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) \, dx \, dy$$

por lo tanto, la funcional de energía es débilmente diferenciable y se puede expresar la diferencial como:

$$DJ[v; \delta v] = (v, \delta v)_{L^2(\Gamma_c \times (0, T))} + \frac{1}{\epsilon} (u(T) - u_T, \delta u(T))_{L^2(\Omega)}$$

esta expresión es útil para mostrar la continuidad. En lo que sigue se debe entender que la norma para los terminos de u y δu es la norma en $L^2(\Omega)$ y para los terminos en v y δv la norma usada es la que corresponde al espacio $L^2(\Gamma_c \times (0, T))$.

Sean v_1 y $v_2 \in L^2(\Gamma_c \times (0, T))$ con correspondientes soluciones del problema directo u_1 y u_2 , entonces usando la desigualdad de Cauchy tenemos

$$|DJ[v_1; \delta v] - DJ[v_2; \delta v]| \leq \|v_1 - v_2\| \|\delta v\| + \frac{1}{\epsilon} \|u_1(T) - u_2(T)\| \|\delta u(T)\|$$

ahora si $v_1 \rightarrow v_2$ entonces, $u_1 \rightarrow u_2$ por la continuidad de la solución respecto de las condiciones de frontera para la ecuación de calor y transporte, esto asegura que

$$|DJ[v_1; \delta v] - DJ[v_2; \delta v]| \rightarrow 0,$$

por lo tanto $DJ[v; \delta v]$ es continua en v . Un argumento similar muestra que $DJ[v; \delta v]$ es continua en δv y por lo tanto es una funcional lineal acotada.

Por el teorema 1.4 se concluye que la funcional de energía es fuertemente diferenciable.

q.e.d.

1.3 Existencia y Unicidad del Control Optimo.

En lo que sigue se establece la existencia y unicidad del mínimo de la funcional de energía sobre el espacio $L^2(\Gamma_c \times (0, T))$, para esto usamos los siguientes teoremas.

Teorema 1.6 Sea H un espacio de Hilbert y $J : H \rightarrow \mathbb{R}$ una funcional.

a) si $J[v]$ es convexa entonces todo mínimo local es un mínimo global.

b) si $J[v]$ es estrictamente convexa entonces el mínimo global es único (si existe).

Demostración

Sea $u \in H$ un mínimo local, entonces existe $\epsilon > 0$ tal que $\forall v$ que cumpla $\|v - u\| < \epsilon$ se tiene que $J[u] \leq J[v]$. Supongamos que no es mínimo global, entonces existe $w \in H$ tal que $J[w] < J[u]$. Como $J[v]$ es convexa al tomar $\lambda \in [0, 1]$ se cumple:

$$J[\lambda u + (1 - \lambda)w] \leq \lambda J[u] + (1 - \lambda)J[w] < \lambda J[u] + (1 - \lambda)J[u] = J[u]$$

definimos $v = \lambda u + (1 - \lambda)w$ donde $\lambda = 1 - \epsilon/2\|u - w\|$, con esta elección $\lambda \in [0, 1]$ y por lo tanto, $J[v] < J[u]$, sin embargo

$$\|v - u\| = (1 - \lambda)\|u - w\| = \frac{\epsilon}{2} < \epsilon,$$

lo cual contradice que u es mínimo local, por lo tanto, u tiene que ser mínimo global.

Para demostrar la unicidad supongamos que v_1 y v_2 son mínimos globales diferentes, es decir $J[v_1] = J[v_2] \leq J[v] \quad \forall v \in H$. Ahora si tomamos $\lambda \in (0, 1)$ fijo y considerando que $J[v]$ es estrictamente convexa tenemos

$$J[\lambda v_1 + (1 - \lambda)v_2] < \lambda J[v_1] + (1 - \lambda)J[v_2] = \lambda J[v_1] + (1 - \lambda)J[v_1] = J[v_1]$$

lo cual contradice que v_1 es mínimo global, por lo tanto $v_1 = v_2$.

q.e.d.

Los siguientes dos teoremas dan condiciones necesarias y suficientes sobre el mínimo, de hecho el segundo teorema se utiliza para justificar un criterio de paro en el algoritmo que desarrollaremos en el siguiente capítulo para calcular el control óptimo (mínimo).

Teorema 1.6.1 (Condición necesaria) Sea H espacio de Hilbert y $J : H \rightarrow \mathbb{R}$ una funcional fuertemente diferenciable. Si v^* es mínimo entonces $\delta J[v^*; \delta v] = 0 \quad \forall \delta v$.

Demostración

Definimos $f(\lambda) = J[v^* + \lambda\delta v]$, calculamos el cociente

$$\frac{f(\lambda + h) - f(\lambda)}{h} = \frac{J[v^* + \lambda\delta v + h\delta v] - J[v^* + \lambda\delta v]}{h},$$

tomando límite cuando $h \rightarrow 0$ en ambos lados obtenemos

$$f'(\lambda) = \delta J[v^* + \lambda\delta v; \delta v],$$

es decir, f es derivable $\forall \lambda$, de hecho

$$f'(0) = \delta J[v^*; \delta v].$$

Por otra parte, la función f tiene un mínimo en $\lambda = 0$, ya que

$$f(0) = J[v^*] \leq J[v^* + \lambda\delta v] = f(\lambda) \quad \forall \lambda$$

entonces por cálculo en una variable $f'(0) = 0$, es decir, $\delta J[v^*; \delta v] = 0 \quad \forall \delta v$ ya que δv se tomó arbitrario.

q.e.d.

Teorema 1.6.2 (Condiciones suficientes) Sea H espacio de Hilbert y $J : H \rightarrow \mathbb{R}$ una funcional estrictamente convexa, con primera variación continua $\delta J[v; \delta v]$. Si existe $v^* \in H$ tal que

$$\delta J[v^*; \delta v] = 0 \quad \forall \delta v$$

entonces v^* es el mínimo global único de $J[v]$.

Demostración

Sea $z \in H$ cualquiera, definimos $w = z - v^*$ y $f(\lambda) = J[v^* + \lambda w]$, calculamos el cociente:

$$\frac{f(\lambda + h) - f(\lambda)}{h} = \frac{J[v^* + \lambda w + h w] - J[v^* + \lambda w]}{h},$$

tomando límite cuando $h \rightarrow 0$ en ambos lados obtenemos

$$f'(\lambda) = \delta J[v^* + \lambda w; w],$$

es decir, f es derivable con continuidad, de hecho,

$$f'(0) = \delta J[v^*; w] = 0.$$

Por otra parte f es estrictamente convexa en R .

Sean $\lambda_1, \lambda_2 \in R$ y $t \in (0, 1)$,

$$f(t\lambda_1 + (1-t)\lambda_2) = J[v^* + (t\lambda_1 + (1-t)\lambda_2)w] =$$

$$J[t(v^* + \lambda_1 w) + (1-t)(v^* + \lambda_2 w)] < tJ[v^* + \lambda_1 w] + (1-t)J[v^* + \lambda_2 w]$$

$$\Rightarrow f(t\lambda_1 + (1-t)\lambda_2) < tf(\lambda_1) + (1-t)f(\lambda_2).$$

Ahora siendo f clase $C^1(R)$ y $f'(0) = 0$, entonces $\lambda = 0$ es máximo, mínimo ó punto de inflexión. Si $\lambda = 0$ fuera máximo ó punto de inflexión, existiría un intervalo donde f fuera concava, lo cual es imposible ya que f es estrictamente convexa en R . Por lo anterior $\lambda = 0$ es un mínimo local. Por el teorema 1.6, $\lambda = 0$ es mínimo global único; en particular

$$J[v^*] = f(0) < f(1) = J[z]$$

y como z es arbitrario tenemos que v^* es el mínimo global único de $J[v]$ en H .

q.e.d.

Definición. Sea $J[v]$ la funcional de energía, definimos

$$J^* = \inf\{J[v] | v \in L^2(\Gamma_c \times (0, T))\},$$

$$C_\mu = \{v \in L^2(\Gamma_c \times (0, T)) | J[v] \leq J^* + \mu\} \quad , \quad \mu > 0$$

El número real J^* existe ya que $J[v] \geq 0 \quad \forall v$, y C_μ es no vacío para cada $\mu \geq 0$ por definición del ínfimo.

Lema 1.1 C_μ es un conjunto convexo cerrado y acotado en el espacio de Hilbert $L^2(\Gamma_c \times (0, T))$

Demostración

Sean v_1 y $v_2 \in C_\mu$ y $\lambda \in [0, 1]$, entonces

$$J[\lambda v_1 + (1 - \lambda)v_2] \leq \lambda J[v_1] + (1 - \lambda)J[v_2] \leq$$

$$\lambda(J^* + \mu) + (1 - \lambda)(J^* + \mu) = J^* + \mu$$

por lo tanto, $\lambda v_1 + (1 - \lambda)v_2 \in C_\mu$. Esto prueba la convexidad.

Sea $\{v_n\}_{n=1}^\infty$ una sucesión en C_μ convergente en $L^2(\Gamma_c \times (0, T))$ y definimos z como su límite, es decir $\lim_{n \rightarrow \infty} v_n = z$. Como $J[v]$ es continua entonces $J[v_n] \rightarrow J[z]$, por lo tanto, $\forall \epsilon > 0$ existe N tal que $|J[v_N] - J[z]| < \epsilon$ o en forma equivalente:

$$J[z] < \epsilon + J[v_N] \leq \epsilon + (J^* + \mu) \quad \forall \epsilon > 0$$

esto implica que $J[z] \leq J^* + \mu$ y por lo tanto, $z \in C_\mu$. Esto prueba que C_μ es cerrado.

La funcional de energía se puede escribir como:

$$J[v] = \frac{1}{2} \|v\|_{L^2(\Gamma_c \times (0, T))}^2 + \frac{1}{2\epsilon} \|u(T) - u_T\|_{L^2(\Omega)}^2 \geq \frac{1}{2} \|v\|_{L^2(\Gamma_c \times (0, T))}^2 ;$$

con esta desigualdad y usando la definición de C_μ obtenemos:

$$\|v\|^2 \leq 2(J^* + \mu) \quad \forall v \in C_\mu ,$$

por lo cual concluimos que C_μ es acotado.

q.e.d.

Si el mínimo global v^* de la funcional de energía existe entonces $v^* \in C_\mu$ para $\mu > 0$ fijo, entonces vamos a minimizar a $J[v]$ restringiéndonos a este tipo de conjuntos con propiedades importantes.

Definición. Sea H un espacio de Hilbert, H^* su dual y $J : H \rightarrow R$ una funcional, definimos varios conceptos:

a) Una sucesión $\{u_n\}_{n=1}^\infty$ en H es débilmente convergente si existe $u \in H$, tal que $L(u_n) \rightarrow L(u) \quad \forall L \in H^*$.

b) Un conjunto $C \subset H$ es débilmente cerrado, si para toda sucesión $\{v_n\}_{n=1}^\infty$ débilmente convergente de elementos en C , se tiene que el límite v de la sucesión pertenece a C .

c) Un conjunto $C \subset H$ es débilmente compacto, si para toda sucesión $\{v_n\}_{n=1}^{\infty}$ de elementos en C existe una subsucesión $\{v_{n_k}\}_{k=1}^{\infty}$ que converge débilmente a un elemento de C .

d) Una funcional $J[v]$ es semicontinua inferior débilmente (s.i.d.) en $u \in H$ si para cada sucesión $\{v_n\}_{n=1}^{\infty}$ que converge débilmente a u se tiene que $J[u] \leq \liminf_{n \rightarrow \infty} J[v_n]$.

Teorema 1.7 Sea H espacio de Hilbert y $C \subset H$ cerrado y convexo, entonces C es débilmente cerrado.

Demostración

Sea $\{v_n\}_{n=1}^{\infty}$ una sucesión débilmente convergente a v , supongamos que v no está en C , entonces por el Teorema de la proyección (ver Apéndice) existe $p_v \in C$ tal que $(v - p_v, v_n - p_v) \leq 0 \quad \forall n$, con esto definimos $L \in H^*$ como $L(u) = (v - p_v, u)$ que cumple

$$\begin{aligned} L(v_n) &= (v - p_v, v_n) = (v - p_v, v_n - p_v) + (v - p_v, p_v) \leq (v - p_v, p_v) \\ &\Rightarrow L(v_n) \leq L(p_v) \quad \forall n, \end{aligned}$$

también cumple

$$\begin{aligned} L(v) &= (v - p_v, v) = (v - p_v, v - p_v) + (v - p_v, p_v) = \|v - p_v\|^2 + L(p_v) \\ &\Rightarrow L(v) > L(p_v) \end{aligned}$$

por lo tanto, $L(v_n) \leq L(p_v) < L(v) \quad \forall n$, lo cual es una contradicción ya que $L(v_n) \rightarrow L(v)$, entonces $v \in C$.

q.e.d.

Teorema 1.8 Sea H espacio de Hilbert y $C \subset H$ cerrado, acotado y convexo, entonces C es débilmente compacto.

Demostración

Como C es un conjunto acotado, existe $N > 0$ tal que $\frac{1}{N}C \subset B$, donde B es la bola unitaria cerrada. Por el Teorema de Banach-Alaoglu (ver Apéndice) B es débilmente compacto. Sea $\{x_n\}$ una sucesión en C , entonces $\{\frac{1}{N}x_n\}$ es sucesión en B , por lo tanto, existe una subsucesión $\{\frac{1}{N}x_{n_j}\}$ que converge débilmente a $y \in B$. Por el teorema (1.7) $\frac{1}{N}C$ es débilmente cerrado y por lo tanto $y \in \frac{1}{N}C$, podemos escribir $y = \frac{1}{N}x$ para algún $x \in C$. Multiplicando los elementos por N concluimos que $\{x_{n_j}\}$ converge débilmente a $x \in C$.

q.e.d.

Lema 1.2 La funcional de energía (1.5) es semicontinua inferior débilmente (s.i.d.) en el espacio de Hilbert $H = L^2(\Gamma_c \times (0, T))$.

Demostración

Sea $v \in H$ y supongamos que no es s.i.d., entonces existe una sucesión $\{v_n\}_{n=1}^{\infty}$ de elementos en H que converge débilmente a v pero $\liminf_{n \rightarrow \infty} J[v_n] < J[v]$. Como J^* es ínfimo, entonces en particular

$$J^* \leq \liminf_{n \rightarrow \infty} J[v_n] < J[v]$$

podemos elegir $\mu > 0$ para que se cumpla

$$\liminf_{n \rightarrow \infty} J[v_n] < J^* + \mu < J[v] \Rightarrow \inf_{k \geq n} J[v_k] < J^* + \mu \quad \forall n,$$

entonces existe una subsucesión $\{v_{n_j}\}_{j=1}^{\infty}$ tal que $J[v_{n_j}] \leq J^* + \mu$, es decir, $v_{n_j} \in C_\mu \quad \forall j$, ya que C_μ es cerrado y convexo, el teorema (1.7) garantiza que es débilmente cerrado y como la subsucesión $\{v_{n_j}\}_{j=1}^{\infty}$ converge débilmente a v , entonces se concluye que $v \in C_\mu$, es decir $J[v] \leq J^* + \mu$, lo cual contradice la elección de μ . Esto prueba que la funcional de energía es s.i.d.

q.e.d.

Teorema 1.9 La funcional de energía (1.5) tiene un mínimo único en el espacio de Hilbert $H = L^2(\Gamma_c \times (0, T))$.

Demostración

La unicidad fue probada en el teorema 1.6, por lo tanto, nos concentramos en la existencia.

Consideremos un conjunto C_μ con $\mu > 0$ fijo. Construimos la sucesión $\{v_n\}_{n=1}^{\infty}$ en C_μ en la forma:

$$J[v_n] < J^* + \frac{\mu}{n} \leq J^* + \mu \quad \forall n$$

esto es posible ya que J^* es un ínfimo. Esta construcción garantiza que $J[v_n] \rightarrow J^*$.

C_μ es cerrado, acotado y convexo, entonces por el teorema 1.8 es débilmente compacto, entonces existe una subsucesión $\{v_{n_j}\}_{j=1}^{\infty}$ que converge débilmente a un elemento $v^* \in C_\mu$, como $J[v]$ es s.i.d., entonces $J[v^*] \leq \liminf_{j \rightarrow \infty} J[v_{n_j}] = \lim_{n \rightarrow \infty} J[v_n] = J^*$, ya que J^* es el ínfimo concluimos que $J[v^*] = J^*$ y por lo tanto, v^* es el mínimo.

q.e.d.

Para un tratamiento mas general de minimización de funcionales convexas sobre espacios de Hilbert ver [9].

1.4 Método de Solución.

Como se mostró en la sección 1.3, el control óptimo existe y es único para cada $\epsilon > 0$ fijo, es decir, la funcional de energía:

$$J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u_T - u(T))^2 dx dy$$

tiene un mínimo global sobre $H = L^2(\Gamma_c \times (0, T))$ que depende de ϵ . A continuación justificaremos la introducción del parámetro ϵ , cuyo manejo será importante para obtener mejores aproximaciones al estado $u_T(x, y)$.

Cuando planteamos el problema inverso como un problema de control aproximado, tenemos en mente que el control óptimo será, en este caso, de menor energía (menor gasto) que en el caso de controlabilidad exacta, la razón es la siguiente:

Sea $\Sigma = \{v \in H | u(x, y, T) = u_T(x, y), (x, y) \in \Omega\}$, es decir, Σ es el conjunto de controles para los cuales la solución del problema directo se distribuye al tiempo $T > 0$ como $u_T(x, y)$. En este caso el problema de control óptimo (exacto) es:

$$\text{Min}_{v \in \Sigma} J_e[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt$$

Este problema tiene solución única, y los argumentos son los mismos que los descritos en la sección 1.2 y 1.3 para el problema de control aproximado, sólo que hay que aplicar los argumentos en el conjunto Σ , que es cerrado y convexo, por lo cual no hay variantes. (Suponemos que $\Sigma \neq \emptyset$).

Sea v_e el control óptimo exacto y v^* el control óptimo aproximado, entonces tomando en cuenta que $\Sigma \subset H$ y $J_e[v] = J[v] \quad \forall v \in \Sigma$, tenemos

$$J_e[v_e] = \inf_{v \in \Sigma} J_e[v] = \inf_{v \in \Sigma} J[v] \geq \inf_{v \in H} J[v] = J[v^*],$$

podemos escribir en términos de integrales:

$$\frac{1}{2} \int_0^T \int_{\Gamma_c} v^{*2} d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u_T - u^*(T))^2 dx dy \leq \frac{1}{2} \int_0^T \int_{\Gamma_c} v_e^2 d\gamma dt,$$

ya que la segunda integral es no negativa, obtenemos

$$\int_0^T \int_{\Gamma_\epsilon} v_*^2 d\gamma dt \leq \int_0^T \int_{\Gamma_\epsilon} v_\epsilon^2 d\gamma dt,$$

esto demuestra que el control óptimo aproximado es de menor energía que el control óptimo exacto para cualquier $\epsilon > 0$. Este hecho fundamental nos permite centrar nuestra atención únicamente en el error de aproximación del estado que se desea alcanzar $u_T(x, y)$, ya que el gasto de energía a quedado acotado. Esta es la razón de usar un parámetro ϵ , que para valores pequeños da más peso a la integral

$$\int_{\Omega} (u(T) - u_T)^2 dx dy,$$

de esta forma esperamos que al minimizar la funcional de energía $J[v]$, estemos minimizando el error de aproximación al estado u_T que deseamos alcanzar.

En el siguiente ejemplo, para una función de una variable, hacemos ver la influencia que esperamos tenga el parámetro ϵ , en la funcional de energía.

$$f(x) = x^2 + \frac{1}{\epsilon}(x - \mu)^2 \quad \epsilon > 0, \quad x \geq 0 \text{ y } \mu > 0,$$

calculamos el mínimo para esta función:

$$f'(x) = 2x + \frac{2}{\epsilon}(x - \mu) = 0 \iff x = \frac{\mu}{1 + \epsilon}$$

$$f''(x) = 2 + \frac{2}{\epsilon} > 0 \quad \forall x \Rightarrow x_{min} = \frac{\mu}{1 + \epsilon},$$

si $\epsilon \rightarrow 0$ entonces $x_{min} \rightarrow \mu$, es decir, que para valores de ϵ pequeños, el mínimo de la función tiende al mínimo de la parte dominante o de mayor peso de la función $f(x)$, que en este caso es $(x - \mu)^2$.

Por lo anteriormente descrito, nuestro problema es minimizar la funcional de energía $J[v]$, para valores pequeños de ϵ . El método que usaremos se conoce como gradiente conjugado. La convergencia y el algoritmo se establecen en el siguiente teorema.

Teorema 1.10 Sea H un espacio de Hilbert con producto escalar (\cdot, \cdot) y norma inducida $\|v\| = (v, v)^{\frac{1}{2}}$, $v \in H$. Sea $J : H \rightarrow R$ una funcional conveza, continuamente diferenciable con mínimo único, entonces el problema:

$$\text{Min}J[v]$$

$$v \in H,$$

se aproxima generando con el siguiente algoritmo una sucesión $\{v_n\}$ que converge en H al óptimo v^* , para cualquier condición inicial v^0 que se tome.

Algoritmo de gradiente conjugado

1) Hallar $g^0 \in H$ tal que:

$$(g^0, \delta v) = \delta J[v^0; \delta v] \quad \forall \delta v \in H;$$

2) $w^0 := g^0$.

Para $n \geq 0$ hallamos g^{n+1} , w^{n+1} y v^{n+1} como sigue:

3) Hallar $\rho_n \in R$ tal que:

$$J[v^n - \rho_n w^n] \leq J[v^n - \rho w^n] \quad \forall \rho \in R;$$

4) $v^{n+1} := v^n - \rho_n w^n$;

5) Hallar $g^{n+1} \in H$ tal que

$$(g^{n+1}, \delta v) = \delta J[v^{n+1}; \delta v] \quad \forall \delta v \in H;$$

si

$$\frac{\|g^{n+1}\|}{\|g^0\|} \leq Tol,$$

entonces $v^* := v^{n+1}$,

de otra forma:

$$6) \quad \gamma = \frac{(g^{n+1}, g^{n+1})}{(g^n, g^n)};$$

7) $w^{n+1} := g^{n+1} + \gamma w^n$;

8) Tomar $n \leftarrow n + 1$ y regresar a 3).

La aplicación de este algoritmo a la funcional de energía es posible ya que cumple con las hipótesis del teorema 1.10, las cuales se demostraron en la sección 1.2.

Ver [6] para aplicaciones de este algoritmo con una discretización en espacio a través del método de Elemento Finito. En este trabajo la discretización y aplicación del método se lleva a cabo con Diferencias Finitas.

CAPITULO II

2 Control en la Frontera tipo Dirichlet para la Ecuación de Calor

En este capítulo se calcula el control óptimo tipo Dirichlet (mínimo de la funcional de energía) para la ecuación de calor a través del método de gradiente conjugado, mostramos la discretización correspondiente del algoritmo y finalmente resultados numéricos de algunos ejemplos.

Especificamos en la *frontera* porque no es ésta la única forma de intentar controlar un sistema físico, en el caso del control de la temperatura de un cuerpo, parece congruente intentar un control sobre la frontera o una parte de esta, mientras que en modelos ambientales sería mejor un control puntual dentro del dominio del sistema. Un control tipo Dirichlet varía la temperatura a lo largo del tiempo sobre una parte de la frontera para influir dentro de el dominio, mientras que un control Neumann mantendría el flujo de calor a lo largo del tiempo en la frontera para lograr el mismo efecto, en el capítulo IV se trabaja con un control Neumann para la ecuación de transporte.

El problema por resolver y al cual nos referiremos en todo el capítulo es el siguiente

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy, \quad \epsilon > 0,$$

para $v \in L^2[\Gamma_c \times (0, T)]$,
sujeto a:

$$\left. \begin{array}{lll} \frac{\partial u}{\partial t} - \Delta u = f & \text{en} & \Omega \times (0, T) \\ au + \frac{\partial u}{\partial n} = ag & \text{en} & \Gamma_1 \times (0, T) \\ u = v & \text{en} & \Gamma_c \times (0, T) \\ u(0) = u_0(x, y) & \text{en} & \Omega \end{array} \right\} \quad (2.1)$$

En todo lo que sigue \mathbb{H} denota al espacio de Hilbert $L^2[\Gamma_c \times (0, T)]$.

2.1 Cálculo de la Ecuación Adjunta

En el algoritmo de gradiente conjugado, es necesario utilizar la primera variación de la funcional de energía en los pasos 1) y 5) (ver sección 1.4), en tales intrucciones buscamos un vector $g^k \in \mathbb{H}$ tal que

$$(g^k, \delta v) = \delta J[v^k; \delta v] \quad \forall \delta v \in \mathbb{H}$$

este vector g^k llamado gradiente de la funcional existe de forma única ya que por el teorema de representación de Riez la funcional lineal acotada $\delta J[v; \delta v]$ puede expresarse con el producto interior de \mathbb{H} de forma única.

El cálculo de este vector g^k es indispensable pues determina (salvo conjugación) la dirección de máximo descenso, que es la técnica para minimizar la funcional de energía, para facilitar el cálculo de g^k vamos a modificar la expresión de la primera variación para que solo tenga un producto interior.

Según lo calculado en el capítulo I (ecuación 1.6) la primera variación es:

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} v \delta v d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) dx dy$$

donde δu es la variación de la solución u del problema directo (2.1), que corresponde a la variación δv del control v , δu se calcula por la linealidad de la ecuación de estado y condiciones de frontera e iniciales como solución del problema directo

$$\frac{\partial \delta u}{\partial t} - \Delta \delta u = 0 \quad \Omega \times (0, T) \quad (2.2)$$

$$a \delta u + \frac{\partial \delta u}{\partial n} = 0 \quad \Gamma_1 \times (0, T) \quad (2.3)$$

$$\delta u = \delta v \quad \Gamma_c \times (0, T) \quad (2.4)$$

$$\delta u(0) = 0 \quad \Omega \quad (2.5)$$

Ahora introducimos una función $P = P(x, y, t)$ con suficiente suavidad cuyas características se determinan con el siguiente desarrollo:

Multiplicando (2.2) por P e integrando, obtenemos

$$\int_0^T \int_{\Omega} P \frac{\partial \delta u}{\partial t} - \int_0^T \int_{\Omega} P \Delta \delta u = 0,$$

integraremos por partes en tiempo la primera integral y aplicamos la segunda fórmula de Green en la segunda (ver apéndice) para obtener:

$$\int_{\Omega} \left[P \delta u \Big|_0^T - \int_0^T \delta u \frac{\partial P}{\partial t} \right] - \int_0^T \left[\int_{\Omega} \delta u \Delta P + \int_{\Gamma} P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right] = 0,$$

agrupando términos y considerando (2.5) tenemos

$$\int_{\Omega} P(T) \delta u(T) + \int_0^T \int_{\Omega} \delta u \left(-\frac{\partial P}{\partial t} - \Delta P \right) - \int_0^T \int_{\Gamma} \left[P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right] = 0,$$

pedimos a P que cumpla $-\frac{\partial P}{\partial t} - \Delta P = 0$ en $\Omega \times (0, T)$, entonces la segunda integral se hace cero, separamos la tercera integral en las secciones de frontera Γ_c y Γ_1 .

$$\int_{\Omega} P(T) \delta u(T) - \int_0^T \int_{\Gamma_c} P \frac{\partial \delta u}{\partial n} - \int_0^T \int_{\Gamma_1} P \frac{\partial \delta u}{\partial n} + \int_0^T \int_{\Gamma_c} \delta u \frac{\partial P}{\partial n} + \int_0^T \int_{\Gamma_1} \delta u \frac{\partial P}{\partial n} = 0$$

pedimos que $P = 0$ en $\Gamma_c \times (0, T)$, con esto la segunda integral es cero. Usamos (2.3) en la tercera integral y (2.4) en la cuarta, obtenemos

$$\int_{\Omega} P(T) \delta u(T) + \int_0^T \int_{\Gamma_1} \delta u \left(aP + \frac{\partial P}{\partial n} \right) + \int_0^T \int_{\Gamma_c} \delta v \frac{\partial P}{\partial n} = 0,$$

pedimos a P que cumpla $aP + \frac{\partial P}{\partial n} = 0$ en $\Gamma_1 \times (0, T)$, de esta forma la segunda integral es cero, también requerimos que $P(T) = \frac{1}{\epsilon} (u(T) - u_T)$ en Ω , entonces

$$\int_0^T \int_{\Gamma_c} \delta v \frac{\partial P}{\partial n} d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) dx dy = 0,$$

restando esta última ecuación de (1.6, sección 1.2) obtenemos

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} \left(v - \frac{\partial P}{\partial n} \right) \delta v d\gamma dt, \quad (2.6)$$

que es la formulación adecuada de la primera variación para usar en Gradiente Conjugado.

La función P que aparece en la ecuación (2.6) la calculamos resolviendo el problema directo conocido como *ecuación adjunta*

$$\left. \begin{aligned} -\frac{\partial P}{\partial t} - \Delta P &= 0 & \text{en} & \Omega \times (0, T) \\ aP + \frac{\partial P}{\partial n} &= 0 & \text{en} & \Gamma_1 \times (0, T) \\ P &= 0 & \text{en} & \Gamma_c \times (0, T) \\ P(T) &= \frac{1}{\epsilon}(u(T) - u_T) & \text{en} & \Omega \end{aligned} \right\} \quad (2.7)$$

Hay que observar que este no es un problema con valor inicial, si no, con valor final.

2.2 Aplicación del Método de Gradiente Conjugado

Una vez que formulamos la primera variación como una sola integral (2.6) (producto interior en \mathbb{H}), podemos aplicar el algoritmo de gradiente conjugado (Teorema 1.12) para nuestro problema.

Tomamos $v^0 \in \mathbb{H}$, puede ser $v^0 = 0$.

Paso 1) Hallar $g^0 \in \mathbb{H}$ tal que $\forall \delta v \in \mathbb{H}$

$$(g^0, \delta v) = \delta J[v^0; \delta v],$$

en términos de integrales requerimos que

$$\begin{aligned} \int_0^T \int_{\Gamma_c} g^0 \delta v d\gamma dt &= \int_0^T \int_{\Gamma_c} \left(v^0 - \frac{\partial P}{\partial n} \right) \delta v d\gamma dt \quad \forall \delta v \\ \Rightarrow \int_0^T \int_{\Gamma_c} \left\{ g^0 - \left(v^0 - \frac{\partial P}{\partial n} \right) \right\} \delta v d\gamma dt &= 0 \quad \forall \delta v \\ \Rightarrow g^0 &= v^0 - \frac{\partial P}{\partial n} \end{aligned}$$

se toma $\frac{\partial P}{\partial n}$ solo en $\Gamma_c \times (0, T)$.

El cálculo de P se realiza en dos pasos:

(1.i) Resolver para obtener $u^0(T)$

$$\begin{aligned} \frac{\partial u^0}{\partial t} - \Delta u^0 &= f & \text{en} & \Omega \times (0, T) \\ au^0 + \frac{\partial u^0}{\partial n} &= ag & \text{en} & \Gamma_1 \times (0, T) \\ u^0 &= v^0 & \text{en} & \Gamma_c \times (0, T) \\ u^0(0) &= u_0(x, y) & \text{en} & \Omega \end{aligned}$$

Es decir, resolvemos el problema directo (2.1) para el control v^0 .

(1.ii) Resolver para obtener P

$$\begin{aligned} -\frac{\partial P}{\partial t} - \Delta P &= 0 & \text{en} & \Omega \times (0, T) \\ aP + \frac{\partial P}{\partial n} &= 0 & \text{en} & \Gamma_1 \times (0, T) \\ P &= 0 & \text{en} & \Gamma_c \times (0, T) \\ P(T) &= \frac{1}{\epsilon}(u^0(T) - u_T) & \text{en} & \Omega, \end{aligned}$$

es decir, resolvemos la ecuación adjunta (2.7) con la solución $u^0(T)$ calculada en (1.i).

(1.iii) Hacemos la asignación

$$g^0 = v^0 - \frac{\partial P}{\partial n}$$

donde $\frac{\partial P}{\partial n}$ solo se toma en $\Gamma_c \times (0, T)$.

Paso 2) Hacemos la asignación

$$\omega^0 = g^0$$

la cual define la dirección de descenso, hasta aquí tenemos calculados v^0, g^0, ω^0 .

Para $n \geq 0$ calculamos v^{n+1}, g^{n+1} y ω^{n+1} de v^n, g^n, ω^n de la siguiente forma:

Paso 3) Hallar $\rho_n \in \mathbb{R}$ que minimice

$$J[v^n - \rho \omega^n] \quad \forall \rho \in \mathbb{R}.$$

este es un problema de minimización de una función de una variable real ρ de valores reales ya que v^n y ω^n están fijas. Por la linealidad, al control $v^n - \rho\omega^n$ le corresponde la solución $u^n - \rho z^n$ del problema directo (2.1).

(3.i) z^n se calcula como solución del problema

$$\begin{aligned} \frac{\partial z^n}{\partial t} - \Delta z^n &= 0 & \text{en} & \Omega \times (0, T) \\ az^n + \frac{\partial z^n}{\partial n} &= 0 & \text{en} & \Gamma_1 \times (0, T) \\ z^n &= \omega^n & \text{en} & \Gamma_c \times (0, T) \\ z^n(0) &= 0 & \text{en} & \Omega \end{aligned}$$

$$J[v^n - \rho\omega^n] = \frac{1}{2} \int_0^T \int_{\Gamma_c} (v^n - \rho\omega^n)^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} [(u^n - \rho z^n)(T) - u_T]^2 dx dy$$

$$\begin{aligned} \frac{d}{d\rho} J[v^n - \rho\omega^n] &= - \int_0^T \int_{\Gamma_c} (v^n - \rho\omega^n)\omega^n d\gamma dt \\ &\quad - \frac{1}{\epsilon} \int_{\Omega} [u^n(T) - \rho z^n(T) - u_T] z^n(T) dx dy \end{aligned}$$

$$\frac{d^2}{d\rho^2} J[v^n - \rho\omega^n] = \int_0^T \int_{\Gamma_c} (\omega^n)^2 d\gamma dt + \int_{\Omega} (z^n(T))^2 dx dy > 0 \quad \forall \rho \in \mathbb{R}$$

esto significa que la solución de la ecuación $\frac{d}{d\rho} J[v^n - \rho\omega^n] = 0$ es el mínimo ρ_n buscado, es decir:

(3.ii)

$$\rho_n = \frac{\int_0^T \int_{\Gamma_c} v^n \omega^n d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u^n(T) - u_T) z^n(T) dx dy}{\int_0^T \int_{\Gamma_c} (\omega^n)^2 d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (z^n(T))^2 dx dy}$$

Paso 4) Hacemos la asignación

$$v^{n+1} = v^n - \rho_n \omega^n$$

que es el punto donde más ha descendido la funcional $J[v]$ en la dirección ω^n .

Paso 5) Hallar $g^{n+1} \in \mathbb{H}$ tal que $\forall \delta v \in \mathbb{H}$

$$(g^{n+1}, \delta v) = \delta J[v^{n+1}, \delta v]$$

(5.i) como en el paso 1 resolvemos el problema directo (2.1) con el control v^{n+1} para obtener $u^{n+1}(T)$.

(5.ii) Resolvemos la ecuación adjunta (2.7) con $u^{n+1}(T)$ para obtener P .

(5.iii) Hacemos la asignación

$$g^{n+1} = v^{n+1} - \frac{\partial P}{\partial n}$$

y como antes $\frac{\partial P}{\partial n}$ sólo se considera en $\Gamma_c \times (0, T)$.

Si $\|g^{n+1}\|/\|g^0\| \leq \text{Tol}$ entonces el óptimo es

$$v^* = v^{n+1},$$

de otra forma,

$$\text{Paso 6) } \gamma = \frac{(g^{n+1}, g^{n+1})}{(g^n, g^n)}$$

Paso 7) Hacemos la asignación $\omega^{n+1} = g^{n+1} + \gamma \omega^n$.

Paso 8) Tomar $n \leftarrow n + 1$ y regresar al paso 3).

La convergencia del algoritmo se garantiza por que el control óptimo v^* es único y la funcional de energía cumple con los requisitos pedidos en el teorema 1.10.

Concluimos esta sección con un comentario sobre el criterio de paro en el paso 5 del algoritmo, este sugiere que cuando el gradiente g^{n+1} sea pequeño (cercano a cero), el correspondiente control v^{n+1} se tome como el mínimo, esto es correcto ya que el teorema 1.6.2 garantiza que cuando la primera variación es cero (o el gradiente) entonces el control correspondiente es el mínimo global.

2.3 Discretización del Problema Directo

En la sección 2.2 se desglosó el algoritmo de gradiente conjugado para nuestro problema de difusión, en éste se observa que es necesario resolver de forma iterativa problemas directos asociados a la ecuación de calor (pasos 1.i, 1.ii, 3.i, 5.i, 5.ii).

Hay que resolver el problema directo (2.1) para diferentes lados derechos (pasos 1.i, 3.i, 5.i), pero conservando los dominios y los lados izquierdos de las ecuaciones.

El otro problema directo por resolver es el de la ecuación adjunta, que es un problema con valor final, es decir, se necesita cumplir una condición al tiempo $T > 0$ (pasos 1.ii, 5.ii).

$$\begin{aligned}
 -\frac{\partial P}{\partial t} - \Delta P &= 0 && \text{en} && \Omega \times (0, T) \\
 aP + \frac{\partial P}{\partial n} &= 0 && \text{en} && \Gamma_1 \times (0, T) \\
 P &= 0 && \text{en} && \Gamma_c \times (0, T) \\
 P(T) &= \frac{1}{\epsilon}(u(T) - u_T) && \text{en} && \Omega
 \end{aligned}$$

Para resolver este problema, lo escribimos como un problema con valor inicial haciendo un cambio de variable en tiempo y dejando fijas las variables espaciales

$$t := T - \tau, \quad q(x, y, \tau) := P(x, y, T - \tau) \quad 0 \leq \tau \leq T.$$

$$\Rightarrow \frac{\partial q}{\partial \tau} = \frac{\partial P}{\partial t} \cdot \frac{\partial t}{\partial \tau} = -\frac{\partial P}{\partial t}$$

y todas las derivadas espaciales $\left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial^2}{\partial x^2}, \text{etc.}\right)$ son iguales. Además $q(x, y, 0) = P(x, y, T) \forall (x, y) \in \Omega$. Con este cambio de variable la ecuación adjunta se escribe como:

$$\begin{aligned}
 \frac{\partial q}{\partial \tau} - \Delta q &= 0 && \text{en} && \Omega \times (0, T) \\
 aq + \frac{\partial q}{\partial n} &= 0 && \text{en} && \Gamma_1 \times (0, T) \\
 q &= 0 && \text{en} && \Gamma_c \times (0, T) \\
 q(0) &= \frac{1}{\epsilon}(u(T) - u_T) && \text{en} && \Omega
 \end{aligned}$$

que nuevamente es un problema directo como (2.1) pero para diferentes lados derechos.

Esto nos lleva a concluir que para implementar computacionalmente el algoritmo de gradiente conjugado necesitamos un programa que resuelva el problema directo (2.1) en forma eficiente y versatil, es decir, que sea posible correrlo para diferentes lados derechos y que de buenos resultados en su salida ($u(x, y, t)$), ya que ésta se usa en otros pasos del proceso iterativo.

El método de discretización que usamos para resolver el problema directo (2.1) es el de diferencias finitas tanto en tiempo como en espacio, lo cual se facilita ya que el dominio de trabajo $\Omega = [0, 1] \times [0, 1]$ es el cuadrado unitario que representa una lámina donde se desea controlar la distribución de temperatura.

Para discretizar en tiempo usamos la fórmula de Gear modificada, que es una fórmula en diferencias que aproxima la primera derivada usando dos puntos hacia atrás igualmente espaciados y con un error de truncamiento de orden dos

$$f'(x_{i+2}) = \frac{\frac{3}{2}f(x_{i+2}) - 2f(x_{i+1}) + \frac{1}{2}f(x_i)}{\Delta x} + o(\Delta x^2),$$

donde $x_{i+2} = x_{i+1} + \Delta x$, $x_{i+1} = x_i + \Delta x$ y $\Delta x > 0$ es el tamaño de paso.

Este método que es de dos pasos necesita de un método de un paso para arrancar, en este caso usamos la fórmula de Euler hacia atrás, que aproxima la primera derivada usando un punto hacia atrás y tiene un error de truncamiento de primer orden.

$$f'(x_{i+1}) = \frac{f(x_{i+1}) - f(x_i)}{\Delta x} + o(\Delta x).$$

Para discretizar en espacio las primeras y segundas derivadas, usamos fórmulas de diferencias centradas para tres puntos igualmente espaciados con error de truncamiento de orden dos

$$f''(x_i) = \frac{f(x_{i-1}) - 2f(x_i) + f(x_{i+1}))}{h^2} + o(h^2)$$

$$f'(x_i) = \frac{f(x_{i+1}) - f(x_{i-1}))}{2h} + o(h^2)$$

donde $x_{i-1} = x_i - h$, $x_{i+1} = x_i + h$ y $h > 0$ es el tamaño de paso.

Estas fórmulas se deducen fácilmente usando el teorema de Taylor para una variable.

Discretización de la Ecuación de Estado.

Sea $\Delta t > 0$ el tamaño de paso en tiempo y $t_n = n\Delta t$ $n = 0, 1, 2, \dots$ los niveles de tiempo, denotamos por $u^k = u(x, y, t_k)$ y por $f^k = f(x, y, t_k)$.

Aplicando la fórmula de Gear modificada a la ecuación de calor en el nivel de tiempo $n + 1$ ($n \geq 1$) obtenemos

$$\frac{\frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1}}{\Delta t} + o(\Delta t)^2 - \Delta u^{n+1} = f^{n+1}$$

despreciando el término de error y acomodando términos tenemos:

$$\frac{3}{2\Delta t}u^{n+1} - \Delta u^{n+1} = f^{n+1} + \frac{2}{\Delta t}u^n - \frac{1}{2\Delta t}u^{n-1}$$

lo cual nos indica que en cada nivel de tiempo hay que resolver un problema elíptico.

Definimos $F^{n+1} = f^{n+1} + \frac{2}{\Delta t}u^n - \frac{1}{2\Delta t}u^{n-1}$, entonces hay que resolver

$$\frac{3}{2\Delta t}u^{n+1} - \Delta u^{n+1} = F^{n+1} \quad \text{en } \Omega.$$

Sea $h = \frac{1}{N+1}$ el tamaño de paso en espacio, $x_i = ih$ $i = 0, N + 1$ y $y_j = jh$ $j = 0, N + 1$ las coordenadas de los nodos de la malla sobre $\Omega = [0, 1] \times [0, 1]$.

Vamos a denotar por $u_{ij}^k = u^k(x_i, y_j)$ y $f_{ij}^k = f^k(x_i, y_j)$, aplicando las fórmulas de diferencias centradas al operador Δu en los nodos (x_i, y_j) $i = 1, N; j = 1, N$, tenemos:

$$\frac{3}{2\Delta t}u_{ij}^{n+1} - \frac{u_{i-1j}^{n+1} - 2u_{ij}^{n+1} + u_{i+1j}^{n+1}}{h^2} - \frac{u_{ij-1}^{n+1} - 2u_{ij}^{n+1} + u_{ij+1}^{n+1}}{h^2} + o(h^2) = F_{ij}^{n+1}. \quad (2.8)$$

Despreciando el término de error y agrupando términos tenemos:

$$\left(\frac{3}{2\Delta t} + \frac{4}{h^2} \right) u_{ij}^{n+1} - \frac{1}{h^2}u_{i+1j}^{n+1} - \frac{1}{h^2}u_{i-1j}^{n+1} - \frac{1}{h^2}u_{ij+1}^{n+1} - \frac{1}{h^2}u_{ij-1}^{n+1} = F_{ij}^{n+1};$$

$i = 1, N, \quad j = 1, N.$

Para completar el problema, discretizamos las condiciones de frontera sobre los puntos de la malla ya definidos y en el nivel de tiempo $n + 1$.

La condición de frontera tipo Dirichlet sobre Γ_c

$$u_{i\sigma}^{n+1} = v_i^{n+1} = v(x_i, t_{n+1}) \quad i = 1, N \quad (2.9)$$

condición de frontera mixta en $\{1\} \times [0, 1]$

$$au^{n+1} + \frac{\partial u^{n+1}}{\partial n} = ag^{n+1} \Rightarrow au^{n+1} + \nabla u^{n+1} \cdot \hat{n} = ag^{n+1}$$

donde $\hat{n} = (1, 0)$

$$\begin{aligned} \Rightarrow au^{n+1} + \frac{\partial u^{n+1}}{\partial x} &= ag^{n+1} \\ \Rightarrow au_{N+1j}^{n+1} + \frac{\frac{3}{2}u_{N+1j}^{n+1} - 2u_{Nj}^{n+1} + \frac{1}{2}u_{N-1j}^{n+1}}{h} + o(h^2) &= ag_{N+1j}^{n+1} \end{aligned}$$

despreciando $o(h^2)$ y agrupando tenemos:

$$u_{N+1j}^{n+1} = \frac{2h}{3 + 2ah} \left(ag_{N+1j}^{n+1} + \frac{2}{h}u_{Nj}^{n+1} - \frac{1}{2h}u_{N-1j}^{n+1} \right) \quad j = 1, N \quad (2.10)$$

hay que observar que discretizamos $\frac{\partial u^{n+1}}{\partial x}$ con un esquema de orden dos (Gear modificado), para ser consistentes con las demás discretizaciones en espacio.

La discretización de las otras condiciones de frontera son semejantes, en $[0, 1] \times \{1\}$ obtenemos:

$$u_{iN+1}^{n+1} = \frac{2h}{3 + 2ah} \left(ag_{iN+1}^{n+1} + \frac{2}{h}u_{iN}^{n+1} - \frac{1}{2h}u_{iN-1}^{n+1} \right) \quad i = 1, N \quad (2.11)$$

y en $\{0\} \times [0, 1]$

$$u_{oj}^{n+1} = \frac{2h}{3 + 2ah} \left(ag_{oj}^{n+1} + \frac{2}{h}u_{1j}^{n+1} - \frac{1}{2h}u_{2j}^{n+1} \right) \quad j = 1, N \quad (2.12)$$

La ecuación (2.8) junto con (2.9), (2.10), (2.11) y (2.12) forman el *problema discreto* en el nivel de tiempo $n + 1$ para el problema directo (2.1).

Al escribir la ecuación (2.8) para cada $i = 1, N$ y $j = 1, N$ e incorporando las condiciones (2.9), (2.10), (2.11) y (2.12), obtenemos un sistema lineal de N^2 ecuaciones con N^2 incógnitas. Al resolver dicho sistema obtenemos una

aproximación discreta a $u(x, y, t_{n+1})$ formada por $\{u_{ij}^{n+1}\}$ $i = 1, N, j = 1, N$, finalmente, obtenemos las incógnitas en la frontera usando las ecuaciones (2.10), (2.11) y (2.12). Si lo que se desea es obtener una aproximación para $u(x, y, T)$, entonces se resuelve este problema $M - 1$ veces, donde M debe cumplir que $M\Delta t = T$, decimos $M - 1$ veces ya que el primer nivel de tiempo se calcula en otra forma, la razón es la siguiente. El esquema de Gear modificado para discretizar en tiempo nos dá un método de solución de dos pasos en cada nivel, esto es claro en la definición de

$$F^{n+1} = f^{n+1} + \frac{2}{\Delta t}u^n - \frac{1}{2\Delta t}u^{n-1} \quad n \geq 1$$

que es el lado derecho del problema elíptico por resolver.

Para $n = 1$ necesitamos conocer u^0 y u^1 (de hecho las aproximaciones discretas), u^0 se conoce, ya que es la condición inicial del problema (2.1), pero u^1 debe ser calculado por un método de un solo paso, en este caso usamos la fórmula de Euler hacia atrás, es decir, discretizamos como:

$$\frac{\partial u^1}{\partial t} = \frac{u^1 - u^0}{\Delta t} + o(h).$$

Siguiendo los pasos análogos a la discretización con el esquema de Gear, obtenemos que el problema (2.1) para el primer nivel de tiempo es:

$$\left(\frac{1}{\Delta t} + \frac{4}{h^2}\right)u_{ij}^1 - \frac{1}{h^2}u_{i-1j}^1 - \frac{1}{h^2}u_{i+1j}^1 - \frac{1}{h^2}u_{ij-1}^1 - \frac{1}{h^2}u_{ij+1}^1 = G_{ij} \quad i = 1, N; j = 1, N$$

donde $G_{ij} = f_{ij}^1 + \frac{1}{\Delta t}u_{ij}^0$ junto con las ecuaciones (2.9), (2.10), (2.11) y (2.12) para $n = 0$, ya que usamos la misma discretización para la frontera.

Como en el caso anterior, estas ecuaciones nos definen un sistema lineal $N^2 \times N^2$, que al resolver nos dan $\{u_{ij}^1\}$ $i = 1, N, j = 1, N$ la aproximación discreta a u^1 necesaria para arrancar el método de dos pasos.

2.4 Aspectos de Algebra Lineal Numérica en la Implementación de Gradiente Conjugado.

En la sección 2.3 establecimos la necesidad de resolver el problema directo (2.1) de forma eficiente y versátil ya que este problema aparece repetidamente en el algoritmo de gradiente conjugado. Para calcular la solución de

(2.1), discretizamos en tiempo y espacio con fórmulas de diferencias finitas y obtuvimos el problema discreto:

Resolver para cada $n \geq 1$

$$\left(\frac{3}{2\Delta t} + \frac{4}{h^2}\right) u_{ij}^{n+1} - \frac{1}{h^2} u_{i+1j}^{n+1} - \frac{1}{h^2} u_{i-1j}^{n+1} - \frac{1}{h^2} u_{ij+1}^{n+1} - \frac{1}{h^2} u_{ij-1}^{n+1} = F_{ij}^{n+1}, \quad i = 1, N; \quad j = 1, N. \quad (2.8)$$

$$\left\{ \begin{array}{l} u_{i0}^{n+1} = v_i^{n+1} \quad i = 1, N \end{array} \right. \quad (2.9)$$

$$\left\{ \begin{array}{l} u_{N+1j}^{n+1} = \frac{2h}{3+2ah} \left(ag_{N+1j}^{n+1} + \frac{2}{h} u_{Nj}^{n+1} - \frac{1}{2h} u_{N-1j}^{n+1} \right) \quad j = 1, N \end{array} \right. \quad (2.10)$$

$$\left\{ \begin{array}{l} u_{iN+1}^{n+1} = \frac{2h}{3+2ah} \left(ag_{iN+1}^{n+1} + \frac{2}{h} u_{iN}^{n+1} - \frac{1}{2h} u_{iN-1}^{n+1} \right) \quad i = 1, N \end{array} \right. \quad (2.11)$$

$$\left\{ \begin{array}{l} u_{0j}^{n+1} = \frac{2h}{3+2ah} \left(ag_{0j}^{n+1} + \frac{2}{h} u_{1j}^{n+1} - \frac{1}{2h} u_{2j}^{n+1} \right) \quad j = 1, N \end{array} \right. \quad (2.12)$$

donde $F_{ij}^{n+1} = f_{ij}^{n+1} + \frac{2}{\Delta t} u_{ij}^n - \frac{1}{2\Delta t} u_{ij}^{n-1}$.

Además $\{u_{ij}^1\}$ $i = 1, N; \quad j = 1, N$ se obtiene resolviendo

$$\left(\frac{1}{\Delta t} + \frac{4}{h^2}\right) u_{ij}^1 - \frac{1}{h^2} u_{i+1j}^1 - \frac{1}{h^2} u_{i-1j}^1 - \frac{1}{h^2} u_{ij+1}^1 - \frac{1}{h^2} u_{ij-1}^1 = G_{ij}, \quad i = 1, N; \quad j = 1, N. \quad (2.13)$$

junto con (2.9), (2.10), (2.11) y (2.12) para $n = 0$, donde

$$G_{ij} = f_{ij}^1 + \frac{1}{\Delta t} u_{ij}^0.$$

Para cada $n \geq 1$ el problema (2.8)-(2.12) consiste en resolver un sistema lineal de N^2 ecuaciones y N^2 incógnitas, la matriz \mathbb{A} del sistema y el lado derecho \mathbf{b}^{n+1} se muestran a continuación.

$$\mathbb{A} = \begin{bmatrix} A & C & & & \\ D & B & D & & \\ & \ddots & \ddots & \ddots & \\ & & D & B & D \\ & & & C & A \end{bmatrix}_{N^2 \times N^2}$$

$$\begin{aligned}
b_1 &= \begin{bmatrix} F_{11}^{n+1} - c_2 v_1 - c_2 c_3 g_{01}^{n+1} \\ F_{12}^{n+1} - c_2 c_3 g_{02}^{n+1} \\ \vdots \\ F_{1N}^{n+1} - c_2 c_3 g_{0N-1}^{n+1} \\ F_{1N}^{n+1} - c_2 c_3 (g_{1N+1}^{n+1} + g_{0N}^{n+1}) \end{bmatrix} \\
b_k &= \begin{bmatrix} F_{k1}^{n+1} - c_2 v_k \\ F_{k2}^{n+1} \\ \vdots \\ F_{kN-1}^{n+1} \\ F_{kN}^{n+1} - c_2 c_3 g_{kN+1}^{n+1} \end{bmatrix} \quad k = 2, \quad N-1 \\
b_N &= \begin{bmatrix} F_{N1}^{n+1} - c_2 v_N - c_2 c_3 g_{N+11}^{n+1} \\ F_{N2}^{n+1} - c_2 c_3 g_{N+12}^{n+1} \\ \vdots \\ F_{NN-1}^{n+1} - c_2 c_3 g_{N+1N-1}^{n+1} \\ F_{NN}^{n+1} - c_2 c_3 (g_{N+1N}^{n+1} + g_{NN+1}^{n+1}) \end{bmatrix}
\end{aligned}$$

El problema discreto consiste entonces en resolver el sistema $\mathbb{A}U^{n+1} = \mathbf{b}^{n+1}$ para cada nivel de tiempo, donde $U^{n+1} = \{u_{ij}^{n+1}\}$ $i = 1, N; j = 1, N$.

Es muy importante notar que la matriz \mathbb{A} no depende del nivel de tiempo en el que se esté resolviendo el problema, ya que sus entradas dependen solo de las constantes a , h y Δt que se fijan a lo largo de todo el proceso de solución; lo que cambia al resolver en cada nivel de tiempo es el lado derecho del sistema \mathbf{b}^{n+1} , ya que este depende de F^{n+1} y g^{n+1} .

La matriz \mathbb{A} es dominante diagonalmente estrictamente, entonces es invertible, para verificarlo solo hay que mostrar tres condiciones que corresponden a tres renglones de la matriz; la dominancia en los otros renglones es inmediata.

i) $|c_1| > 4|c_2|$.

$$|c_1| = \frac{3}{2\Delta t} + \frac{4}{h^2} \text{ y } 4|c_2| = \frac{4}{h^2} \Rightarrow |c_1| > 4|c_2|.$$

ii) $|c_1 + c_2 c_4| > 2|c_2| + |c_2 + c_2 c_5|$.

$$|c_1 + c_2 c_4| = \frac{1}{h^2} \left(4 + \frac{3h^2}{2\Delta t} - \frac{4}{3 + 2ah} \right)$$

$$2|c_2| + |c_2 + c_2c_5| = \frac{1}{h^2} \left(2 + 1 - \frac{1}{3 + 2ah} \right)$$

$$|c_1 + c_2c_4| - (2|c_2| + |c_2 + c_2c_5|) = \frac{1}{h^2} \left(1 + \frac{3h^2}{2\Delta t} - \frac{3}{3 + 2ah} \right) > 0.$$

$$\text{iii) } |c_1 + 2c_2c_4| > 2|c_2 + c_2c_5|.$$

$$|c_1 + 2c_2c_4| = \frac{1}{h^2} \left(4 + \frac{3h^2}{2\Delta t} - \frac{8}{3 + 2ah} \right).$$

$$2|c_2 + c_2c_5| = \frac{1}{h^2} \left(2 - \frac{2}{3 + 2ah} \right)$$

$$|c_1 + 2c_2c_4| - 2|c_2 + c_2c_5| = \frac{1}{h^2} \left(2 + \frac{3h^2}{2\Delta t} - \frac{6}{3 + 2ah} \right) > 0.$$

El hecho de que la matriz \mathbb{A} sea estacionaria y domine diagonalmente, muestra que el sistema lineal $\mathbb{A}U^{n+1} = \mathbf{b}^{n+1}$ tiene solución única para cualquier lado derecho \mathbf{b}^{n+1} , es decir en cualquier nivel de tiempo.

El problema discreto (2.13), (2.9), (2.10), (2.11) y (2.12) que necesitamos resolver para el cálculo de $U^1 = \{u_{ij}^1\} i = 1, N, j = 1, N$; también consiste en resolver un sistema lineal $\mathbb{A}_0U^1 = \mathbf{b}^1$ de N^2 incógnitas y N^2 ecuaciones. La matriz \mathbb{A}_0 es igual a la matriz \mathbb{A} salvo que hay que cambiar c_1 por $c_0 = \frac{1}{\Delta t} + \frac{4}{h^2}$, la propiedad de ser dominante diagonalmente estrictamente se preserva (el argumento es el mismo que para \mathbb{A}) y por lo tanto $\mathbb{A}_0U^1 = \mathbf{b}^1$ tiene solución única.

El lado derecho \mathbf{b}^1 es igual a \mathbf{b}^{n+1} salvo que hay que cambiar F_{ij}^{n+1} por G_{ij} y la función g^{n+1} se toma en $n = 0, i = 0, N + 1; j = 0, N + 1$.

Para calcular la solución de estos sistemas de ecuaciones usamos el método de *Gauss-Seidel por Bloques*, este método iterativo se adapta en forma sencilla a una matriz con las características de \mathbb{A} (\mathbb{A}_0), es decir, tridiagonal por bloques y los bloques de la diagonal son a su vez matrices tridiagonales.

El algoritmo general para resolver $\mathbb{A}x = \mathbf{b}$ donde \mathbb{A} es una matriz con las características antes mencionadas es

Para $k = 0, k_{\max}$

$$1. \text{ Resolver } Ax_1^{(k+1)} = b_1 - Cx_2^{(k)};$$

$$\text{Para } i = 2, \dots, N - 1$$

$$7) \quad q_{N-1} := p.$$

Para resolver $LUx = \mathbf{b}$, hacemos el cambio de variable $y = Ux$, entonces resolvemos primero $Ly = \mathbf{b}$ por sustitución hacia adelante

$$1) \quad y_1 := \frac{b_1}{l_1}$$

Para $i = 2, \dots, N-1$

$$2) \quad y_i := \frac{b_i - sy_{i-1}}{l_i}$$

$$3) \quad y_N := \frac{b_N - py_{N-1}}{l_N}$$

Ahora resolvemos $Ux = y$ por sustitución hacia atrás

$$1) \quad x_N := y_N$$

Para $i = N-1, \dots, 1$

$$2) \quad x_i := y_i - u_i x_{i+1}$$

Como consecuencia de que A es estacionaria (y por lo cual A y B lo son), tenemos que las factorizaciones LU de A y B se hacen una sola vez y se almacenan en vectores para usarse en la solución del problema directo (2.1) en cada nivel de tiempo.

Con esto se completa la descripción de la solución numérica del problema directo (2.1), el cual es parte esencial del algoritmo de gradiente conjugado, el programa escrito en FORTAN 77 que incorpora estas ideas aparece en el Apéndice 2 con el nombre de RECUCAL.FOR, ejemplos de la solución numérica de problemas directos con este programa, aparecen en la siguiente sección.

En la sección 2.2, se desglosó el algoritmo de gradiente conjugado, el cual se programó en FORTAN 77 y aparece en el Apéndice 2 con el nombre de GRACON.FOR. Este programa llama a RECUCAL.FOR según lo indica el algoritmo en los pasos 1.i, 1.ii, 3.i, 5.i y 5.ii.

Otros cálculos importantes en gradiente conjugado son las integrales en $L^2(\Gamma_c \times (0, T))$ y $L^2(\Omega)$ (productos interiores), que aparecen en los pasos 3.ii y 6, estas integrales corresponden a funciones de dos variables sobre rectángulos; su cálculo se realizó con el método de Trapecio que tiene un

error de truncamiento $o(h^2)$, lo cual es congruente con las discretizaciones hechas en espacio. El algoritmo es material estándar y se puede ver en [5], los programas correspondientes escritos en FORTRAN 77 aparecen en el Apéndice 2 con los nombres de PIL20T y PIL2.FOR.

2.5 Resultados Numéricos

En esta sección se presentan resultados numéricos y gráficas de ejemplos particulares tanto para problemas directos como para problemas inversos.

Ejemplo 1.

Consideramos el problema directo (2.1) con los siguientes datos:

$$f(x, y, t) = [x^2(x-1)^2 + (y-1)^2] - 4(1+t)[3x^2 - 3x + 1],$$

$$v(x, t) = (x^2(x-1)^2 + 1)(1+t)$$

$$u_0(x, y) = x^2(x-1)^2 + (y-1)^2.$$

Al resolver este problema con el programa RECUCAL.FOR para el caso $a = 0$ y hasta $T = 1$ se obtiene la siguiente tabla de resultados:

	h	Δt	M	TGS	IGS	$ERROR$
<i>a</i>	$\frac{1}{10}$	$\frac{1}{10}$	10	10^{-16}	170	7.23×10^{-2}
<i>b</i>	$\frac{1}{16}$	$\frac{1}{20}$	20	10^{-16}	240	3.066×10^{-2}
<i>c</i>	$\frac{1}{32}$	$\frac{1}{100}$	100	10^{-16}	260	8.148×10^{-3}
<i>d</i>	$\frac{1}{40}$	$\frac{1}{100}$	100	10^{-16}	325	5.317×10^{-3}

Donde $h > 0$ es el tamaño de paso en la discretización e espacio; $\Delta t > 0$ es el tamaño de paso en la discretización de tiempo; M representa los niveles de tiempo, tal que $M\Delta t = T$; TGS es la tolerancia en el método iterativo de Gauss-Seidel para resolver los sistemas lineales; IGS son las iteraciones necesarias para alcanzar la TGS ; $error = \|u(T) - u_T\|_\infty$.

El tiempo de máquina en una PC-486 para obtener la solución de este problema con una discretización como en c) es de 120 sg. (aproximadamente), y aumenta conforme la discretización en espacio y tiempo es mas fina. Los errores mostrados en la tabla son congruentes, ya que el método de discretización en espacio y tiempo tiene un error de truncamiento $o(h^2 + \Delta t^2)$. Las

gráficas 2.1.1 y 2.1.2 al final de la sección muestran los cortes de la solución exacta ($u(x, y, t) = [x^2(x - 1)^2 + (y - 1)^2](1 + t)$) y la solución aproximada para la discretización c) de la tabla.

Considerar el problema inverso asociado.

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy, \quad \epsilon > 0$$

Sujeto a:

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f && \text{en } \Omega \times (0, T) \\ au + \frac{\partial u}{\partial n} &= ag && \text{en } \Gamma_1 \times (0, T) \\ u &= v && \text{en } \Gamma_c \times (0, T) \\ u(0) &= u_0(x, y) && \text{en } \Omega \end{aligned}$$

donde a , $f(x, y, t)$ y $u_0(x, y)$ están definidas como arriba. Para este problema consideramos que la distribución por alcanzar al tiempo $T = 1$ es

$$u_T(x, y) = 2[x^2(x - 1)^2 + (y - 1)^2].$$

Los resultados numéricos obtenidos en esta parte no son satisfactorios ya que no se logra la convergencia del método, se puede decir que el algoritmo sufre un estancamiento, es decir, el criterio de paro de gradiente conjugado $\|g^k\|/\|g^0\|$ el cual se busca se anule no desciende más allá de una cota. Buscamos que se anule el gradiente (g^k) o la primera variación ya que esto es una condición suficiente para obtener el mínimo (Teorema 1.6.2). Para ejemplificar esto consideramos la corrida con los datos $h = \frac{1}{16}$, $\Delta t = \frac{1}{100}$, $M = 100$, $TGS = 10^{-16}$, $\epsilon = 10^{-4}$; en este caso el proceso se estanca después de 12 iteraciones en el valor $\|g^k\|/\|g^0\| = 0.0532$. La gráfica 2.1.3 muestra el control exacto ($v^*(x, t) = (x^2(x - 1)^2 + 1)(1 + t)$) y el control calculado, así como los cortes de la distribución u_T y la distribución calculada. Los errores para estos son:

$$\|v^* - v\|_{\infty} = 0.02$$

$$\|u_T - u(T)\|_{\infty} = 0.42.$$

El tiempo de máquina utilizado en una PC-486 es de aproximadamente 20 min., esto se debe a que cada iteración de Gradiente Conjugado requiere resolver 3 problemas directos como el descrito al principio de este ejemplo.

Antes de dar una explicación de este comportamiento desarrollamos otro ejemplo.

Ejemplo 2.

Consideramos el problema directo (2.1) con los siguientes datos:

$$\begin{aligned} f(x, y, t) &= -\frac{10}{(1+t)^2}(y-1)^2 - \frac{20}{1+t}, \\ v(x, t) &= \frac{10}{1+t}, \\ u_0(x, y) &= 10(y-1)^2. \end{aligned}$$

Al resolver este problema con el programa RECUCAL.FOR para el caso $a = 0$ y hasta $T = 1$ se obtiene la siguiente tabla de resultados

	h	Δt	M	TGS	IGS	$ERROR$
a	$\frac{1}{10}$	$\frac{1}{10}$	10	10^{-16}	175	2.806×10^{-4}
b	$\frac{1}{16}$	$\frac{1}{20}$	20	10^{-16}	255	7.042×10^{-5}
c	$\frac{1}{32}$	$\frac{1}{100}$	100	10^{-16}	230	2.539×10^{-6}
d	$\frac{1}{40}$	$\frac{1}{100}$	100	10^{-16}	360	2.501×10^{-6}

El tiempo de máquina requerido en una PC-486 para obtener la solución de este problema con una discretización como en c) es de 120 seg. aproximadamente. El tiempo de cómputo aumenta conforme la discretización es más fina.

Esto se debe a que la dimensión de los sistemas lineales por resolver es mayor, y por lo tanto la solución requiere de más operaciones aritméticas.

Los errores mostrados en la tabla indican que la solución aproximada es correcta, es decir, se resuelve el problema directo en forma satisfactoria.

La gráfica 2.2.1 al final de la sección muestra el corte de la solución exacta ($u(x, y, t) = (1+t)[x^2(x-1)^2 + (y-1)^2]$) y la solución aproximada para la discretización c) de la tabla.

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy, \quad \epsilon > 0$$

Sujeto a:

$$\begin{aligned} \frac{\partial u}{\partial t} - \Delta u &= f && \text{en } \Omega \times (0, T) \\ au + \frac{\partial u}{\partial n} &= ag && \text{en } \Gamma_1 \times (0, T) \\ u &= v && \text{en } \Gamma_c \times (0, T) \\ u(0) &= u_0(x, y) && \text{en } \Omega \end{aligned}$$

donde a , $f(x, y, t)$ y $u_0(x, y)$ están definidas como arriba. Para este problema consideramos que la distribución por alcanzar al tiempo $T = 1$ es

$$u_T(x, y) = 5(y - 1)^2.$$

Para el cálculo del control óptimo (v) de este problema usamos el programa GRACON.FOR, el cual es una implementación del método de gradiente conjugado analizado en la sección 2.2.

Al igual que el ejemplo 1, los resultados numéricos en esta parte no son satisfactorios, ya que no se logra la convergencia del método, de hecho el proceso iterativo sufre un estancamiento. Para la corrida con los datos:

$$h = \frac{1}{16}, \quad \Delta t = \frac{1}{100}, \quad M = 100 \quad TGS = 10^{-16} \quad \epsilon = 10^{-4}$$

el proceso se estanca después de 16 iteraciones y el criterio de paro toma el valor $\|g^k\|/\|g^0\| = 0.0428$.

La gráfica 2.2.2 muestra el control exacto ($v^* = \frac{10}{1+t}$) y distribución u_T y la distribución calculada. Los errores correspondientes son:

$$\|v^* - v\|_\infty = 0.32$$

$$\|u_T - u(T)\|_\infty = 0.37$$

El tiempo de máquina utilizado en una PC-486 es de aproximadamente 28 min., es importante observar que cada iteración de gradiente conjugado requiere resolver 3 problemas directos como el descrito al principio de este ejemplo.

Estos dos ejemplos muestran que existe una deficiencia en la solución de los problemas inversos. El método de gradiente conjugado para minimizar una funcional es eficiente, por lo cual el error debe estar en la implementación

de éste. Básicamente, lo que se hace en una iteración del método de gradiente conjugado es resolver 3 problemas directos, calcular productos interiores (integrales dobles) y calcular gradientes.

Por lo visto en los ejemplos 1 y 2 concluimos que la solución de los problemas directos se realiza en forma satisfactoria. Por otra parte, el cálculo de productos interiores con el método de Trapecio es material estándar y las rutinas utilizadas (PIL20T, PILZ) se revisaron satisfactoriamente. Esto nos lleva a concluir que el error se encuentra en el cálculo del gradiente:

$$g^k = v^k - \frac{\partial P}{\partial n} \quad \text{en } \Gamma_c \times (0, T)$$

donde P es la solución de la ecuación adjunta.

El cálculo de P es correcto ya que es solución de un problema directo, de hecho el error está en el cálculo de la derivada normal $\frac{\partial P}{\partial n}$, esta derivada se aproxima con fórmulas de diferencias finitas hacia adelante, es decir, para calcular la derivada en puntos de la frontera usamos puntos en el interior de Ω ; estas fórmulas producen pérdida de cifras en los datos y de hecho son inestables, es decir, conforme la discretización es más fina ($h \rightarrow 0$), los resultados numéricos son más inexactos.

En los ejemplos 1 y 2 el cálculo de $\frac{\partial P}{\partial n}$ se realiza con una fórmula de tres puntos:

$$\frac{\partial P}{\partial n} = \nabla P \cdot n = -\frac{\partial P}{\partial n} = -\frac{2P_{i2} - \frac{1}{2}P_{i3} - \frac{3}{2}P_{i1}}{h}$$

para cada $i = 0, N + 1$.

Los resultados obtenidos se dan en la siguiente tabla:

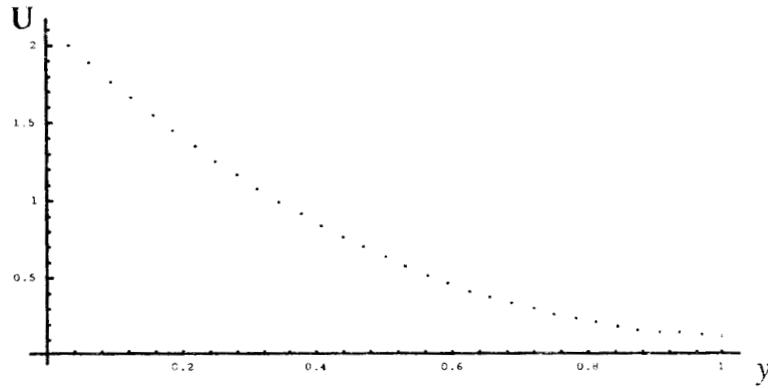
h	Δt	M	Error del problema directo($T = 1$)	Error en la derivada normal($T = 1$)
$\frac{1}{32}$	$\frac{1}{100}$	100	2.539×10^{-6}	8.001×10^{-3}
$\frac{1}{40}$	$\frac{1}{100}$	100	2.501×10^{-6}	7.477×10^{-3}
$\frac{1}{80}$	$\frac{1}{100}$	100	2.003×10^{-6}	1.902×10^{-2}
$\frac{1}{100}$	$\frac{1}{100}$	100	1.905×10^{-6}	3.807×10^{-2}

Los datos para resolver el problema directo son los del ejemplo 2.

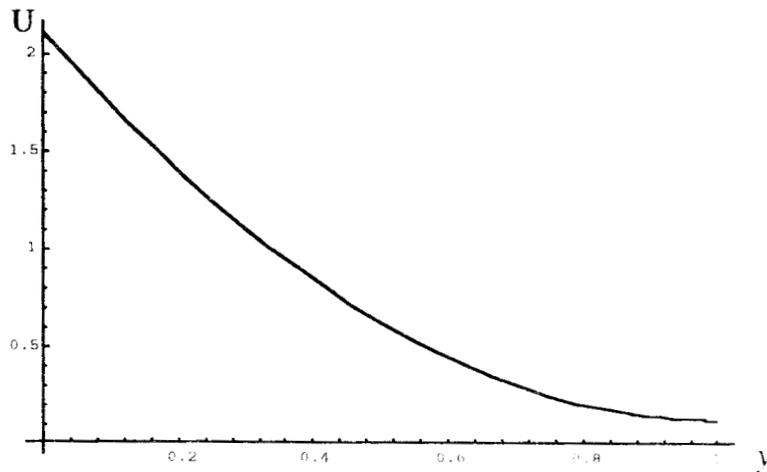
Los experimentos numéricos del cálculo de la derivada normal confirman que el gradiente (g^k) se calcula en forma deficiente, esto descompone el proceso de gradiente conjugado en su parte más importante ya que g^k define la dirección de descenso en la búsqueda del mínimo.

En las conclusiones abordamos las alternativas para este cálculo que como se ha visto es muy importante.

PROBLEMA DIRECTO: ECUACION DE DIFUSION



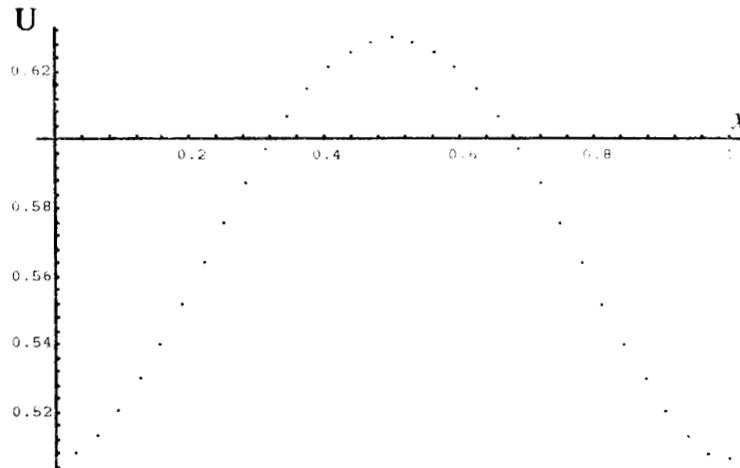
SOLUCION APROXIMADA (Corte para $x=1/2$ y $T=1$)



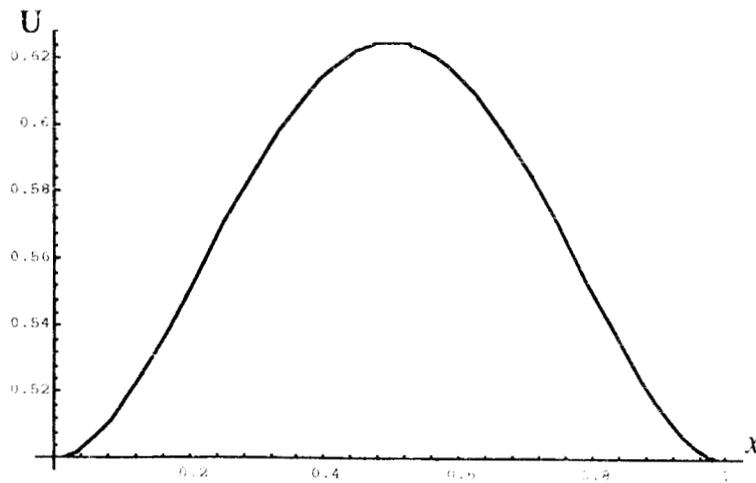
SOLUCION EXACTA (Corte para $x=1/2$ y $T=1$)

Grafica 2.1.1

PROBLEMA DIRECTO: ECUACION DE DIFUSION.



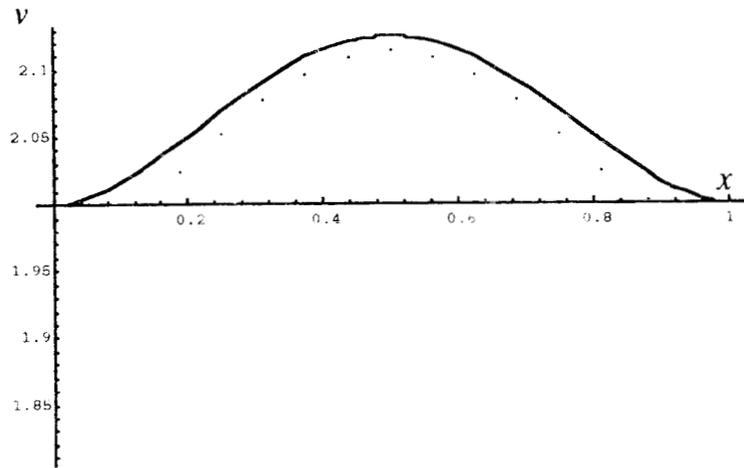
SOLUCION APROXIMADA (Corte para $y=1/2$ y $T=1$)



SOLUCION EXACTA (Corte para $y=1/2$ y $T=1$)

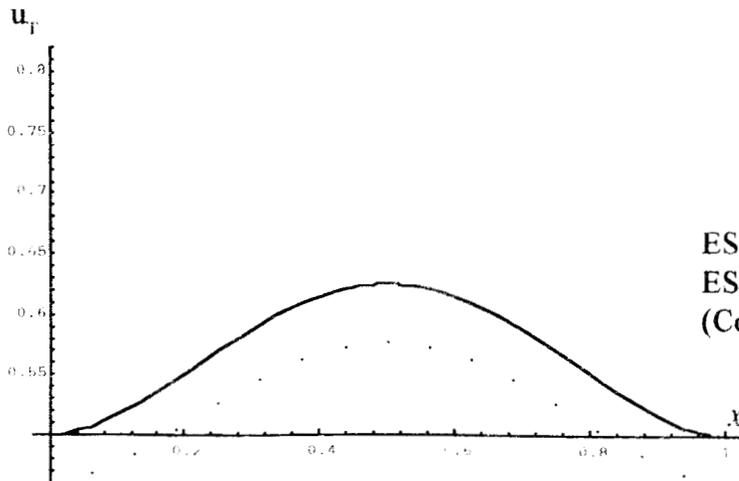
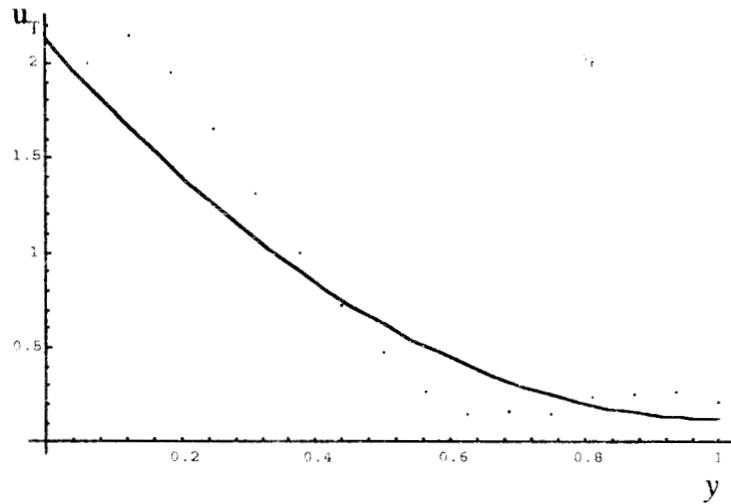
Grafica 2.1.2

PROBLEMA INVERSO: ECUACION DE DIFUSION



CONTROL EXACTO VS CONTROL APROXIMADO (Corte para $T=1$)

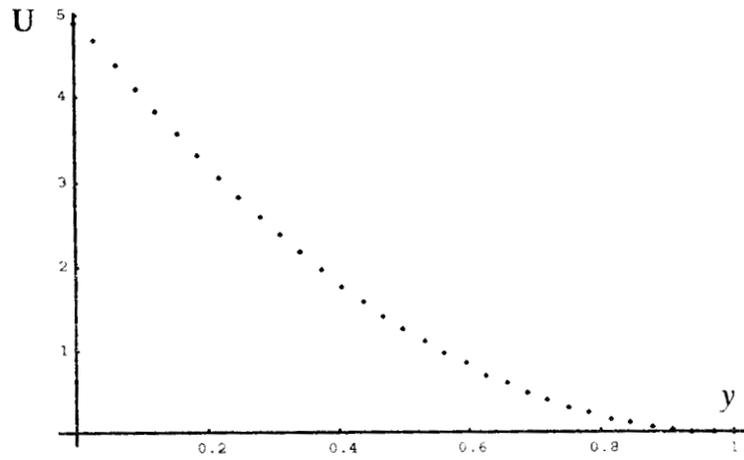
ESTADO EXACTO ($T=1$) VS ESTADO APROXIMADO (Corte para $x=1/2$)



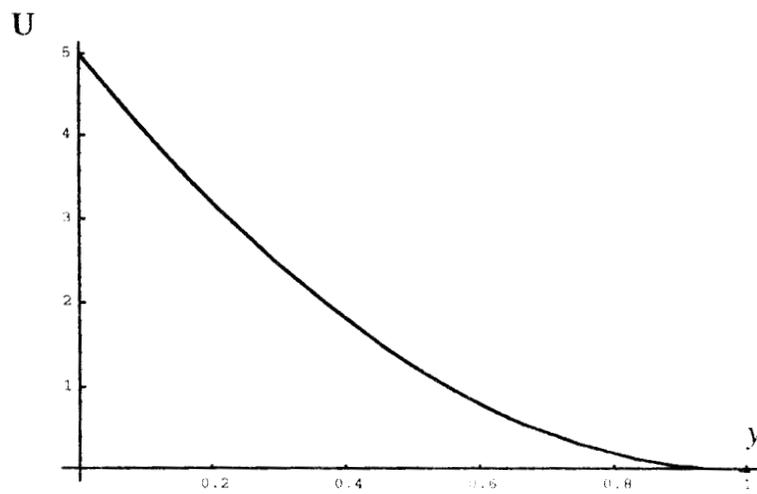
ESTADO EXACTO ($T=1$) VS ESTADO APROXIMADO (Corte para $y=1/2$)

Grafica 2.1.3

PROBLEMA DIRECTO: ECUACION DE DIFUSION



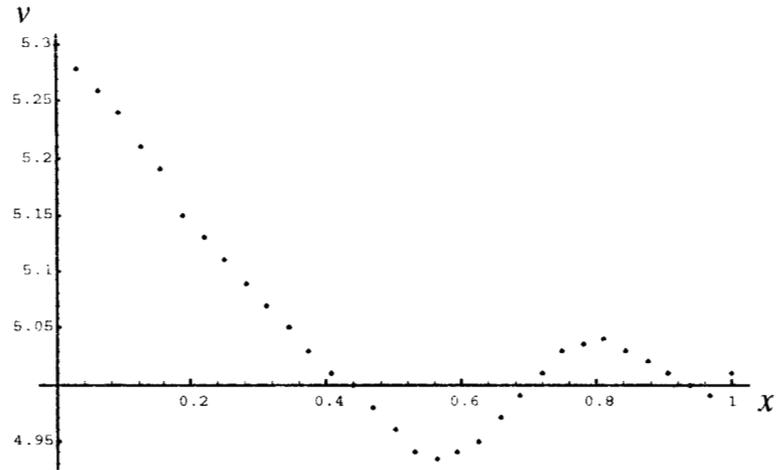
SOLUCION APROXIMADA (Corte para $x=1/2$ y $T=1$)



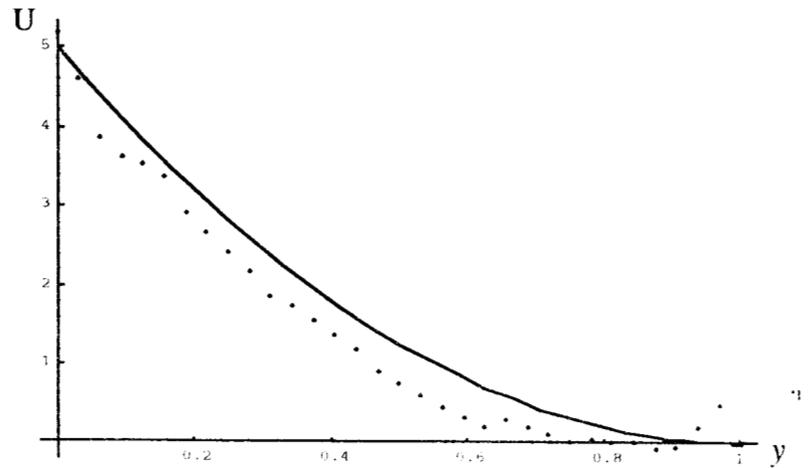
SOLUCION EXACTA (Corte para $x=1/2$ y $T=1$)

Grafica 2.2.1

PROBLEMA INVERSO: ECUACION DE DIFUSION



CONTROL EXACTO ($\nu=5$) VS CONTROL APROXIMADO (Corte para $T=1$)



ESTADO EXACTO ($T=1$) VS ESTADO APROXIMADO (Corte para $x=1/2$)

Grafica 2.2.2

CAPITULO III

3 Control en la frontera tipo Dirichlet para la Ecuación de Transporte

En este capítulo calculamos el mínimo de la funcional de energía en el espacio $H = L^2(\Gamma_c \times (0, T))$, solo que ahora la restricción es la ecuación de transporte, la cual modela la distribución de la temperatura en un fluido incompresible ($\text{div}(\mathbf{w}) = 0$) de velocidad \mathbf{w} conocida. El problema de control óptimo para esta ecuación lo describimos en la sección 1.1. El método de solución es a través del algoritmo de gradiente conjugado, mostramos la aplicación de éste y la discretización del problema directo correspondiente, al final del capítulo damos los resultados numéricos.

Los cálculos y ejemplos mostrados en el capítulo anterior indican que es arduo, pero no imposible, controlar el comportamiento interno de un proceso de difusión desde la frontera. Muchos fenómenos físicos pueden ser vistos como una combinación de procesos de difusión y advección, entonces la pregunta natural es ¿cómo afecta el fenómeno advectivo la contrabilidad interna de un sistema? Recordamos que el fenómeno de difusión está dado por Δu y el de advección por $\mathbf{w} \cdot \nabla u$.

El problema a resolver y al cual nos referimos en todo el capítulo es el siguiente

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy \quad \epsilon > 0,$$

para $v \in H$, sujeto a:

$$\left. \begin{aligned} \frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u &= f, \text{div}(\mathbf{w}) = 0 && \text{en } \Omega \times (0, T) \\ au + \beta \frac{\partial u}{\partial n} &= ag && \text{en } \Gamma_1 \times (0, T) \\ u &= v && \text{en } \Gamma_c \times (0, T) \\ u(0) &= u_0(x, y) && \text{en } \Omega \end{aligned} \right\} \quad (3.1)$$

donde H denota al espacio Hilbert $L^2[\Gamma_c \times (0, T)]$.

Hay que notar que el problema (3.1), a diferencia del problema (2.1) (el cual es un problema potencial), es un problema no potencial lineal, aún en el

caso de que la condición de frontera sea solamente de Dirichlet. Esto conduce a que la matriz del sistema algebraico en el problema totalmente discreto sea no simétrica.

Como en el capítulo anterior, el control es de tipo Dirichlet, es decir, buscamos controlar la distribución de temperatura interna del fluido por medio del control de temperatura en la frontera.

3.1 Cálculo de la Ecuación Adjunta

En la sección 2.1 del capítulo II notamos que para aplicar el algoritmo de gradiente conjugado al problema de control óptimo con la ecuación de difusión, es necesario escribir la primera variación de la funcional de energía en términos únicamente del producto interior de H , en dicho cálculo aparece la nueva función que satisface una ecuación importante conocida como *ecuación adjunta*. Para el problema de control óptimo con la ecuación de transporte repetimos el proceso, es decir, calculamos $g^k \in H$ tal que

$$(g^k, \delta v) = \delta J[v^k; \delta v] \quad \forall \delta v \in H,$$

la existencia y unicidad de dicha representación la asegura el teorema de representación de Riez.

La funcional a minimizar en ambos procesos, difusión y difusión advección, es la misma, y como ambas ecuaciones de estado son lineales tenemos que la primera variación también. En el teorema 1.5 se obtuvo (ecuación 1.6) que la primera variación es

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} v \delta v d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) dx dy$$

donde δu es la variación de la solución u del problema directo (3.1), que corresponde a la variación δv del control v , δu se calcula por la linealidad de la ecuación de estado y condiciones de frontera e iniciales como solución del problema directo

$$\frac{\partial \delta u}{\partial t} - \beta \Delta \delta u + \mathbf{w} \cdot \nabla \delta u = 0, \quad \text{en } \Omega \times (0, T) \quad (3.2)$$

$$a \delta u + \beta \frac{\partial \delta u}{\partial n} = 0 \quad \text{en } \Gamma_1 \times (0, T) \quad (3.3)$$

$$\delta u = \delta v \quad \text{en } \Gamma_c \times (0, T) \quad (3.4)$$

$$\delta u(0) = 0 \quad \text{en } \Omega \quad (3.5)$$

Introducimos una función $P = P(x, y, t)$ definida para $(x, y) \in \Omega$ y $t \geq 0$, con suficiente suavidad y cuyas características se determinan en el siguiente desarrollo.

Multiplicamos (3.2) por P e integramos

$$\int_0^T \int_{\Omega} P \frac{\partial \delta u}{\partial t} - \int_0^T \int_{\Omega} \beta P \Delta \delta u + \int_0^T \int_{\Omega} P \mathbf{w} \cdot \nabla \delta u = 0,$$

en la primera integral hacemos una integración por partes en tiempo y en la segunda aplicamos la segunda fórmula de Green para obtener

$$\int_{\Omega} \left[P \delta u \Big|_0^T - \int_0^T \frac{\delta u \partial P}{\partial t} \right] - \beta \int_0^T \left[\int_{\Omega} \delta u \Delta P + \int_{\Gamma} \left(P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right) \right] + \int_0^T \int_{\Omega} P \mathbf{w} \cdot \nabla \delta u = .$$

usando (3.5) y $\nabla(P\delta u) = P\nabla\delta u + \delta u\nabla P$ en la última integral, podemos escribir

$$\int_{\Omega} P(T) \delta u(T) + \int_0^T \int_{\Omega} \delta u \left(-\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P \right) - \beta \int_0^T \int_{\Gamma} \left(P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right) + \int_0^T \int_{\Omega} \mathbf{w} \cdot \nabla(P\delta u) = 0,$$

pedimos a P que cumpla

$$\frac{-\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P = 0 \text{ en } \Omega \times (0, T),$$

entonces tenemos

$$\int_{\Omega} P(T) \delta u(T) - \beta \int_0^T \int_{\Gamma} \left(P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right) + \int_0^T \int_{\Omega} \mathbf{w} \cdot \nabla(P\delta u) = 0,$$

por otra parte, tomando en cuenta que $\text{div}(\mathbf{w}) = 0$ tenemos

$$\text{div}(\mathbf{w}(P\delta u)) = \mathbf{w} \cdot \nabla(P\delta u) + (P\delta u)\text{div}(\mathbf{w}) = \mathbf{w} \cdot \nabla(P\delta u)$$

aplicando esto a la última integral y usando el teorema de la divergencia, podemos escribir

$$\int_{\Omega} P(T) \delta u(T) - \beta \int_0^T \int_{\Gamma} \left(P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right) + \int_0^T \int_{\Gamma} (P\delta u) \mathbf{w} \cdot \mathbf{n} = 0,$$

las últimas dos integrales las separamos en las secciones de frontera Γ_c y Γ_1

$$\int_{\Omega} P(T)\delta u(T) - \beta \int_0^T \int_{\Gamma_c} P \frac{\partial \delta u}{\partial n} - \beta \int_0^T \int_{\Gamma_1} P \frac{\partial \delta u}{\partial n} + \beta \int_0^T \int_{\Gamma_c} \delta u \frac{\partial P}{\partial n} + \beta \int_0^T \int_{\Gamma_1} \delta u \frac{\partial P}{\partial n} + \int_0^T \int_{\Gamma_c} (P\delta u)\mathbf{w} \cdot \mathbf{n} + \int_0^T \int_{\Gamma_1} (P\delta u)\mathbf{w} \cdot \mathbf{n} = 0,$$

pedimos que $P = 0$ en $\Gamma_c \times (0, T)$. Aplicamos (3.3) y (3.4) para escribir

$$\int_{\Omega} P(T)\delta u(T) + \beta \int_0^T \int_{\Gamma_c} \frac{\partial P}{\partial n} \delta v + \int_0^T \int_{\Gamma_1} (aP + \beta \frac{\partial P}{\partial n} + P\mathbf{w} \cdot \mathbf{n})\delta u = 0,$$

pedimos que $aP + \beta \frac{\partial P}{\partial n} + \mathbf{w} \cdot \mathbf{n}P = 0$ en $\Gamma_1 \times (0, T)$ y $P(T) = \frac{1}{\epsilon}(u(T) - u_T)$ en Ω , entonces

$$\int_0^T \int_{\Gamma_c} \beta \frac{\partial P}{\partial n} \delta v + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T)\delta u(T) = 0,$$

restando esta última igualdad de la ecuación (1.6) obtenemos

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} \left(v - \beta \frac{\partial P}{\partial n} \right) \delta v d\gamma dt \quad (3.6)$$

que es la formulación adecuada de la primera variación de la funcional de energía para utilizar en gradiente conjugado.

La función P que aparece en (3.6) se calcula resolviendo el problema directo siguiente, conocido como *ecuación adjunta* (caso transporte).

$$\left. \begin{aligned} -\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P &= 0 & \text{en } \Omega \times (0, T) \\ aP + \beta \frac{\partial P}{\partial n} + \mathbf{w} \cdot \mathbf{n}P &= 0 & \text{en } \Gamma_1 \times (0, T) \\ P &= 0 & \text{en } \Gamma_c \times (0, T) \\ P(T) &= \frac{1}{\epsilon}(u(T) - u_T) & \text{en } \Omega \end{aligned} \right\} \quad (3.7)$$

El problema (3.7) es un problema con valor final, y como en el problema de difusión lo resolvemos como un problema de valor inicial mediante un cambio de variable.

3.2 Aplicación del Método de Gradiente Conjugado

En la sección 2.2 del capítulo anterior, aplicamos el algoritmo de gradiente conjugado (Teorema 1.12) al problema de control óptimo en el proceso de difusión, en este capítulo lo aplicamos al proceso de difusión-advección, el desarrollo es similar, ya que la funcional a minimizar y su primera variación son semejantes.

Paso 1) Hallar $g^0 \in H$ tal que $\forall \delta v \in H$

$$(g^0, \delta v) = \delta J[v^0; \delta v],$$

en términos de productos interiores podemos escribir (3.6) como

$$\delta J[v; \delta v] = \left(v - \beta \frac{\partial P}{\partial n}, \delta v \right),$$

igualando expresiones obtenemos

$$g^0 = v^0 - \beta \frac{\partial P}{\partial n}.$$

El cálculo de P se realiza en dos pasos:

1.i) Resolver para obtener $u^0(T)$

$$\begin{aligned} \frac{\partial u^0}{\partial t} - \beta \Delta u^0 + \mathbf{w} \cdot \nabla u^0 &= f, & \text{en } \Omega \times (0, T) \\ au^0 + \beta \frac{\partial u^0}{\partial n} &= ag & \text{en } \Gamma_1 \times (0, T) \\ u^0 &= v^0 & \text{en } \Gamma_c \times (0, T) \\ u^0(0) &= u_0(x, y) & \text{en } \Omega, \end{aligned}$$

es decir, resolvemos el problema directo (3.1) para el control v^0 .

1.ii) Resolver para obtener P

$$\begin{aligned} -\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P &= 0 & \text{en } \Omega \times (0, T) \\ aP + \beta \frac{\partial P}{\partial n} + \mathbf{w} \cdot nP &= 0 & \text{en } \Gamma_1 \times (0, T) \\ P &= 0 & \text{en } \Gamma_c \times (0, T) \\ P(T) &= \frac{1}{\epsilon}(u(T) - u_T) & \text{en } \Omega, \end{aligned}$$

es decir, resolvemos la ecuación adjunta (3.7) con la solución $u^0(T)$ calculada en 1.i).

1.iii) Hacemos la asignación

$$g^0 = v^0 - \frac{\partial P}{\partial n}$$

donde $\frac{\partial P}{\partial n}$ solo se toma en $\Gamma_c \times (0, T)$.

Paso 2) Hacemos la asignación

$$\omega^0 = g^0$$

la cual define la dirección de descenso, hasta aquí tenemos calculados v^0, g^0, ω^0 .

Para $n \geq 0$ calculamos v^{n+1}, g^{n+1} y ω^{n+1} de v^n, g^n y ω^n de la siguiente forma

Paso 3) Hallar $\rho_n \in \mathbb{R}$ que minimice

$$J[v^n - \rho\omega^n] \quad \forall \rho \in \mathbb{R},$$

el cálculo de ρ_n se hace en dos pasos

3.i) Calcular z^n como solución del problema

$$\begin{aligned} \frac{\partial z^n}{\partial t} - \beta \Delta z^n + \mathbf{w} \cdot \nabla z^n &= 0 & \text{en } \Omega \times (0, T) \\ az^n + \beta \frac{\partial z^n}{\partial n} &= 0 & \text{en } \Gamma_1 \times (0, T) \\ z^n &= \omega^n & \text{en } \Gamma_c \times (0, T) \\ z^n(0) &= 0 & \text{en } \Omega \end{aligned}$$

3.ii)

$$\rho_n = \frac{\int_0^T \int_{\Gamma_c} v^n \omega^n d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u^n(T) - u_T) z^n(T) dx dy}{\int_0^T \int_{\Gamma_c} (\omega^n)^2 d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (z^n(T))^2 dx dy},$$

los detalles de este cálculo son los mismos que para el problema de difusión, pueden verse en la sección 2.2.

Paso 4) Hacemos la asignación

$$v^{n+1} = v^n - \rho_n \omega^n$$

que es el punto donde más ha descendido la funcional $J[v]$ en la dirección ω^n .

Paso 5) Hallar $g^{n+1} \in H$ tal que $\forall \delta v \in H$

$$(g^{n+1}, \delta v) = \delta J[v^{n+1}; \delta v].$$

5.i) Resolvemos (3.1) con el control v^{n+1} para obtener $u^{n+1}(T)$.

5.ii) Resolvemos la ecuación adjunta (3.7) con $u^{n+1}(T)$, para obtener P .

5.iii) Hacemos la asignación

$$g^{n+1} = v^{n+1} - \beta \frac{\partial P}{\partial n} \text{ en } \Gamma_c \times (0, T).$$

Si $\|g^{n+1}\|/\|g^0\| \leq \text{Tol}$ entonces el óptimo es $v^* = v^{n+1}$, de otra forma,

$$\text{Paso 6) } \gamma = \frac{(g^{n+1}, g^{n+1})}{(g^n, g^n)}$$

Paso 7) Hacemos la asignación

$$\omega^{n+1} = g^{n+1} + \gamma \omega^n,$$

Paso 8) Tomar $n \leftarrow n + 1$ y regresar a 3).

La funcional de energía tiene un mínimo único en H y cumple con las hipótesis del teorema 1.10, entonces la convergencia está asegurada y el criterio de paro usado es que el gradiente se anule, lo cual por el teorema 1.6.2 sabemos que es condición suficiente para el cálculo del mínimo.

3.3 Discretización del Problema Directo.

La aplicación del método de gradiente conjugado requiere resolver la ecuación de difusión-advección (problema directo 3.1) y la ecuación adjunta (3.7) en forma repetitiva, por lo cual requerimos programas eficientes y versátiles para cada uno de estos problemas.

La ecuación adjunta (3.7) es un problema de valor final, es decir, debemos cumplir una condición al tiempo $T > 0$. Para resolver este problema lo pasamos a uno de valor inicial con un cambio de variable en tiempo y dejando fijas las variables espaciales.

$$t := T - \tau, \quad q(x, y, \tau) := P(x, y, T - t) \quad 0 \leq \tau \leq T$$

$$\implies \frac{\partial q}{\partial \tau} = \frac{\partial P}{\partial t} \cdot \frac{\partial t}{\partial \tau} = -\frac{\partial P}{\partial t}$$

y todas las derivadas espaciales son iguales, además

$$q(x, y, 0) = P(x, y, T) \quad \forall (x, y) \in \Omega.$$

Con este cambio la ecuación adjunta se escribe como

$$\left. \begin{aligned} \frac{\partial q}{\partial t} - \beta \Delta q - \mathbf{w} \cdot \nabla q &= 0 & \text{en } \Omega \times (0, T) \\ a q + \beta \frac{\partial q}{\partial n} + \mathbf{w} \cdot \mathbf{n} q &= 0 & \text{en } \Gamma_1 \times (0, T) \\ q &= 0 & \text{en } \Gamma_c \times (0, T) \\ q(0) &= \frac{1}{\epsilon} (u(T) - u_T) & \text{en } \Omega \end{aligned} \right\} \quad (3.8)$$

que es el problema con valor inicial por resolver.

El método de discretización que usamos para resolver los problemas directos (3.1) y (3.8) es el de diferencias finitas, lo cual se facilita ya que el dominio de trabajo es $\Omega = [0, 1] \times [0, 1]$. Para discretizar en tiempo usamos los esquemas de Gear modificado y Euler hacia atrás, mientras que en espacio usamos fórmulas de diferencias centradas para aproximar la primera y segunda derivada. En la sección 2.3 hemos descrito estas ecuaciones.

Discretizamos solo (3.1) ya que (3.8) se puede ver como un caso particular.

Discretización de la Ecuación de Estado.

Sea $\Delta t > 0$ el tamaño de paso en tiempo y $t_n = n\Delta t$, $n = 0, 1, 2, \dots$ los niveles de tiempo, denotamos por $u^k = u(x, y, t_k)$ y por $f^k = f(x, y, t_k)$.

Aplicando el esquema de Gear modificado para discretizar $\frac{\partial u}{\partial t}$ en la ecuación de difusión-advención y en el nivel de tiempo $n + 1$ ($n \geq 1$), obtenemos

$$\frac{\frac{3}{2}u^{n+1} - 2u^n + \frac{1}{2}u^{n-1}}{\Delta t} - \beta \Delta u^{n+1} + \mathbf{w} \cdot \nabla u^{n+1} + o(\Delta t^2) = f^{n+1}$$

acomodando términos y despreciando el término de error, tenemos

$$\frac{3}{2\Delta t}u^{n+1} - \beta \Delta u^{n+1} + \mathbf{w} \cdot \nabla u^{n+1} = F^{n+1} \quad \text{en } \Omega$$

donde $F^{n+1} = f^{n+1} + \frac{2}{\Delta t}u^n - \frac{1}{2\Delta t}u^{n-1}$, que es el problema semidiscreto que hay que resolver en espacio, para cada nivel de tiempo.

Sea $h = \frac{1}{N+1}$ el tamaño de paso en espacio, $x_i = ih$ $i = 0, N + 1$ y $y_j = jh$ $j = 0, N + 1$ las coordenadas de los nodos (x_i, y_j) de la malla sobre $\Omega = [0, 1] \times [0, 1]$. Denotamos por $u_{ij}^k = u_{ij}^k$ y $f_{ij}^k = f^k(x_i, y_j)$, si discretizamos $\frac{\partial u^{n+1}}{\partial x}$, $\frac{\partial u^{n+1}}{\partial y}$, $\frac{\partial^2 u^{n+1}}{\partial x^2}$ y $\frac{\partial^2 u^{n+1}}{\partial y^2}$ con fórmulas centradas de orden dos y despreciamos el término de error, tenemos

$$c_1 u_{ij}^{n+1} + c_2 u_{i+ij}^{n+1} + c_3 u_{i-ij}^{n+1} + c_4 u_{ij+1}^{n+1} + c_5 u_{ij-1}^{n+1} = F_{ij}^{n+1}; \quad i = 1, N, \quad j = 1, N. \quad (3.9)$$

si consideramos la velocidad $\mathbf{w} = (\omega_1, \omega_2)$ con ω_1 y ω_2 constantes, entonces las constantes c_r $r = 1, 5$ se definen como

$$\begin{aligned} c_1 &= \frac{3}{2\Delta t} + \frac{4\beta}{h^2}, \\ c_2 &= \frac{\omega_1}{2h} - \frac{\beta}{h^2}, \\ c_3 &= -\frac{\beta}{h^2} - \frac{\omega_1}{2h}, \\ c_4 &= \frac{\omega_2}{2h} - \frac{\beta}{h^2}, \\ c_5 &= -\frac{\beta}{h^2} - \frac{\omega_2}{2h}. \end{aligned}$$

La discretización de las condiciones de frontera sobre los puntos de la malla antes definida y en cada nivel de tiempo $n + 1$, son semejantes a lo hecho para el problema de difusión, en Γ_c obtenemos

$$u_{i0}^{n+1} = v_i^{n+1} \quad i = 1, N \quad (3.10)$$

Γ_1 se divide en tres secciones, en $\{1\} \times [0, 1]$ tenemos

$$u_{N+1}^{n+1} = \frac{2h}{3\beta + 2ah} \left(ag_{N+1j}^{n+1} + \frac{2\beta}{h} u_{Nj}^{n+1} - \frac{\beta}{2h} u_{N-1j}^{n+1} \right) \quad j = 1, N \quad (3.11)$$

para $[0, 1] \times \{1\}$ obtenemos

$$u_{iN+1}^{n+1} = \frac{2h}{3\beta + 2ah} \left(ag_{iN+1}^{n+1} + \frac{2\beta}{h} u_{iN}^{n+1} - \frac{\beta}{2h} u_{iN-1}^{n+1} \right) \quad i = 1, N \quad (3.12)$$

finalmente, en $\{0\} \times [0, 1]$

$$u_{0j}^{n+1} = \frac{2h}{3\beta + 2ah} \left(ag_{0j}^{n+1} + \frac{2\beta}{h} u_{1j}^{n+1} - \frac{\beta}{2h} u_{2j}^{n+1} \right) \quad j = 1, N \quad (3.13)$$

Para cada una de las discretizaciones de las derivadas que provienen de $\frac{\partial u^{n+1}}{\partial n}$, se utiliza un esquema de Gear modificado, el cual tiene un error de truncamiento de orden dos, esto es consistente con las discretizaciones en Ω .

La ecuación (3.9) junto con (3.10), (3.11), (3.12) y (3.13) forman el *problema discreto* en el nivel de tiempo $n + 1$ para el problema directo (3.1).

Al escribir la ecuación (3.9) para $i = 1, N$ y $j = 1, N$ e incorporando las condiciones (3.10), (3.11), (3.12) y (3.13) obtenemos un sistema lineal de N^2 ecuaciones con N^2 incógnitas. Al resolver dicho sistema, obtenemos una aproximación discreta de $u(x, y, t_{n+1})$, formada por $\{u_{ij}^{n+1}\}$ $i = 1, N$ y $j = 1, N$, finalmente obtenemos las incógnitas en la frontera usando las ecuaciones (3.11), (3.12) y (3.13).

La discretización en tiempo con Gear modificado define un método de dos pasos, es decir, para aproximar u^{n+1} necesitamos u^n y u^{n-1} , esto se aprecia en la definición del lado derecho del problema semidiscreto,

$$F^{n+1} = f^{n+1} + \frac{2}{\Delta t} u^n - \frac{1}{2\Delta t} u^{n-1}.$$

Para arrancar el método necesitamos u^0 y u^1 , u^0 es la condición inicial ($u^0 = u_0(x, y)$), entonces hay que calcular u^1 con un método de un solo paso, como en el problema de difusión, usamos Euler hacia atrás para discretizar en tiempo la ecuación de difusión-advención, en espacio usamos diferencias centradas y las condiciones de frontera quedan como en el caso anterior, así tenemos

$$c_1^1 u_{ij}^1 + c_2^1 u_{i+1j}^1 + c_3^1 u_{i-1j}^1 + c_4^1 u_{ij+1}^1 + c_5^1 u_{ij-1}^1 = G_{ij} \quad i = 1, N, \quad j = 1, N \quad (3.14)$$

donde $G_{ij} = f_{ij}^1 + \frac{1}{\Delta t} u_{ij}^0$.

El problema (3.14) junto con (3.10), (3.11), (3.12) y (3.13) forman un sistema lineal $N^2 \times N^2$, las constantes c_r , $r = 2, 5$ tienen el mismo significado que en (3.9) y $c_1^1 = \frac{1}{\Delta t} + \frac{4\beta}{h^2}$.

Queda claro que en cada nivel de tiempo resolvemos un sistema lineal $N^2 \times N^2$ y para calcular la aproximación a $u(x, y, T) = u(T)$ resolvemos entonces M sistemas lineales, donde M cumple $T = M\Delta T$.

M y N son grandes ya que Δt y h se requieren pequeños para tener buenas aproximaciones, esto se traduce entonces en un arduo trabajo computacional para resolver el problema directo. Problemas directos en forma iterativa se resuelven dentro de gradiente conjugado, por lo tanto esperamos un tiempo de cómputo grande.

3.4 Aspectos de Algebra Lineal Numérica en la Implementación de Gradiente Conjugado

La solución del problema (3.1) en forma numérica la obtenemos resolviendo el sistema lineal $N^2 \times N^2$, descrito por las ecuaciones (3.9) ... (3.13), en cada nivel de tiempo $n = 2 \dots M$ y para $n = 1$ resolvemos (3.14) junto con (3.10), (3.11), (3.12) y (3.13).

Escribimos estos sistemas como:

$$\begin{cases} \mathbb{A}_0 U^1 = \mathbf{b}^1 \\ \mathbb{A} U^{n+1} = \mathbf{b}^{n+1} \end{cases} \quad n = 1 \dots M - 1 \quad (3.15)$$

donde $U^k = \{u_{ij}^k\}$ $i = 1, N; j = 1, N$, es la aproximación discreta a $u(x, y, t_k)$.

Las matrices \mathbb{A} y \mathbb{A}_0 no dependen del nivel de tiempo en el cual se calcula la solución, sólo los lados derechos en los sistemas, esto facilita la solución de (3.1) numéricamente, ya que las factorizaciones que de estas matrices se requieran sólo se hacen una vez y se almacenan.

La estructura de \mathbb{A} y \mathbb{A}_0 es tridiagonal por bloques, es decir,

$$\mathbb{A} = \begin{bmatrix} A & C & & & & \\ D' & B & D & & & \\ & \ddots & \ddots & \ddots & & \\ & & D' & B & D & \\ & & & C' & A' & \end{bmatrix}_{N^2 \times N^2}$$

Los bloques A, A', B, C, C', D y D' son de tamaño $N \times N$, además los bloques en la diagonal principal son tridiagonales, mientras que C, C', D y D' son diagonales, es decir,

$$A = \begin{bmatrix} c_1 + c_3 c_7 & c_4 & & & \\ & c_5 & & & \\ & & \ddots & \ddots & \ddots \\ & & & c_5 & c_1 + c_3 c_7 & c_4 \\ & & & c_5 + c_4 c_8 & c_1 + c_7(c_3 + c_4) & \end{bmatrix}_{N \times N}$$

$$B = \begin{bmatrix} c_1 & c_4 & & & \\ & c_5 & & & \\ & & \ddots & \ddots & \ddots \\ & & & c_5 & c_1 & c_4 \\ & & & c_5 + c_4 c_8 & c_1 + c_4 c_7 & \end{bmatrix}_{N \times N}$$

$$A' = \begin{bmatrix} c_1 + c_2 c_7 & c_4 & & & \\ & c_5 & & & \ddots \\ & & \ddots & \ddots & c_4 \\ & & & c_5 & c_1 + c_2 c_7 & c_4 \\ & & & c_5 + c_4 c_8 & c_1 + c_7(c_2 + c_4) & \end{bmatrix}_{N \times N}$$

$$C = (c_2 + c_3 c_8)\mathbb{I}, \quad C' = (c_3 + c_2 c_8)\mathbb{I}, \quad D = c_2 \mathbb{I} \text{ y } D' = c_3 \mathbb{I},$$

donde \mathbb{I} es la matriz identidad $N \times N$. La matriz \mathbb{A}_0 es igual a \mathbb{A} salvo que hay que cambiar c_1 por c_1^1 . Las constantes c_1^1 y c_r , $r = 1, 8$ se definieron en la sección 3.3.

Los vectores del lado derecho \mathbf{b} y \mathbf{b}^{n+1} tienen la siguiente estructura,

$$\mathbf{b}^{n+1} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N-1} \\ b_N \end{bmatrix}$$

$$\begin{aligned}
\mathbf{b}_1 &= \begin{bmatrix} F_{11}^{n+1} - c_3 c_6 g_{01}^{n+1} - c_5 v_1 \\ F_{12}^{n+1} - c_3 c_6 g_{02}^{n+1} \\ \vdots \\ F_{1N-1}^{n+1} - c_3 c_6 g_{0N}^{n+1} \\ F_{1N}^{n+1} - c_3 c_6 g_{0N}^{n+1} - c_4 c_6 g_{1N+1}^{n+1} \end{bmatrix}_{N \times 1} \\
b_k &= \begin{bmatrix} F_{k1}^{n+1} - c_5 v_k \\ F_{k2}^{n+1} \\ \vdots \\ F_{kN-1}^{n+1} \\ F_{kN}^{n+1} - c_4 c_6 g_{kN+1}^{n+1} \end{bmatrix}_{N \times 1} \quad k = 2, N-1 \\
b_N &= \begin{bmatrix} F_{N1}^{n+1} - c_2 c_6 g_{N+11}^{n+1} - c_5 v_N \\ F_{N2}^{n+1} - c_2 c_6 g_{N+12}^{n+1} \\ \vdots \\ F_{NN-1}^{n+1} - c_2 c_6 g_{N+11}^{n+1} - c_5 v_N \\ F_{NN}^{n+1} - c_2 c_6 g_{N+1N}^{n+1} - c_4 c_6 g_{NN+1}^{n+1} \end{bmatrix}_{N \times 1}
\end{aligned}$$

el vector \mathbf{b}^1 tiene la misma estructura que \mathbf{b}^{n+1} , solo hay que cambiar F_{ij}^{n+1} por G_{ij} $i = 1, N; j = 1, N$ y g_{ij}^{n+1} por g_{ij}^1 $i = 0, N+1; j = 0, N+1$.

En el problema de difusión las matrices asociadas a los sistemas lineales para resolver el problema directo (2.1) dominantes diagonalmente estrictamente, lo cual es condición suficiente para su solución. Cuando se introduce el fenómeno advectivo $\mathbf{w} \cdot \nabla u$ entonces las matrices \mathbb{A} y \mathbb{A}_0 asociadas a los sistemas lineales para resolver el problema directo (3.1) ya no son dominantes diagonalmente estrictamente en forma incondicional, este es el problema numérico asociado a un modelo más complejo. Sin embargo, el problema que nos interesa controlar en este trabajo es cuando el fenómeno es fuertemente advectivo, es decir, cuando $\beta \rightarrow 0$ en la ecuación de transporte. En este caso demostramos que las matrices \mathbb{A} y \mathbb{A}_0 son dominantes diagonalmente estrictamente en forma condicional, para esto necesitamos que se cumplan

las siguientes condiciones para \mathbb{A} ,

$$\begin{aligned} |c_1 + c_3c_7| &> |c_4| + |c_5| + |c_2 + c_3c_8| \\ |c_1 + c_3c_7 + c_4c_7| &> |c_5 + c_4c_8| + |c_2 + c_3c_8| \\ |c_1| &> |c_2| + |c_3| + |c_4| + |c_5| \\ |c_1 + c_4c_7| &> |c_2| + |c_3| + |c_5 + c_4c_8| \\ |c_1 + c_2c_7| &> |c_4| + |c_5| + |c_3 + c_2c_8| \\ |c_1 + c_2c_7 + c_4c_7| &> |c_5 + c_4c_8| + |c_3 + c_2c_8| \end{aligned}$$

que son satisfechos si la discretización cumple la condición

$$\frac{3h}{2\Delta t} > |\omega_1| + |\omega_2|$$

donde ω_1 y ω_2 son las componentes de la velocidad del fluido.

Las condiciones para \mathbb{A}_0 son las mismas solo que hay que cambiar c_1 por c_1^1 , en este caso son satisfechas bajo la condición

$$\frac{h}{\Delta t} > |\omega_1| + |\omega_2|. \quad (3.16)$$

Desde el punto de vista de la solución numérica de los sistemas lineales (3.15), el problema de difusión-advección es más difícil, ya que no hay condiciones suficientes incondicionales en las matrices que aseguren la solución, es decir, no son simétricas ni dominantes diagonalmente estrictamente, ni positivas definidas (para valores grandes de ω_1 ó ω_2 , algunos términos en la diagonal de \mathbb{A} y \mathbb{A}_0 son negativos). Por esta razón nos restringimos al control en el proceso de difusión advección cuando el fenómeno es fuertemente advectivo ($\beta \rightarrow 0$) y discretizamos bajo la condición (3.16), con el fin de asegurar la solución de los sistemas (3.15) y factorizaciones LU de los bloques de la diagonal, esto último es importante ya que el método de solución de los sistemas (3.15) lo hacemos usando el método iterativo de Gauss-Seidel por bloques, que es una forma de aprovechar los bloques $N \times N$ de ceros en las matrices y lo más recomendable en sistemas lineales grandes. Aclaremos que se realizaron cálculos con β no muy pequeño y se tuvo solución única, esto es consecuencia de que el problema continuo tiene solución única para $\beta > 0$. El algoritmo para resolver $\mathbb{A}x = \mathbf{b}$, donde \mathbb{A} tiene la estructura antes mencionada es,

Para $k = 0, \dots, k_{\max}$

Algoritmo de sustitución hacia adelante.

1. $y_1 := b_1$

Para $i = 2, \dots, N$

2. $y_i := b_i - \ell_{i-1}y_{i-1}$

Algoritmo de sustitución hacia atrás.

1. $x_N := y_N/u_N$

Para $i = N - 1, \dots, 1$

2. $x_i := \frac{y_i - tx_{i+1}}{u_i}$

Con esto se completa la descripción de la solución numérica del problema (3.1), el cual es parte esencial del algoritmo de Gradiente Conjugado, el programa que incorpora estas ideas fue escrito en FORTRAN 77 y aparece en el Apéndice 2 con el nombre de RECUTRA.FOR, ejemplos de la solución numérica de problemas directos con este programa, aparecen en la siguiente sección.

En la sección 3.2 se desglosó el algoritmo de gradiente conjugado para el problema de difusión-advección, el cual se programó en FORTRAN 77 y aparece en el apéndice 2 con el nombre de GRATRA.FOR. Este programa llama a RECUTRA.FOR según lo indica el algoritmo en los pasos 1.i, 1.ii, 3.i, 5i y 5.ii.

Otros cálculos importantes son las integrales en $L^2(\Omega)$ y $L^2(\Gamma_c \times (0, T))$ (productos interiores), que aparecen en los pasos 3.i y 6, estos cálculos como en el problema de difusión se obtienen con los programas PILZ.FOR y PILZOT.FOR que aparecen en el apéndice 2.

3.5 Resultados Numéricos

En esta sección se presentan resultados numéricos y gráficas de ejemplos particulares, tanto para problemas directos como para problemas inversos.

Ejemplo 1.

Considerar el problema directo (3.1) con los siguientes datos:

$$\begin{aligned}f(x, y, t) &= \frac{20}{(1+t)}(y-1-\beta) - \frac{10}{(1+t)^2}(y-1)^2, \\v(x, t) &= \frac{10}{1+t}, \\u_0(x, y) &= 10(y-1)^2.\end{aligned}$$

Al resolver este problema con el programa RECUTRA.FOR para el caso $a = 0$, $\mathbf{w} = (1, 1)$ y $\beta = 1$ y hasta $T = 1$ se obtienen los siguientes resultados:

	h	Δt	M	TGS	IGS	$ERROR$
a	$\frac{1}{10}$	0.1	10	10^{-12}	115	8.8532×10^{-4}
b	$\frac{1}{10}$	0.01	100	10^{-12}	22	7.9927×10^{-6}
c	$\frac{1}{32}$	0.01	100	10^{-12}	152	8.0224×10^{-6}
d	$\frac{1}{40}$	0.01	100	10^{-12}	46	8.0136×10^{-6}

Donde $h > 0$ es el tamaño de paso en la discretización en tiempo; $\Delta t > 0$ es el tamaño de paso en la discretización en tiempo; M representa los niveles de tiempo tal que $M\Delta t = T = 1$; TGS es la tolerancia en el método iterativo de Gauss-Seidel para resolver los sistemas lineales; IGS son las iteraciones necesarias para alcanzar la TGS y el Error se calcula como la norma infinito, es decir,

$$\text{Error} = \|u(T) - u_T\|_{\infty}.$$

El tiempo de máquina en una PC-486 para obtener la solución de este problema con los datos del inciso c) de la tabla anterior es de aproximadamente 140 seg. y aumenta conforme la discretización en espacio y tiempo es más fina.

La siguiente tabla muestra el comportamiento de este problema directo con respecto al parámetro β , los datos fueron generados con la discretización

$$h = \frac{1}{16}, \quad \Delta t = 0.01, \quad M = 100, \quad \text{y} \quad TGS = 10^{-12}.$$

β	IGS	ERROR
0.001	10	4.0488×10^{-4}
0.01	9	2.8080×10^{-4}
0.1	12	9.7166×10^{-5}
1	46	8.0136×10^{-6}
5	240	2.4683×10^{-6}
10	285	1.1375×10^{-6}
100	1130	1.0617×10^{-7}

Los errores en ambas tablas nos indican que se está resolviendo el problema directo en forma eficiente con un método combinado con error de truncamiento $o(h^2 + \Delta t^2)$.

La gráfica 3.1.1 al final de la sección muestra un corte de la solución exacta ($u(x, y, t) = \frac{10}{1+t}(y-1)^2$) y la solución aproximada para la discretización c) de la primera tabla.

Considerar el problema inverso asociado

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy, \quad \epsilon > 0$$

Sujeto a:

$$\begin{aligned} \frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u &= f, \quad \text{div}(W) = 0 && \text{en } \Omega \times (0, T) \\ au + \beta \frac{\partial u}{\partial n} &= ag && \text{en } \Gamma_1 \times (0, T) \\ u &= v && \text{en } \Gamma_c \times (0, T) \\ u(0) &= u_0(x, y) && \text{en } \Omega \end{aligned}$$

donde f , u_0 , a , \mathbf{w} y β están definidas como arriba. Para este problema consideramos que la distribución por alcanzar al tiempo $T = 1$ es

$$u_T(x, y) = 5(y-1)^2.$$

$$u_T(x, y) = 5(y-1)^2.$$

Para el cálculo del control óptimo (v) de este problema usamos el programa GRATRA.FOR, el cual es una implementación del método de Gradiente Conjugado analizado en la sección 3.2.

Los resultados numéricos obtenidos no son satisfactorios, ya que no se logra la convergencia del método, se puede decir que el algoritmo sufre un

estancamiento, es decir el criterio de paro del Gradiente Conjugado $\|g^k\|/\|g^0\|$ el cual se busca se anule no desciende más allá de una cota.

Para ejemplificar esto consideramos la corrida con los datos:

$$h = \frac{1}{32}, \quad \Delta t = \frac{1}{100}, \quad M = 100 \quad TGS = 10^{-12} \quad \epsilon = 10^{-4};$$

en este caso el proceso se estanca después de 8 iteraciones en el valor $\|g^k\|/\|g^0\| = 0.0803$.

La gráfica 3.1.2 muestra el control exacto ($v^* = \frac{10}{1+t}$) y el control calculado, así como un corte de la distribución u_T y la distribución calculada. Los errores para estos son:

$$\|v^* - v\|_\infty = 0.20$$

$$\|u(T) - u_T\|_\infty = 0.48$$

El tiempo utilizado en una PC-486 es de aproximadamente 25 min., esto se debe a que cada iteración de Gradiente conjugado requiere resolver 3 problemas directos como el descrito al principio de este ejemplo.

Ejemplo 2.

Consideremos el problema directo (3.1) con los siguientes datos:

$$f(x, y, t) = x^2(x-1)^2 + (y-1)^2 - 4\delta(1+t)(3x^2 - 3x + 1) \\ + 2(1+t)x(x-1)(2x-1);$$

$$v(x, t) = (1+t)(1+x^2(x-1)^2);$$

$$u_0(x, y) = x^2(x-1)^2 + (y-1)^2$$

Este problema se resolvió con el programa RECUTRA.FOR usando los siguientes datos:

$$a = 0, \quad \mathbf{w} = (1, 0), \quad \beta = 1 \quad \text{y hasta } T = 1,$$

la siguiente es una tabla de resultados para diferentes discretizaciones.

	h	Δt	M	TGS	IGS	$ERROR$
a	$\frac{1}{10}$	0.1	10	10^{-12}	114	8.411×10^{-2}
b	$\frac{1}{16}$	0.01	100	10^{-12}	42	3.5625×10^{-2}
c	$\frac{1}{32}$	0.01	100	10^{-12}	137	9.5047×10^{-3}
d	$\frac{1}{40}$	0.01	100	10^{-12}	206	6.1610×10^{-3}

El tiempo de máquina en una PC-486 es para la discretización d) de aproximadamente 140 seg., ahora consideramos en la siguiente tabla el efecto de variar el parámetro β para la discretización en c).

β	IGS	ERROR
0.001	12	8.2793×10^{-3}
0.01	9	8.3238×10^{-3}
0.1	23	8.6868×10^{-3}
1	137	9.5047×10^{-3}
5	562	1.0210×10^{-2}
10	1005	1.3714×10^{-2}
100	1670	1.9871×10^{-2}

Los errores en ambas tablas son satisfactorias considerando la complejidad de las funciones f y u_0 , además de que el método combinado tiene un error de truncamiento $o(h^2 + \Delta t^2)$. La gráfica 3.2.1 al final de la sección muestra los cortes de la solución exacta ($u(x, y, t) = (t + 1)(x^2(x - 1)^2 + (y - 1)^2)$) y la solución aproximada para la discretización en c) de la primera tabla.

Considerar el problema inverso como en el problema 1 de esta sección con los datos de a , f , u_0 , \mathbf{w} y β dados como arriba. Para este problema consideramos que la distribución por alcanzar el tiempo $T = 1$ es:

$$u_T(x, y) = 2[x^2(x - 1)^2 + (y - 1)^2].$$

Para el cálculo del control óptimo (v) se utilizó el programa GRACON.FOR.

Como en el problema 1 los resultados numéricos no son satisfactorios ya que nuevamente el algoritmo sufre un estancamiento, para los datos:

$$h = \frac{1}{32}, \quad \Delta t = \frac{1}{100}, \quad M = 100 \quad TGS = 10^{-12} \quad \epsilon = 10^{-4}$$

tenemos que el proceso se estanca después de 15 iteraciones en el valor $\|g^k\|/\|g^0\| = 0.079$.

La gráfica 3.2.2 muestra el control exacto ($v^* = (1 + t)(1 + x^2(1 - x)^2)$) y el control calculado así como los cortes de distribución u_T y la distribución calculada. Los errores para éstos son:

$$\|v^* - v\|_\infty = 0.24$$

$$\|u_T - u(T)\|_\infty = 0.43$$

El tiempo de máquina utilizado en esta corrida es aproximadamente 27 min.

Estos dos ejemplos muestran que existe una deficiencia en la solución de los problemas inversos, pero por los resultados del capítulo II sabemos que estos no corresponden a la solución de los problemas directos, sino al cálculo del gradiente para la aplicación del método de gradiente conjugado:

$$g^{(k)} = v^{(k)} - \frac{\partial P}{\partial n} \quad \text{en } \Gamma_c \times (0, T).$$

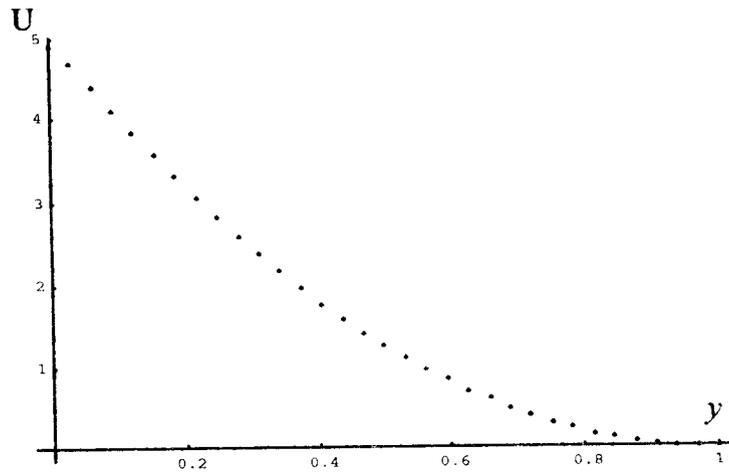
El cálculo P es correcto ya que es solución de un problema directo (ecuación adjunta), el problema está en el cálculo de $\partial P/\partial n$ ya que esta derivada normal se aproxima con fórmulas de diferencias finitas con puntos hacia adelante, las cuales han demostrado ser inestables y producen pérdida de cifras en los datos. El efecto en el proceso de gradiente conjugado es el cálculo incorrecto de la dirección de máximo descenso, en las conclusiones abordamos las alternativas para éste cálculo.

La siguiente tabla (correspondiente al ejemplo 1) muestra el efecto de calcular $\partial P/\partial n$ con diferencias finitas de tres puntos:

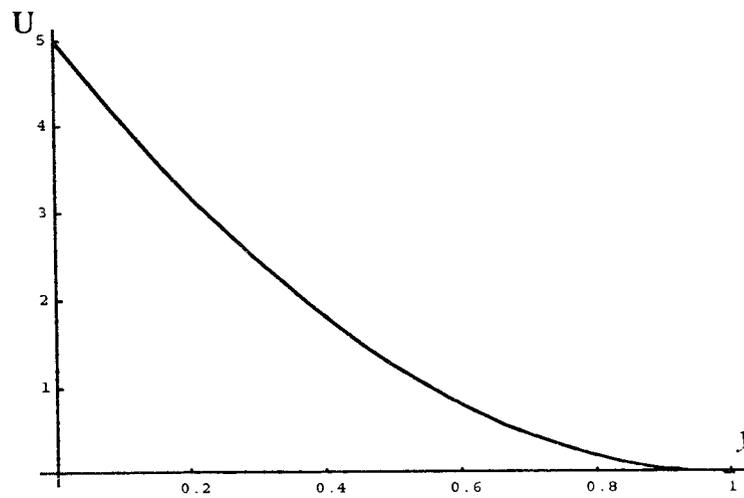
$$\frac{\partial P}{\partial n} = \nabla P \cdot n = -\frac{\partial P}{\partial y} = -\frac{2P_{i2} - \frac{1}{2}P_{i3} - \frac{3}{2}P_{i1}}{h} \quad i = 0, N = 1$$

h	Δt	M	Error del problema directo($T = 1$)	Error en la derivada normal($T = 1$)
$\frac{1}{32}$	$\frac{1}{100}$	100	8.0224×10^{-6}	5.0017×10^{-3}
$\frac{1}{40}$	$\frac{1}{100}$	100	8.0136×10^{-6}	6.813×10^{-3}
$\frac{1}{80}$	$\frac{1}{100}$	100	7.403×10^{-6}	1.389×10^{-2}
$\frac{1}{100}$	$\frac{1}{100}$	100	3.171×10^{-6}	7.378×10^{-2}

PROBLEMA DIRECTO: ECUACION DE DIFUSION-ADVECCION



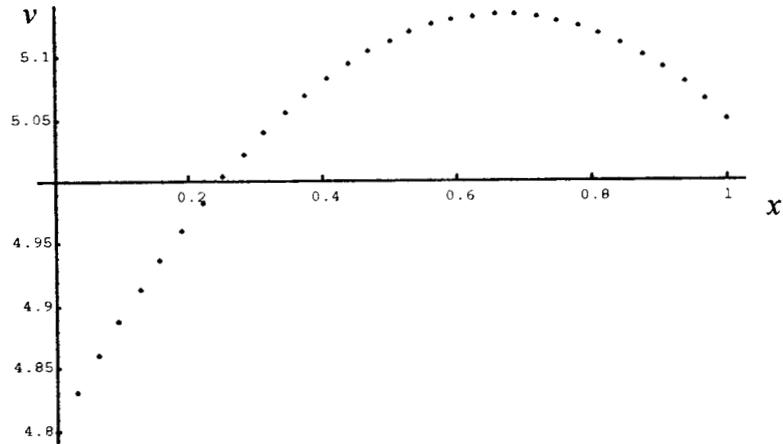
SOLUCION APROXIMADA (Corte para $x=1/2$ y $T=1$)



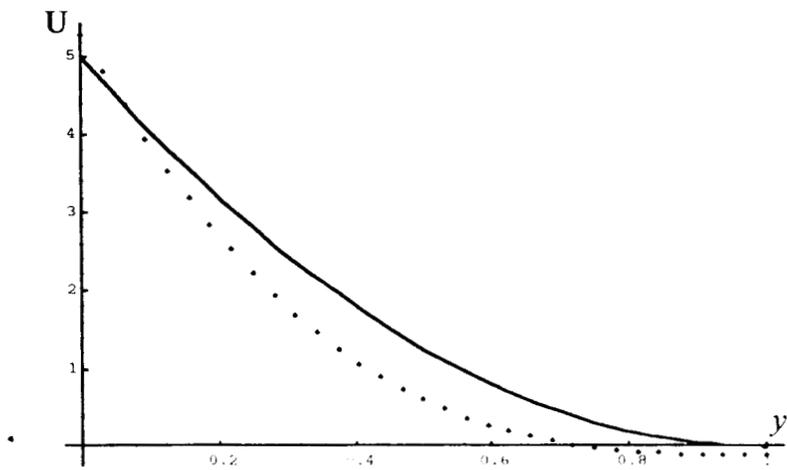
SOLUCION EXACTA (Corte para $x=1/2$ y $T=1$)

Grafica 3.1.1

PROBLEMA INVERSO: ECUACION DE DIFUSION-ADVECCION



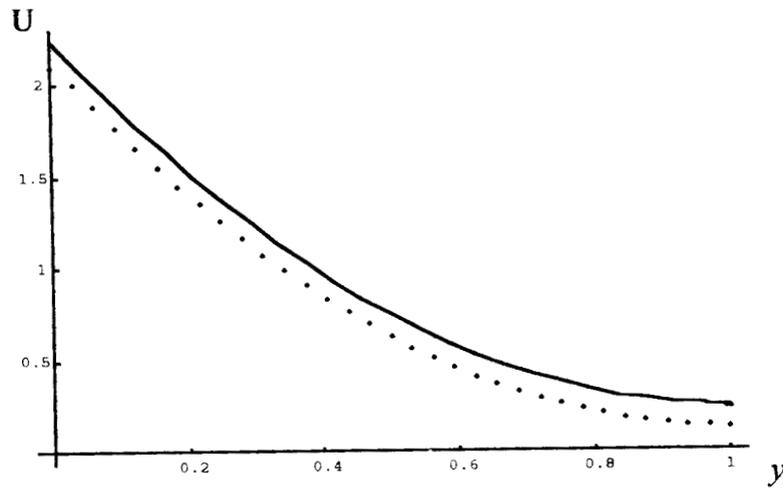
**CONTROL EXACTO ($v=5$) vs CONTROL APROXIMADO
(Corte para $T=1$)**



**ESTADO EXACTO ($T=1$) vs ESTADO APROXIMADO
(Corte para $x=1/2$)**

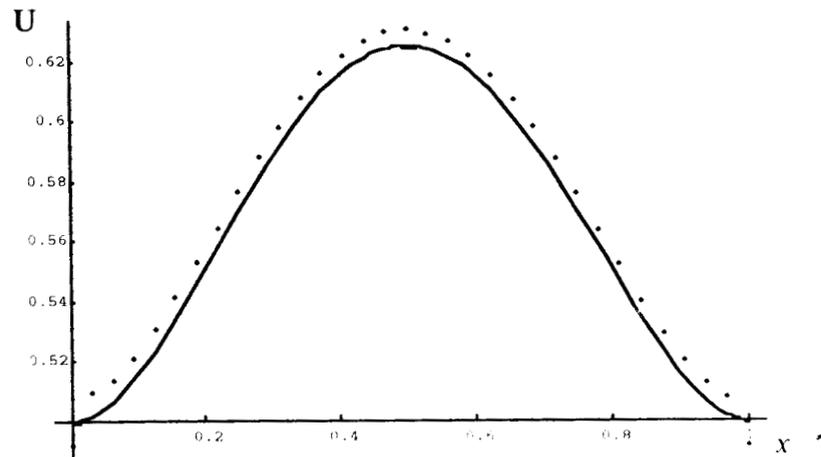
Grafica 3.1.2

PROBLEMA DIRECTO: ECUACION DE DIFUSION-ADVECCION



SOLUCION EXACTA vs APROXIMADA

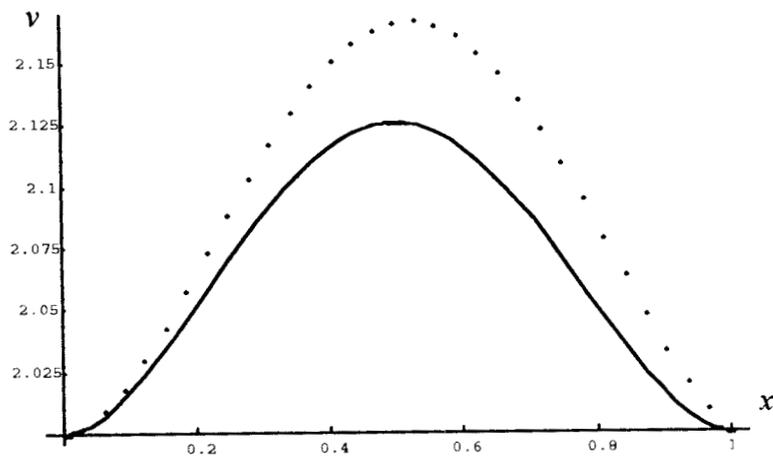
(Corte para $x=1/2$ y $T=1$)



SOLUCION EXACTA vs APROXIMADA

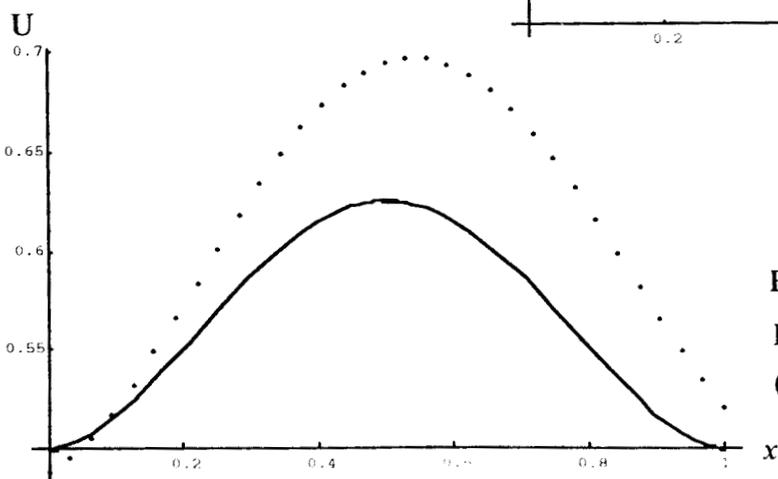
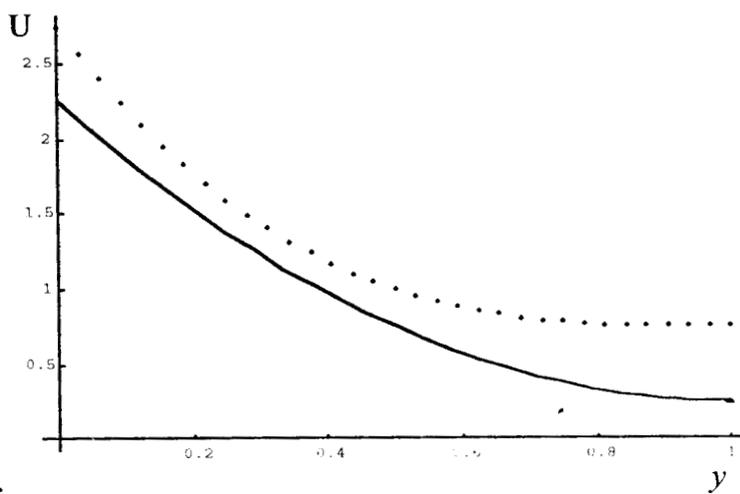
(Corte para $y=1/2$ y $T=1$)

PROBLEMA INVERSO: ECUACION DE DIFUSION-ADVECCION



**CONTROL
EXACTO vs
CONTROL
APROXIMADO
(Corte para $T=1$)**

**ESTADO EXACTO ($T=1$) vs
ESTADO APROXIMADO
(Corte para $x=1/2$)**



**ESTADO EXACTO ($T=1$) vs
ESTADO APROXIMADO
(Corte para $y=1/2$)**

Grafica 3.2.2

CAPITULO IV

4 Control en la Frontera Tipo Neumann para la Ecuación de Transporte

En los capítulos anteriores se trató de controlar el proceso de difusión y difusión-advención a través de un control v del tipo Dirichlet ($u = v$ en $\Gamma_c \times (0, T)$), en este capítulo el control v es tipo Neumann

$$\left(\frac{\partial u}{\partial n} = v \text{ en } \Gamma_c \times (0, T) \right),$$

por lo cual controlamos el flujo de calor a través de una sección de frontera y no la temperatura. Este cambio no modifica la funcional de Energía (1.5) que se desea minimizar, es decir, el problema de control óptimo planteado en el capítulo I es el mismo:

Sea $u_T(x, y)$, $(x, y) \in \Omega$ y $T > 0$ una distribución de temperatura en el dominio Ω . Queremos hallar $v \in L^2(\Gamma_c \times (0, T))$ tal que

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(t) - u_T)^2 dx dt, \quad \epsilon > 0$$

Sujeto a:

$$(4.1) \quad \left\{ \begin{array}{ll} \frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u = f & \text{en } \Omega \times (0, T) \\ au + \beta \frac{\partial u}{\partial n} = ag & \text{en } \Gamma_1 \times (0, T) \\ \frac{\partial u}{\partial n} = v & \text{en } \Gamma_c \times (0, T) \\ u(0) = u_0(x, y) & \text{en } \Omega, \end{array} \right.$$

donde $\text{div}(\mathbf{w}) = 0$.

Sobre este problema de control óptimo asumimos dos cosas, la primera es que para cada $v \in L^2(\Gamma_c \times (0, T))$ existe una función $u(x, y, t)$ definida para $(x, y) \in \Omega$ y $0 < t < T$ que es solución del problema (4.1), los datos

f, g y $u_0(x, y)$ los consideramos cuadrado integrables en su dominio. La segunda consideración es que respecto del control v hay continuidad, es decir, si $v_1 \rightarrow v_2$ entonces $u_1 \rightarrow u_2$. Con estas consideraciones es posible demostrar que el problema de control óptimo tiene solución única v^* en $L^2(\Gamma_c \times (0, T))$, el argumento es el mismo que el mostrado en el capítulo I. Para calcular el control óptimo v^* minimizamos la funcional de Energía (1.5) con el método de gradiente conjugado; este método requiere de la primera variación de la funcional y su gradiente, el cual en el caso de control Neumann sí es diferente por lo cual procedemos a su cálculo.

4.1 Cálculo de la primera variación

Se define la primera variación como:

$$\delta J[v; \delta v] = \lim_{\lambda \rightarrow 0} \frac{J[v + \lambda \delta v] - J[v]}{\lambda}.$$

Por linealidad del problema (4.1) al control $v + \lambda \delta v$ le corresponde el control $u + \lambda \delta u$, donde δu es solución de

$$(4.2) \quad \left\{ \begin{array}{l} \frac{\partial \delta u}{\partial t} - \beta \Delta \delta u + \mathbf{w} \cdot \nabla \delta u = 0 \quad \text{en } \Omega \times (0, T) \\ a \delta u + \beta \frac{\partial \delta u}{\partial n} = 0 \quad \text{en } \Gamma_1 \times (0, T) \\ \frac{\partial \delta u}{\partial n} = \delta v \quad \text{en } \Gamma_c \times (0, T) \\ \delta u(0) = 0 \quad \text{en } \Omega. \end{array} \right.$$

Por otro lado,

$$\begin{aligned} J[v + \lambda \delta v] &= \frac{1}{2} \int_0^T \int_{\Gamma_c} (v + \lambda \delta v)^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) + \lambda \delta u(T) - u_T)^2 dx dy \\ &= \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 + \lambda \int_0^T \int_{\Gamma_c} v \delta v d\gamma dt + \frac{\lambda^2}{2} \int_0^T \int_{\Gamma_c} \delta^2 v d\gamma dt + \\ &+ \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 + \frac{\lambda}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(t) + \frac{\lambda^2}{2\epsilon} \int_{\Omega} \delta^2 u(T) \end{aligned}$$

$$\begin{aligned} \frac{J[v + \lambda \delta v] - J[v]}{\lambda} &= \int_0^T \int_{\Gamma_c} v \delta v dt + \frac{\lambda}{2} \int_0^T \int_{\Gamma_c} \delta^2 v d\gamma dt \\ &\quad + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) + \frac{\lambda}{2\epsilon} \int_{\Omega} \delta^2 u(T) \end{aligned}$$

tomando en ambos lados límite cuando $\lambda \rightarrow 0$ obtenemos:

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} v \delta v dt + \frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) \quad (4.3)$$

La ecuación (4.3) es correcta pero poco útil ya que contiene integrales de diferentes espacios, lo que se busca es una sola integral (producto interior en $\Gamma_c \times (0, T)$) con el fin de deducir la expresión para el gradiente g , la cual será útil en la aplicación del método de gradiente conjugado.

Es decir, requerimos g tal que

$$\delta J[v; \delta v] = (g, \delta v)_{L^2(\Gamma_c \times (0, T))} = \int_0^T \int_{\Gamma_c} g \delta v d\gamma dt$$

el cual existe de forma única por el teorema de representación de Riez. Ahora procedemos a su cálculo.

4.2 Cálculo de la Ecuación Adjunta

Sea $P = P(x, y, t)$ cuyas características se determinan en lo que sigue:

Multiplicando (4.2) por P e integrando, tenemos

$$\int_{\Omega} \int_0^T P \frac{\partial \delta u}{\partial t} - \beta \int_{\Omega} \int_0^T P \Delta \delta u + \int_{\Omega} \int_0^T P \mathbf{w} \cdot \nabla \delta u = 0.$$

Integrando por partes y usando la segunda fórmula de Green

$$\begin{aligned} \int_{\Omega} \left[P \delta u \Big|_0^T - \int_0^T \delta u \frac{\partial P}{\partial t} \right] - \beta \int_0^T \left[\int_{\Omega} \delta u \Delta P + \int_{\Gamma} \left[P \frac{\partial \delta u}{\partial u} - \delta u \frac{\partial P}{\partial n} \right] \right] \\ + \int_{\Omega} \int_0^T P \mathbf{w} \cdot \nabla \delta u = 0, \end{aligned}$$

tomando en cuenta que $\delta u(0) = 0$, se tiene

$$\begin{aligned} \int_{\Omega} P(T)\delta u(T) + \int_{\Omega} \int_0^T \delta u \left(-\frac{\partial P}{\partial t} - \beta \Delta P \right) - \beta \int_0^T \int_{\Gamma} \left[P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right] \\ + \int_{\Omega} \int_0^T P \mathbf{w} \cdot \nabla \delta u = 0 \end{aligned}$$

De $\mathbf{w} \cdot \nabla(P\delta u) = P \mathbf{w} \cdot \nabla \delta u + \delta u \mathbf{w} \cdot \nabla P$, tenemos

$$\int_{\Omega} \int_0^T P \mathbf{w} \cdot \nabla \delta u = \int_{\Omega} \int_0^T \mathbf{w} \cdot \nabla(P\delta u) - \int_{\Omega} \int_0^T \delta u \mathbf{w} \cdot \nabla P.$$

Entonces,

$$\begin{aligned} \int_{\Omega} P(T)\delta u(T) + \int_{\Omega} \int_0^T \delta u \left(-\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P \right) \\ - \beta \int_0^T \int_{\Gamma} \left[P \frac{\partial \delta u}{\partial n} - \delta u \frac{\partial P}{\partial n} \right] + \int_{\Omega} \int_0^T \mathbf{w} \cdot \nabla(P\delta u) = 0, \end{aligned}$$

pedimos a P que cumpla

$$-\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P = 0 \quad \text{en } \Omega \times (0, T),$$

luego

$$\begin{aligned} \int_{\Omega} P(T)\delta u(T) - \beta \int_0^T \int_{\Gamma_c} P \frac{\partial \delta u}{\partial n} - \int_0^T \int_{\Gamma_1} P \beta \frac{\partial \delta u}{\partial n} \\ + \beta \int_0^T \int_{\Gamma_c} \delta u \frac{\partial P}{\partial n} + \beta \int_0^T \int_{\Gamma_1} \delta u \frac{\partial P}{\partial n} \\ + \int_{\Omega} \int_0^T \mathbf{w} \cdot \nabla(P\delta u) = 0 \end{aligned}$$

$$\begin{aligned} \int_{\Omega} P(T)\delta u(T) - \beta \int_0^T \int_{\Gamma_c} P \frac{\partial \delta u}{\partial n} + \beta \int_0^T \int_{\Gamma_c} \delta u \frac{\partial P}{\partial n} \\ + \int_0^T \int_{\Gamma_1} \delta u \left(aP + \beta \frac{\partial P}{\partial n} \right) + \int_{\Omega} \int_0^T \mathbf{w} \cdot \nabla(P\delta u) = 0 \end{aligned}$$

$$\operatorname{div}(\mathbf{w}(P\delta u)) = \mathbf{w} \cdot \nabla(P\delta u) + P\delta u \operatorname{div}(\mathbf{w}) = \mathbf{w} \cdot \nabla(P\delta u),$$

ya que $\operatorname{div}(\mathbf{w}) = 0$.

$$\Rightarrow \int_{\Omega} \int_0^T \mathbf{w} \cdot \nabla(P\delta u) = \int_0^T \int_{\Omega} \operatorname{div}(\mathbf{w}(P\delta u)) = \int_0^T \int_{\Gamma} P\delta u \mathbf{w} \cdot \mathbf{n}.$$

Entonces

$$\begin{aligned} \int_{\Omega} P(T)\delta u(T) - \beta \int_0^T \int_{\Gamma_c} P \frac{\partial \delta u}{\partial n} + \beta \int_0^T \int_{\Gamma_c} \delta u \frac{\partial P}{\partial n} \\ + \int_0^T \int_{\Gamma_1} \delta u \left(aP + \beta \frac{\partial P}{\partial n} + P\mathbf{w} \cdot \mathbf{n} \right) + \int_0^T \int_{\Gamma_c} P\delta u \mathbf{w} \cdot \mathbf{n} = 0. \end{aligned}$$

Pedimos que

$$aP + \beta \frac{\partial P}{\partial n} + P\mathbf{w} \cdot \mathbf{n} = 0 \quad \text{en } \Gamma_1 \times (0, T).$$

$$\int_{\Omega} P(T)\delta u(T) + \int_0^T \int_{\Gamma_c} -\beta P\delta v + \int_0^T \int_{\Gamma_c} \beta \delta u \frac{\partial P}{\partial n} + \int_0^T \int_{\Gamma_c} P\delta u \mathbf{w} \cdot \mathbf{n} = 0$$

$$\int_{\Omega} P(T)\delta u(T) + \int_0^T \int_{\Gamma_c} \delta u \left(\beta \frac{\partial P}{\partial n} + P\mathbf{w} \cdot \mathbf{n} \right) - \int_0^T \int_{\Gamma_c} \beta P\delta v = 0$$

pedimos que $\beta \frac{\partial P}{\partial n} + P\mathbf{w} \cdot \mathbf{n} = 0$ en $\Gamma_c \times (0, T)$ y $P(T) = \frac{1}{\epsilon}(u(T) - u_T)$ en Ω , entonces tenemos

$$\frac{1}{\epsilon} \int_{\Omega} (u(T) - u_T) \delta u(T) - \int_0^T \int_{\Gamma_c} \beta P\delta v = 0 \quad (4.4)$$

restando (4.3) y (4.4) tenemos la expresión adecuada para la primera variación.

$$\delta J[v; \delta v] = \int_0^T \int_{\Gamma_c} (v + \beta P)\delta v = (v + \beta P, \delta v)_{L^2(\Gamma_c \times (0, T))} \quad (4.5)$$

$\Rightarrow g = v + \beta P$ en $\Gamma_c \times (0, T)$.

Donde P se calcula como solución del problema adjunto.

$$(4.6) \quad \left\{ \begin{array}{ll} -\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P = 0 & \text{en } \Omega \times (0, T) \\ aP + \beta \frac{\partial P}{\partial n} + P \mathbf{w} \cdot \mathbf{n} = 0 & \text{en } \Gamma_1 \times (0, T) \\ \beta \frac{\partial P}{\partial n} + P \mathbf{w} \cdot \mathbf{n} = 0 & \text{en } \Gamma_c \times (0, T) \\ P(T) = \frac{1}{c} (u(T) - u_T) & \text{en } \Omega. \end{array} \right.$$

Ya que este es un problema de valores finales, se escribe como un problema de valores iniciales:

$$\tau = T - t$$

$$q(x, y, \tau) := P(x, y, T - \tau) \quad \text{i.e.} \quad q(0) = P(T)$$

$$\frac{\partial q}{\partial \tau} = \frac{\partial P}{\partial t} \cdot \frac{\partial t}{\partial \tau} = -\frac{\partial P}{\partial t},$$

y las derivadas en espacio son iguales, entonces obtenemos:

$$(4.7) \quad \left\{ \begin{array}{ll} \frac{\partial q}{\partial \tau} - \beta \Delta q - \mathbf{w} \cdot \nabla q = 0 & \text{en } \Omega \times (0, T) \\ aq + \beta \frac{\partial q}{\partial n} + q \mathbf{w} \cdot \mathbf{n} = 0 & \text{en } \Gamma_1 \times (0, T) \\ \beta \frac{\partial q}{\partial n} + q \mathbf{w} \cdot \mathbf{n} = 0 & \text{en } \Gamma_c \times (0, T) \\ q(0) = \frac{1}{c} (u(T) - u_T) & \text{en } \Omega. \end{array} \right.$$

Se deben notar dos cosas muy importantes, la primera es que el cálculo del gradiente $g = v + \beta P$ ya no involucra como en los problemas de control tipo Dirichlet a la derivada normal de P , esto es una gran ventaja en la solución numérica ya que no se pierde precisión por el cálculo de la derivada con diferencias finitas, lo cual como se ha visto en los capítulos anteriores fué determinante en el control del sistema. Podemos decir que la precisión

de los datos numéricos en el cálculo de g es la misma precisión con que se resuelve el problema directo (4.6) para obtener P . Esto lleva a la búsqueda de un algoritmo eficiente para resolver los problemas directos y de esta forma obtener la mayor precisión en el cálculo del gradiente, en la sección 4.4 se analiza más este punto.

El segundo punto por notar es que cada vez que sea necesario resolver el problema (4.6) se resuelve el problema (4.7) para obtener $q(x, y, \tau)$, a continuación aplicamos el cambio de variable $P(x, y, t) = q(x, y, T - t)$ para obtener P .

4.3 Aplicación del Método Gradiente Conjugado

En el capítulo I se mostró que la funcional de Energía (1.5) es estrictamente convexa y con primera derivada continua, además el mínimo v^* (control óptimo) existe de forma única, por lo cual el algoritmo siguiente genera una sucesión $\{v^n\}$ en $L^2(\Gamma_c \times (0, T))$ que converge a v^* . El criterio de paro del método es la anulación del gradiente, lo cual para una funcional estrictamente convexa es condición suficiente para el óptimo (Teorema 1.6.2).

Tomar $v^0 \in L^2(\Gamma_c \times (0, T))$ (por ejemplo $v^0 \equiv 0$), definimos

$$H = L^2(\Gamma_c \times (0, T))$$

Paso 1) Hallar $g^0 \in H$ tal que $\forall \delta v \in H$

$$(g^0, \delta v) = \delta J[v^0, \delta v].$$

En la sección 4.2 se mostró que $g^0 = v^0 + \beta P$, este cálculo se hace en tres pasos

1.i) Resolver para obtener $u^0(T)$

$$-\frac{\partial u^0}{\partial t} - \beta \Delta u^0 + \mathbf{w} \cdot \nabla u^0 = f \quad \text{en } \Omega \times (0, T)$$

$$au^0 + \beta \frac{\partial u^0}{\partial n} n = ag \quad \text{en } \Gamma_1 \times (0, T)$$

$$\frac{\partial u^0}{\partial n} = v^0 \quad \text{en } \Gamma_c \times (0, T)$$

$$u^0(0) = u^0(x, y) \quad \text{en } \Omega \quad ,$$

es decir, resolvemos la ecuación (4.1) con el control v^0 .

1.ii) Resolvemos para obtener P

$$-\frac{\partial P}{\partial t} - \beta \Delta P - \mathbf{w} \cdot \nabla P = 0 \quad \text{en } \Omega \times (0, T)$$

$$aP + \beta \frac{\partial P}{\partial n} + P \mathbf{w} \cdot n = 0 \quad \text{en } \Gamma_1 \times (0, T)$$

$$\beta \frac{\partial P}{\partial n} + P \mathbf{w} \cdot n = 0 \quad \text{en } \Gamma_c \times (0, T)$$

$$P(T) = \frac{1}{c} (u^0(T) - u_T) \quad \text{en } \Omega \quad ,$$

es decir, resolvemos la ecuación adjunta (4.6) con la solución $u^0(T)$ calculada en 1.i).

1.iii) Hacemos la asignación

$$g^0 = v^0 + \beta P$$

donde P solo se considera en $\Gamma_c \times (0, T)$.

Paso 2) Hacemos la asignación

$$w^0 = g^0$$

la cual define la dirección de descenso, hasta aquí tenemos calculados v^0, g^0 y ω^0 .

Para $n \geq 0$ calculamos v^{n+1}, g^{n+1} y ω^{n+1} de v^n, g^n y ω^n de la siguiente forma.

Paso 3) Hallar $\rho_n \in \mathbb{R}$ que minimice

$$J[v^n - \rho \omega^n] \quad \forall \rho \in \mathbb{R},$$

el cálculo de ρ_n se hace en dos pasos.

3.i) Calcular z^n como sigue

$$\frac{\partial z^n}{\partial t} - \beta \Delta z^n + \mathbf{w} \cdot \nabla z^n = 0 \quad \text{en } \Omega \times (0, T)$$

$$az^n + \beta \frac{\partial z^n}{\partial n} = 0 \quad \text{en } \Gamma_1 \times (0, T)$$

$$\frac{\partial z^n}{\partial n} = \omega^n \quad \text{en } \Gamma_c \times (0, T)$$

$$z^n(0) = 0 \quad \text{en } \Omega,$$

3.ii)

$$\rho_n = \frac{\int_0^T \int_{\Gamma_c} v^n \omega^n d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (u^n - u_T) z^n dx dy}{\int_0^T \int_{\Gamma_c} (\omega^n)^2 d\gamma dt + \frac{1}{\epsilon} \int_{\Omega} (z^n)^2 dx dy}$$

ρ_n define el punto de minimo sobre la línea $(v^n - \rho_n \omega^n)$ de máximo descenso.

Paso 4) Se define el nuevo control

$$v^{n+1} = v^n - \rho_n \omega^n.$$

Paso 5) Hallar $g^{n+1} \in H$ tal que $\forall \delta v \in H$

$$(g^{n+1}, \delta v) = \delta J[v^{n+1}, \delta v].$$

5.i) Resolvemos (4.1) con el control v^{n+1} para obtener $u^{n+1}(T)$.

5.ii) Resolvemos la ecuación adjunta (4.6) con $u^{n+1}(T)$, para obtener P .

5.iii) Hacemos la asignación

$$g^{n+1} = v^{n+1} + \beta P \quad \text{en } \Gamma_c \times (0, T).$$

Si $\|g^{n+1}\|/\|g^0\| \leq \text{Tol}$, entonces $v^* = v^{n+1}$, de otra forma:

Paso 6) $\gamma = \frac{(g^{n+1}, g^{n+1})}{(g^n, g^n)}$

Paso 7) $\omega^{n+1} = g^{n+1} + \gamma\omega^n,$

Paso 8) Hacer $n \leftarrow n + 1$ y regresar a 3).

Para la aplicación numérica de este algoritmo y por lo tanto para la solución del problema inverso, se necesitan sólo dos tipos de programas, uno que resuelva problemas directos del tipo de la ecuación (4.1) y otro que calcule productos interiores (integrales dobles), el detalle de estos se analiza en la siguiente sección.

4.4 Discretización y Resultados Numéricos

Para la discretización y experimentos numéricos de este capítulo se tomó el problema de difusión, es decir, consideramos $\beta = 1$, $\mathbf{w} = (0, 0)$ y $a = 0$ en la ecuación (4.1) y (4.6). La semidiscretización en el caso general es semejante a lo desarrollado en el capítulo III (sección 3.3).

Por las experiencias numéricas en el control tipo Dirichlet desarrolladas en los capítulos II y III, sabemos que para resolver el problema inverso es muy importante resolver los problemas directos con la mayor precisión posible. Las discretizaciones en tiempo y espacio empleadas para los problemas directos en esos capítulos fueron de Diferencias Finitas, con un error de truncamiento de orden dos. Estas fórmulas fueron suficientes para obtener buenas aproximaciones, por lo cual se vuelven a utilizar en este capítulo a través de una rutina eficiente que a continuación se describe.

La subrutina DEPX4 (FISHPAK) es una subrutina que resuelve ecuaciones diferenciales parciales elípticas (separables); usa aproximaciones por diferencias finitas de orden dos y orden cuatro en dominios rectangulares, es decir, encuentra $u(x, y)$ con $A \leq x \leq B$ y $C \leq y \leq D$ que es solución de:

$$A(x)\frac{\partial^2 u}{\partial x^2} + B(x)\frac{\partial u}{\partial x} + C(x)u + \frac{\partial^2 u}{\partial y^2} = G(x, y) \quad (4., 8)$$

Las condiciones de frontera en $X = A$, $X = B$, $Y = C$ y $Y = D$, pueden ser de tipo periódico, Dirichlet, Neumann o mixtas (Fourier).

La discretización de la ecuación parcial y sus condiciones de frontera las realiza la subrutina SEPX4, con esto se obtiene el sistema lineal cuya solución

es el vector $\{u_{ij}\}$ que aproxima los valores de la función $u(x_i, y_j)$ en una malla sobre el rectángulo $A \leq x \leq B$ y $C \leq y \leq D$. El sistema lineal se resuelve por un método iterativo eficiente conocido como "Reducción cíclica generalizada", una referencia de este método se encuentra en [11]. La documentación de la subrutina DEPX4 se encuentra en el apéndice 2.

Ya que DEPX4 resuelve en forma eficiente problemas elípticos, se ocupa para resolver los problemas directos (4.1) y (4.6), debido a que (4.1) es más general que (4.6) sólo se resuelve (4.1), aplicando el programa que resulta a (4.6).

Para resolver (4.1) discretizamos en tiempo con la fórmula de Gear, pasando el problema parabólico a una sucesión de problemas elípticos cuya solución en cada nivel de tiempo se obtiene con DEPX4. De hecho como la fórmula de Gear determina un esquema de dos pasos, es necesario, para obtener la solución en el primer nivel de tiempo, un esquema de un paso, en este caso usamos la fórmula de Euler hacia atrás. Una vez discretizado en tiempo es necesario escribir los problemas elípticos en la forma (4.8) para aplicar DEPX4, presentamos las dos discretizaciones.

Sea $\Delta t > 0$ el paso de tiempo y M los niveles de tiempo, tal que $M\Delta t = T$, los niveles de tiempo se difinen como $t_{k+1} = (k+1)\Delta t$.

Semidiscretización con Euler.

$$k = 0, 1, \dots, M-1.$$

$$\frac{\partial^2 u^{k+1}}{\partial x^2} - \frac{1}{\Delta t} u^{k+1} + \frac{\partial^2 u^{k+1}}{\partial y^2} = -f^{k+1} - \frac{1}{\Delta t} u^k \quad \text{en } \Omega$$

$$\frac{\partial u^{k+1}}{\partial x} = 0 \quad \text{en } X = 0$$

$$\frac{\partial u^{k+1}}{\partial x} = 0 \quad \text{en } X = 1$$

$$-\frac{\partial u^{k+1}}{\partial y} = v^{k+1} \quad \text{en } Y = 0$$

$$\frac{\partial u^{k+1}}{\partial y} = 0 \quad \text{en } Y = 1$$

$$u^0 = u_0(x, y).$$

Este esquema se aplica una vez ($k = 0$) para encontrar u^1 .

Semidiscretización con Gear.

$k = 1, 2, \dots, M - 1$.

$$\frac{\partial^2 u^{k+1}}{\partial x^2} - \frac{3}{2\Delta t} u^{k+1} + \frac{\partial^2 u^{k+1}}{\partial y^2} = -f^{k+1} - \frac{2}{\Delta t} u^k + \frac{1}{2\Delta t} u^{k-1} \quad \text{en } \Omega$$

$$\frac{\partial u^{k+1}}{\partial x} = 0 \quad \text{en } X = 0$$

$$\frac{\partial u^{k+1}}{\partial x} = 0 \quad \text{en } X = 1$$

$$-\frac{\partial u^{k+1}}{\partial y} = v^{k+1} \quad \text{en } Y = 0$$

$$\frac{\partial u^{k+1}}{\partial y} = 0 \quad \text{en } Y = 1$$

$u^0 = u_0(x, y)$ y u^1 es la solución calculada con Euler.

La subrutina ECALFN.F en FORTRAN 77 que resuelve de acuerdo a los esquemas anteriores el problema directo 4.1 aparece en el Apéndice 2.

Para aplicar el algoritmo de la sección 4.3 (gradiente conjugado) es necesario calcular productos interiores (integrales dobles), estos se calculan por el método de trapecio en dos dimensiones con la subrutina PROINT.F en FORTRAN 77 que aparece en el Apéndice 2.

El programa que resuelve el problema inverso (gradiente conjugado) usa a ECALFN.F y PROINT.F de acuerdo al algoritmo en la sección 4.3, aparece en el Apéndice 2 con el nombre de GRAFN.F. Las aplicaciones de este programa se hacen a continuación.

Ejemplo 1.

Consideramos el problema directo (4.1) con los datos antes mencionados. Se definen f , v y u_0 como sigue:

$$f(x, y, t) = \frac{-10}{(1+t)^2} (y-1)^2 - \frac{20}{(1+t)}$$

$$v(x, t) = \frac{20}{1+t},$$

$$u_0(x, y) = 10(y-1)^2.$$

Este problema se resolvió usando el programa ECALFN.F el cual se describió arriba, la siguiente tabla muestra los resultados:

	h	Δt	M	T	$ERROR$
a	$\frac{1}{10}$	$\frac{1}{10}$	10	1.0	2.7313×10^{-2}
b	$\frac{1}{16}$	$\frac{1}{20}$	20	1.0	7.2656×10^{-3}
c	$\frac{1}{32}$	$\frac{1}{100}$	100	1.0	3.0728×10^{-4}
d	$\frac{1}{32}$	$\frac{1}{500}$	500	1.0	1.2435×10^{-5}
e	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	3.1134×10^{-6}
f	$\frac{1}{64}$	$\frac{1}{100}$	100	1.0	3.0728×10^{-4}

Donde $h > 0$ es el tamaño de paso en espacio; $\Delta t > 0$ es el tamaño de paso en tiempo; M representa los niveles de tiempo de tal forma que $M\Delta t = T$, en esta caso se calcula la solución hasta $T = 1$, finalmente $ERROR = \|u_T - u(T)\|_\infty$.

Estos ejemplos se corrieron en una máquina PC-486, el tiempo de cálculo para la discretización e) de la tabla anterior es aproximadamente 4 min.

Estos resultados numéricos son satisfactorios y proporcionan una base confiable para la solución de los problemas inversos.

Ahora consideremos el problema inverso asociado, es decir,

$$\text{Min}J[v] = \frac{1}{2} \int_0^T \int_{\Gamma_c} v^2 d\gamma dt + \frac{1}{2\epsilon} \int_{\Omega} (u(T) - u_T)^2 dx dy, \quad \epsilon > 0,$$

sujeto a:

$$\frac{\partial u}{\partial t} - \beta \Delta u + \mathbf{w} \cdot \nabla u = f \quad \text{en } \Omega \times (0, T)$$

$$au + \beta \frac{\partial u}{\partial n} = ag \quad \text{en } \Gamma_1 \times (0, T)$$

$$\frac{\partial u}{\partial n} = v \quad \text{en } \Gamma_c \times (0, T)$$

$$u(0) = u_0(x, y) \quad \text{en } \Omega$$

donde $\text{div}(\mathbf{w}) = 0$.

Este problema se resolvió con los datos antes mencionados y el estado por alcanzar hasta $T = 1$ es:

$$u_T = 5(y - 1)^2 \quad \text{en } \Omega.$$

El cálculo del control óptimo (v) se realizó con el programa GRAFN.F descrito en la sección anterior, en este caso los resultados numéricos son satisfactorios lográndose la convergencia del método de gradiente conjugado. La siguiente tabla muestra dichos resultados

	h	Δt	M	T	ϵ	Tol	$E1$	$E2$
a	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.001	0.15	0.0114
b	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.0001	0.022	0.0037

Donde Tol es la tolerancia en el método de gradiente conjugado, $E1$ es el error en el control aproximado, es decir,

$$E1 = \|v^* - v\|_\infty$$

y $E2$ es el error en el estado aproximado, es decir,

$$E2 = \|u(T) - u_T\|_\infty.$$

Ambos ejemplos se obtuvieron en una PC-486 y el tiempo de máquina aproximado para el ejemplo a) es de 35 min, que corresponden a 26 iteraciones y en el ejemplo b) es de 60 min que corresponde a 44 iteraciones.

La gráfica 4.1.1 al final del capítulo muestra al control aproximado y los cortes del estado aproximado con este control, los datos corresponden a la discretización en a).

La gráfica 4.1.2, correspondiente también a la discretización en a), muestra el estado calculado con el control óptimo para $T = 1$.

A diferencia de los ejemplos estudiados en los capítulos anteriores si hay convergencia en el proceso iterativo de gradiente conjugado, la diferencia radica entonces en que el cálculo del gradiente

$$g^{(k)} = v^k + \beta P$$

ya no requiere de la derivada normal de P y por lo tanto no se tiene inestabilidad numérica que causaba su aproximación usando fórmulas de diferencias finitas.

Ejemplo 2.

Consideramos el problema directo (4.1) con los datos siguientes:

$$f(x, y, t) = \{ [e^{x(1-x)} - 1] e^{x(1-x)} (4x^2 - 4x - 1) + e^{2x(1-x)} (1 - 2x)^2 \} \\ - \frac{20}{(1+t)^2} [e^{x(1-x)} - 1],$$

$$v(x, t) = 0,$$

$$u_0(x, y) = 20[e^{x(1-x)} - 1]^2.$$

Usando el programa ECALFN.F para resolver el problema anterior, se genera la siguiente tabla de resultados:

	h	Δt	M	T	$ERROR$
a	$\frac{1}{10}$	$\frac{1}{10}$	10	1.0	1.3413×10^{-1}
b	$\frac{1}{16}$	$\frac{1}{20}$	20	1.0	2.0772×10^{-2}
c	$\frac{1}{32}$	$\frac{1}{100}$	100	1.0	9.4676×10^{-4}
d	$\frac{1}{32}$	$\frac{1}{500}$	500	1.0	7.9815×10^{-4}
e	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	7.8916×10^{-4}
f	$\frac{1}{64}$	$\frac{1}{100}$	100	1.0	5.8152×10^{-5}
g	$\frac{1}{64}$	$\frac{1}{500}$	500	1.0	8.0642×10^{-6}

Analizando la columna correspondiente a los errores de aproximación a la solución exacta $\left(u(x, y, t) = \frac{20}{1+t} [e^{x(1-x)} - 1]^2 \right)$ para $T = 1$, observamos que son satisfactorios y forman una buena base para resolver el problema inverso asociado.

El tiempo de máquina utilizado en una PC-486 es de aproximadamente 3 min para una discretización como en el inciso e) de la tabla anterior.

Ahora consideramos el problema inverso asociado con los datos de f y u_0 dados arriba. El estado por alcanzar en $T = 1$ es:

$$u_T = 10 [e^{x(1-x)} - 1]^2 \quad \text{en } \Omega.$$

El control óptimo (v) se calculó con el programa GRAFN.F y se generó la siguiente tabla de resultados:

	h	Δt	M	T	ϵ	Tol	$E1$	$E2$
a	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.001	0.022	0.0019
b	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.0001	0.022	0.00085

Como en el ejemplo anterior tenemos convergencia del método iterativo de gradiente conjugado, por lo cual el control óptimo calculado genera a su vez una buena aproximación al estado u_T (error $E2$).

El tiempo de máquina utilizado en una PC-486 es de 25 min en la discretización a) que corresponde a 20 iteraciones del método de gradiente conjugado y en el caso de la discretización b) se utilizan 50 min que corresponden a 34 iteraciones del método.

La gráfica 4.2.1 corresponde a la discretización en a) y muestra el control óptimo calculado y cortes del estado alcanzado con el control. La gráfica 4.2.2 muestra el estado alcanzado como una superficie sobre Ω .

Ejemplo 3.

Considerar el problema directo (4.1) con los siguientes datos:

$$\begin{aligned} f(x, y, t) &= [x^2(x-1)^2 + (y-1)^2] - 4(1+t)[3x^2 - 3x + 1], \\ v(x, t) &= 2(1+t), \\ u_0(x, y) &= x^2(x-1)^2 + (y-1)^2. \end{aligned}$$

Usando el programa ECALFN.F para resolver el problema anterior, se genera la siguiente tabla de resultados:

	h	Δt	M	T	$ERROR$
a	$\frac{1}{10}$	$\frac{1}{10}$	10	1.0	3.8770×10^{-3}
b	$\frac{1}{16}$	$\frac{1}{20}$	20	1.0	8.0852×10^{-4}
c	$\frac{1}{32}$	$\frac{1}{100}$	100	1.0	8.7202×10^{-5}
d	$\frac{1}{32}$	$\frac{1}{500}$	500	1.0	1.3425×10^{-4}
e	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	1.3985×10^{-4}
f	$\frac{1}{64}$	$\frac{1}{100}$	100	1.0	1.1972×10^{-5}
g	$\frac{1}{64}$	$\frac{1}{500}$	500	1.0	1.7994×10^{-5}

Analizando la columna correspondiente a los errores de aproximación a la solución exacta ($u(x, y, t) = (1+t)[x^2(x-1)^2 + (y-1)^2]$) para $T = 1$, observamos que son satisfactorios y forman una buena base para resolver el problema inverso asociado. El tiempo de máquina utilizado en una PC-486 es de aproximadamente 3 min para una discretización como en el inciso e) de la tabla anterior.

Consideramos ahora el problema inverso asociado con los datos de f y u_0 dados arriba. El estado por alcanzar en $T = 1$ es

$$u_T = 2 [x^2(x-1)^2 + (y-1)^2] \quad \text{en } \Omega.$$

El control óptimo (v) se calculó con el programa GRAFN.F y se generó la siguiente tabla de resultados:

	h	Δt	M	T	ϵ	Tol	$E1$	$E2$
a	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.001	0.114	0.0559
b	$\frac{1}{32}$	$\frac{1}{1000}$	1000	1.0	0.0001	0.0001	0.070	0.0043

Como en los ejemplos anteriores se tiene convergencia en el método iterativo de gradiente conjugado, por lo cual el control óptimo calculado genera a su vez una buena aproximación al estado u_T (error $E2$).

El tiempo de máquina utilizado en una PC-486 es de 28 min en la discretización a) que corresponde a 24 iteraciones del método de gradiente conjugado y en el caso de la discretización b) se utilizaron 70 min que corresponden a 40 iteraciones del método.

La gráfica 4.3.1 corresponde a la discretización en a) y muestra el control óptimo calculado y los cortes del estado alcanzado con este control. La gráfica 4.3.2 muestra el estado alcanzado como una superficie sobre Ω .

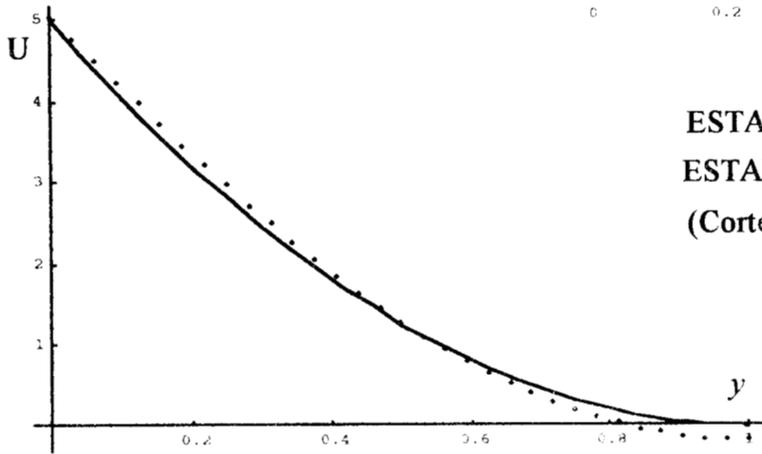
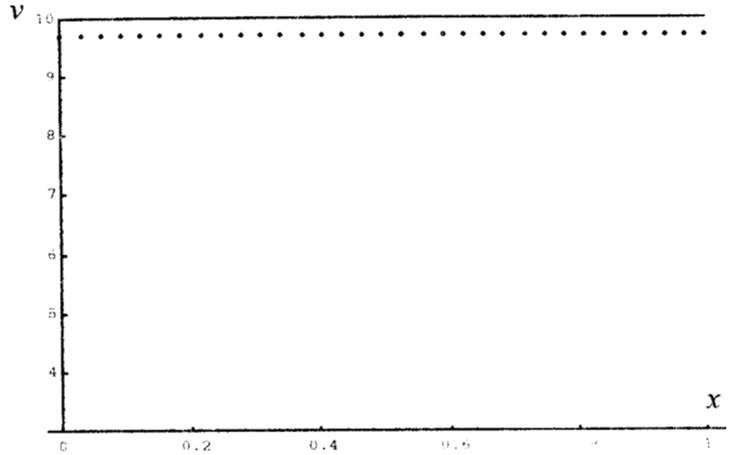
Remarcamos que estos ejemplos muestran la ventaja de que el gradiente

$$g^k = v^k + \beta P$$

no se calcule con la derivada normal de P .

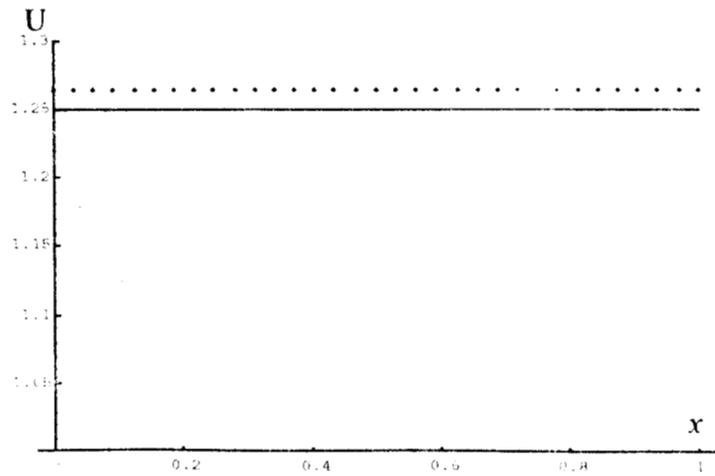
PROBLEMA INVERSO: ECUACION DE DIFUSION

CONTROL EXACTO vs
CONTROL APROXIMADO ($T=1$)



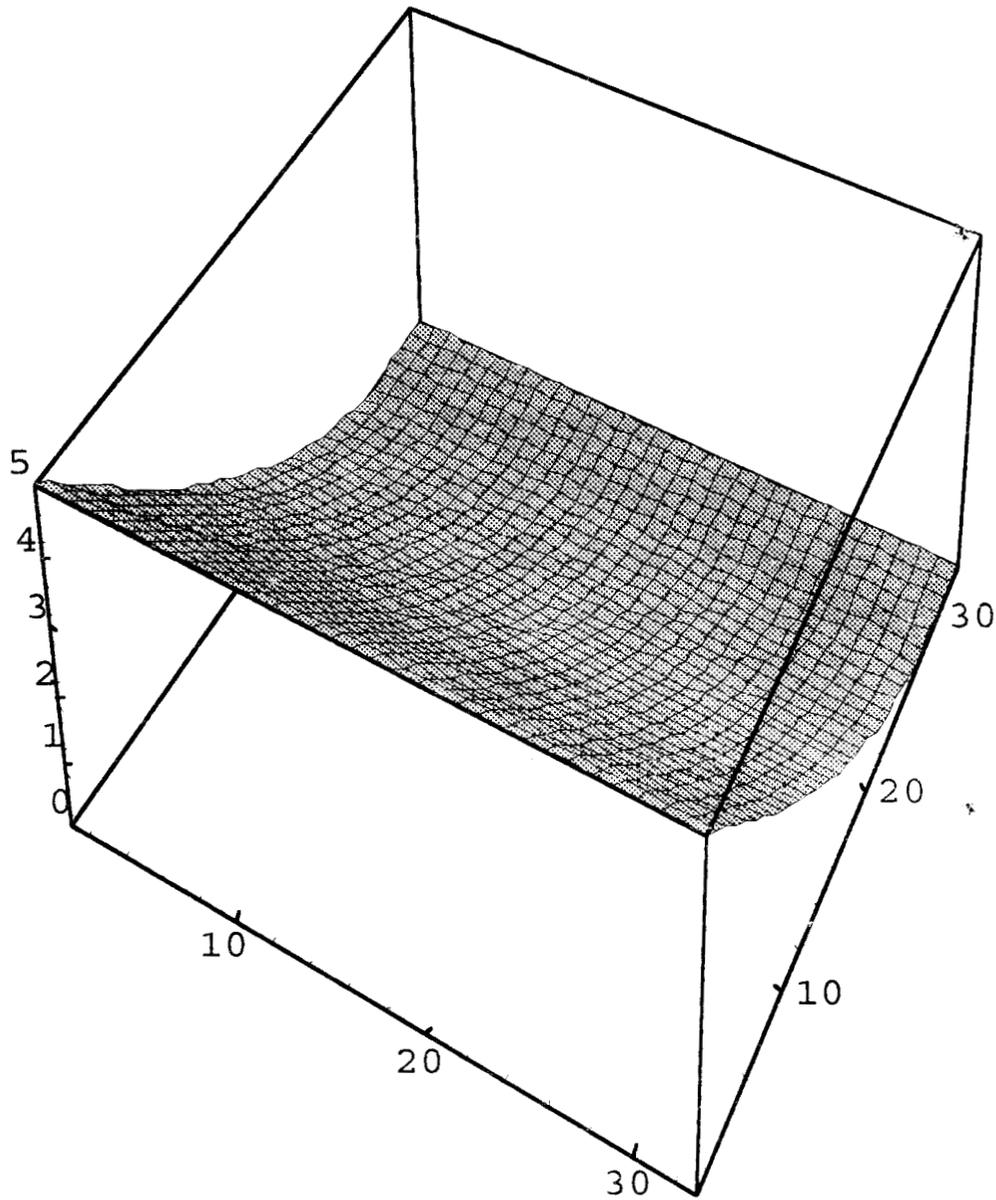
ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $x=1/2$ y $T=1$)

ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $y=1/2$ y $T=1$)



Grafica 4.1.1

PROBLEMA INVERSO: ECUACION DE DIFUSION

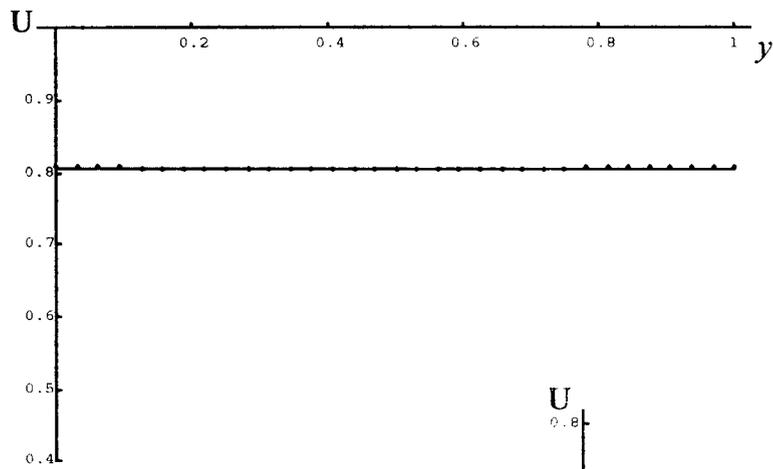
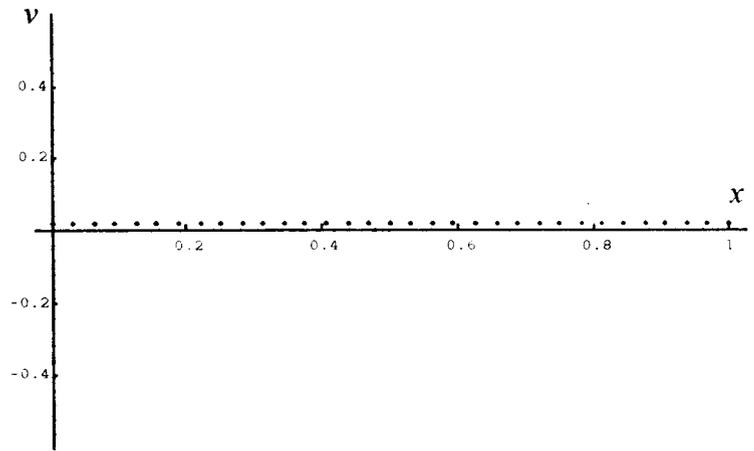


**ESTADO ALCANZADO EN $T=1$
CON CONTROL TIPO NEUMANN**

Grafica 4.1.2

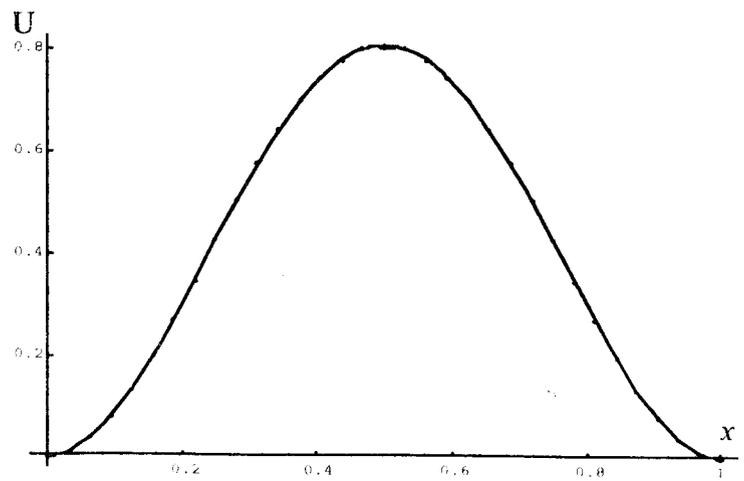
PROBLEMA INVERSO: ECUACION DE DIFUSION

CONTROL EXACTO ($v=0$) vs
CONTROL APROXIMADO
(Corte para $T=1$)



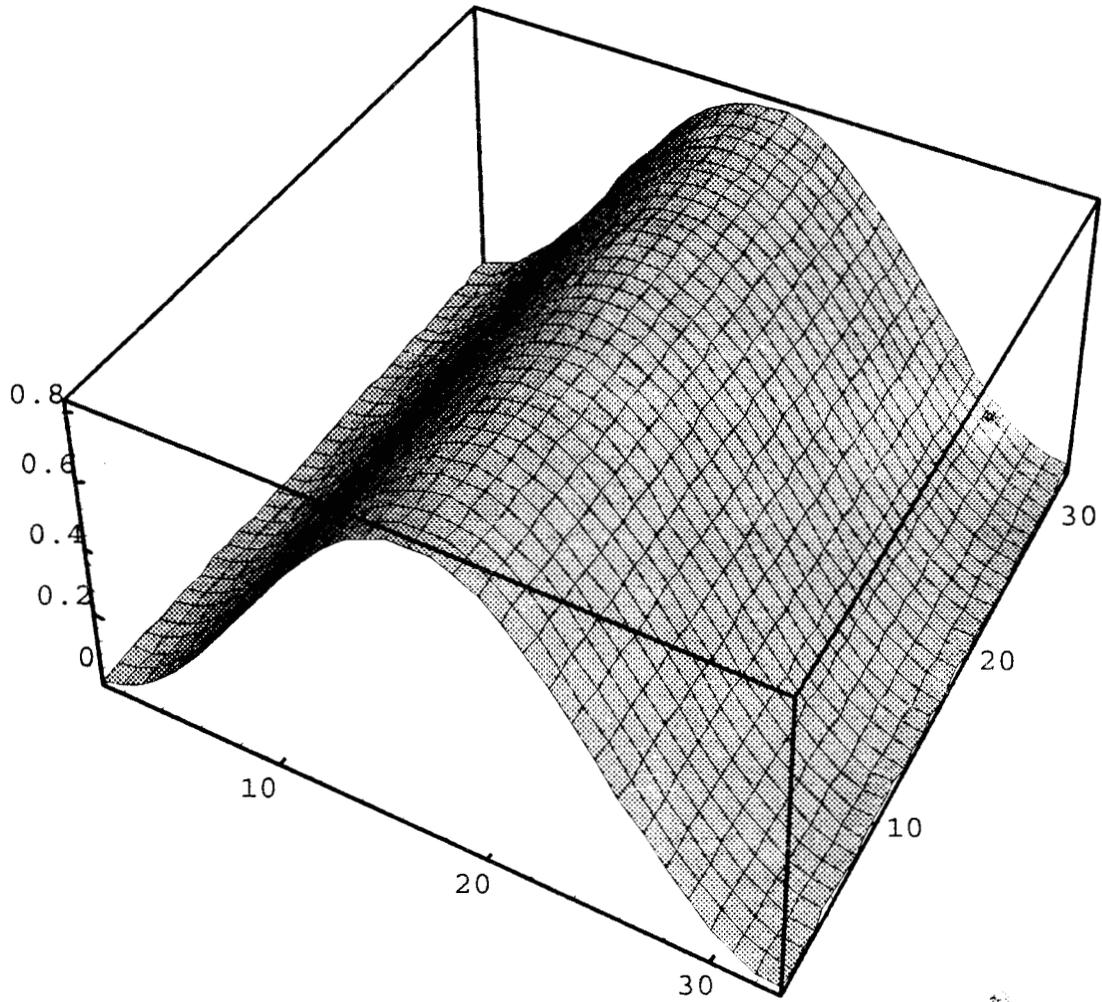
ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $x=1/2$ y $T=1$)

ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $y=1/2$ y $T=1$)



Grafica 4.2.1
96

PROBLEMA INVERSO: ECUACION DE DIFUSION

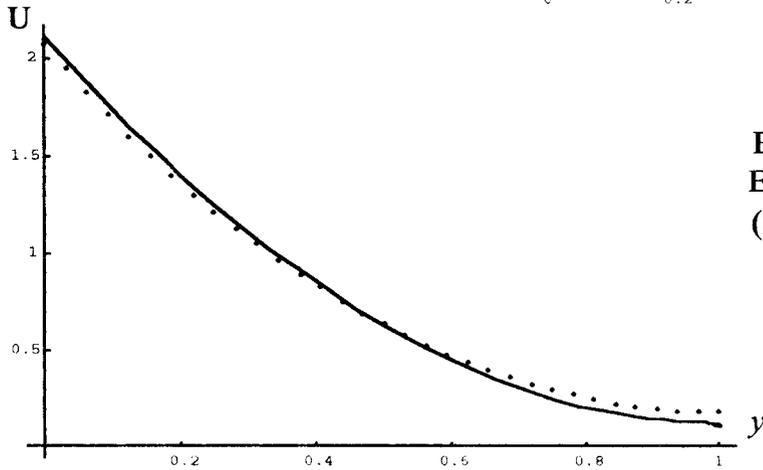
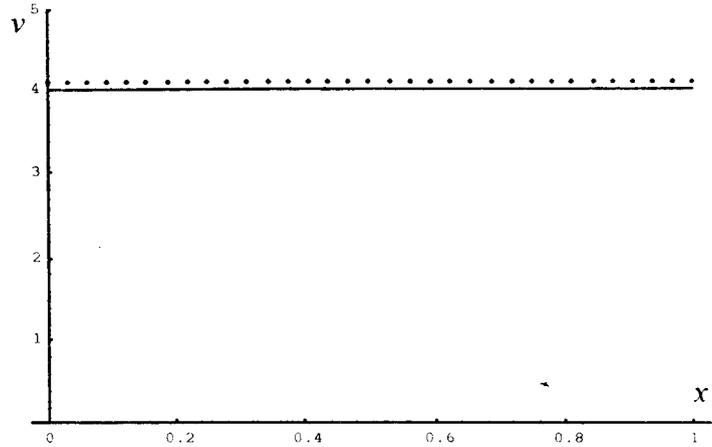


ESTADO ALCANZADO EN $T=1$
CON CONTROL TIPO NEUMANN

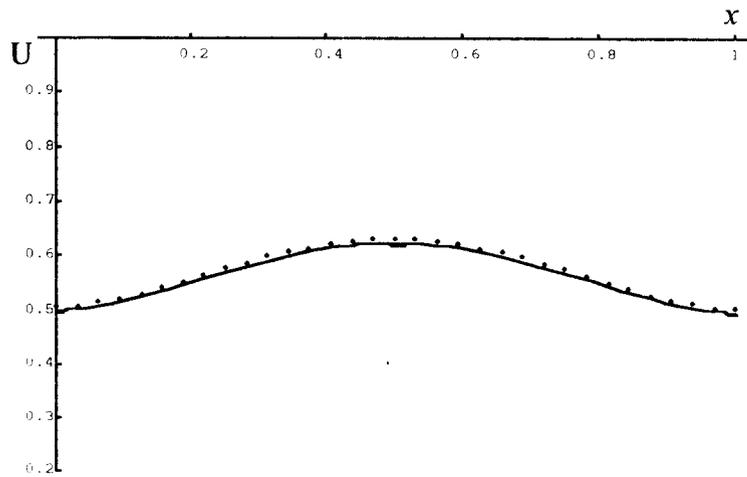
Grafica 4.2.2

PROBLEMA INVERSO: ECUACION DE DIFUSION

CONTROL EXACTO ($v=4$) vs
CONTROL APROXIMADO
(Corte para $T=1$)



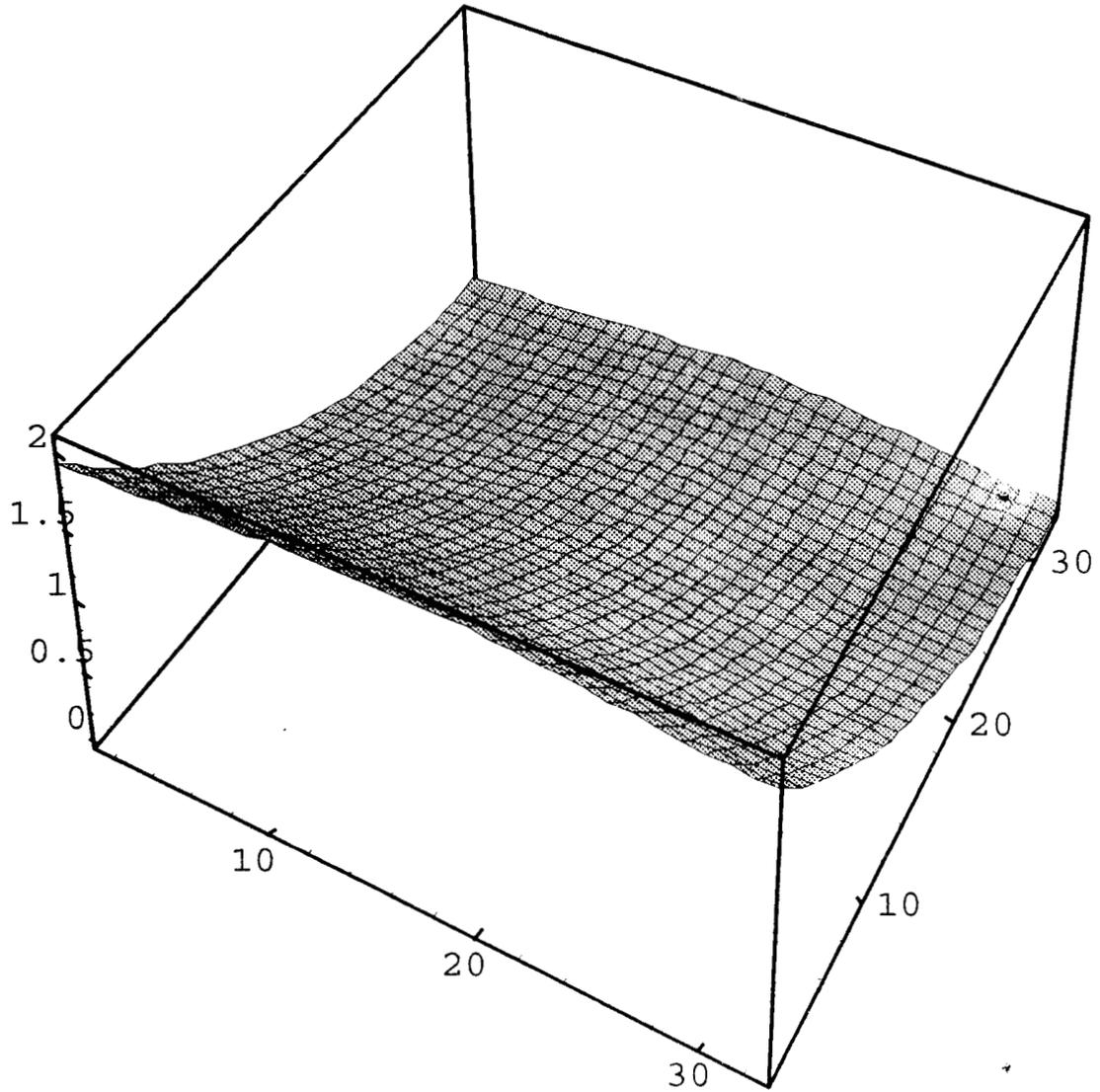
ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $x=1/2$ y $T=1$)



ESTADO EXACTO vs
ESTADO APROXIMADO
(Corte para $y=1/2$ y $T=1$)

Grafica 4.3.1
98

PROBLEMA INVERSO: ECUACION DE DIFUSION



ESTADO ALCANZADO EN $T=1$
CON CONTROL TIPO NEUMANN

Conclusiones

El trabajo presentado trata sobre una clase de problemas inversos relacionados con ecuaciones diferenciales parciales. Los problemas fueron del siguiente tipo: Hallar la acción (condición de frontera) de mínima energía que provoque que el sistema evolucione hasta una distribución dada u_T en un tiempo $T > 0$. La evolución del sistema se da de acuerdo a la evolución parcial que lo modele. Estos problemas de Control Optimo se formularon a través de una funcional (funcional de energía) la cual se busca minimizar. El cálculo del mínimo se realizó con el método de gradiente conjugado. Este método iterativo requiere del gradiente de la funcional y como se aclaró en los capítulos II, III y IV éste es fundamental para una adecuada solución. En los problemas donde se usó un control de tipo Dirichlet en la frontera, los resultados no son satisfactorios ya que el cálculo del gradiente requiere del cálculo de una derivada normal, la cual al aproximarla con fórmulas de diferencias finitas sufre pérdida de cifras significativas (inestabilidad), lo cual provoca un estancamiento en el método de gradiente conjugado. La primera alternativa tomada fué utilizar fórmulas de diferencias finitas que involucren un mayor número de puntos (mayor orden). Todas las fórmulas conocidas se obtienen usando polinomios de interpolación o polinomios de Taylor, por ejemplo, si buscamos aproximar la primera derivada con tres puntos hacia adelante, tenemos:

$$f'(a) = \alpha f(a) + \beta f(a+h) + \gamma f(a+2h) + \delta f(a+3h),$$

donde $h > 0$ es el tamaño de paso, desarrollando con la fórmula de Taylor hasta orden cuatro los términos $f(a+h)$, $f(a+2h)$ y $f(a+3h)$ tenemos:

$$\begin{aligned} f'(a) = & f(a)[\alpha + \beta + \gamma + \delta] + f'(a)[h\beta + 2h\gamma + 3h\delta] + f''(a) \left[\frac{h^2}{2}\beta + 2h^2\gamma + \frac{9}{2}h^2\delta \right] \\ & + f'''(a) \left[\frac{h^3}{3!}\beta + \frac{8h^3}{3!}\gamma + \frac{3^3h^3}{3!}\delta \right] + o(h^3) \end{aligned}$$

igualando términos obtenemos el sistema lineal:

$$\begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & h & 2h & 3h \\ 0 & \frac{h^2}{2} & \frac{2^2}{2}h^2 & \frac{3^2}{2}h^2 \\ 0 & \frac{h^3}{3!} & \frac{2^3}{3!}h^3 & \frac{3^3}{3!}h^3 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

con solución: $\alpha = -\frac{11}{6h}$, $\beta = \frac{3}{h}$, $\gamma = -\frac{3}{2h}$ y $\delta = \frac{1}{3h}$.

Esto nos proporciona una fórmula de orden tres, en forma similar se obtienen fórmulas de orden cuatro, cinco, etc.

Al utilizar este tipo de fórmulas de cálculo de $\frac{\partial P}{\partial n}$ la inestabilidad numérica fué similar, perdiéndose aproximadamente 50% de precisión en los datos. Por consiguiente el comportamiento en el proceso iterativo de gradiente conjugado de nuevo sufre un estancamiento. No encontrando otra alternativa para el cálculo de dicha derivada normal concluimos que el método de diferencias finitas no es el más adecuado para abordar estos problemas de control óptimo, de hecho, según lo expuesto en [6], la mejor forma de resolver el problema de control óptimo usando un control Dirichlet es a través del método de Elemento Finito, en este caso la derivada $\partial P/\partial n$ se aproxima en forma especial utilizando una forma integral.

En el problema donde se utilizó un control tipo Neumann el método de solución usando diferencias finitas, proporciona resultados satisfactorios; en este caso, el cálculo del gradiente no requiere de la derivada normal de P y por lo tanto no hay pérdida de cifras significativas, teniendo como consecuencia la convergencia del método de gradiente conjugado. Hasta la fecha no existe un trabajo semejante con el cual se pueda comparar, por lo cual concluimos que esta es una forma correcta de resolver el problema con control Neumann sin necesidad de grandes requerimientos de memoria y tiempo de cómputo, de hecho lo que nos indican estas experiencias es que una forma sencilla de controlar procesos de difusión y difusión-advección es a través de controlar el flujo de calor en la frontera y no la temperatura, este hecho como se apunta en [6] es más natural, dada la naturaleza de las ecuaciones.

Las ideas expuestas en este trabajo nos motivan a pensar la aplicación del método a otros sistemas con las siguientes características: la ecuación de estado debe ser lineal, la región de solución debe ser rectangular y el gradiente de la Funcional de Energía asociada no debe contener el cálculo de alguna derivada normal. De hecho, estas ideas también pueden fomar la

base para abordar problemas no lineales y acoplados (como las ecuaciones de Navier-Stokes o el modelo Boussinesq), donde la correspondiente funcional a minimizar no necesariamente es convexa; en [20] se explica mas sobre este último punto.

Apéndice A

Teoremas Citados.

1. Teorema de la Proyección.

Sea H espacio de Hilbert, $C \subset H$ cerrado y convexo y v un elemento de H tal que $v \notin C$. Entonces existe $Pv \in C$ tal que

$$(v - Pv, u - Pv) \leq 0 \quad \forall u \in C.$$

2. Teorema de Banach-Alaoglu.

Sea $(E, \|\cdot\|)$ un espacio normado. La bola unitaria cerrada $\mathbf{B}_1(0)$ es compacta en la topología débil.

Se puede consultar la demostración en [10].

3. Segunda Forma del Teorema de Green.

Sea S una superficie cerrada que rodea un volumen V y u, v funciones escalares de clase $C^2(\bar{V})$. Entonces se tiene,

$$\int \int_S \hat{n} \cdot (u \nabla v - v \nabla u) ds = \int \int \int_V (u \Delta v - v \Delta u) dV$$

donde \hat{n} representa el vector normal unitario exterior a S .

Se puede consultar la demostración en [19].

Apéndice B

Este apéndice incluye los listados de los programas utilizados para la simulación numérica de los problemas ejemplo de los capítulos II, III y IV.

- **Capítulo II**

- GRACON.F
- RECUCAL.F
- PIL20T.F
- PIL2.F
- GAUSEI.F
- FACTLU.F
- RESUELVE.F
- ERROR.F

- **Capítulo III**

- GRATRA.F
- RECUTRA.F
- SEIDEL.F
- LUFACF.F
- SOLVE.F

- **Capítulo IV**

- GRAFN.F
- ECALFN.F
- FISHPAK.F
- PROINT.F
- ERROR.F

```

C*****
C  EL SIGUIENTE PROGRAMA RESUELVE EL PROBLEMA DE CONTROL
C  OPTIMO ASOCIADO A LA ECUACION DE CALOR APLICANDO EL
C  METODO DE GRADIENTE CONJUGADO.
C
C  EL SIGNIFICADO DE LAS VARIABLES UTILIZADAS ES EL SIGUIENTE:
C  N=NUMERO DE SUBINTERVALOS PARA LA DISCRETIZACION EN ESPACIO
C  H=TAMAÑO DE PASO EN ESPACIO
C  M=NUMERO DE SUBINTERVALOS PARA LA DISCRETIZACION EN TIEMPO
C  DT=TAMAÑO DE PASO EN TIEMPO
C  TOL=TOLERANCIA EN GAUSS-SEIDEL
C  KMAX=NUMERO MAXIMO DE ITERACIONES EN GAUSS-SEIDEL
C  TOLGC=TOLERANCIA EN GRADIENTE CONJUGADO
C  MAXIT=NUMERO MAXIMO DE ITERACIONES EN GRADIENTE CONJUGADO
C  EPS=PARAMETRO DE PENALIZACION.
C  V=VECTOR DE CONTROL
C  U0=VECTOR DE CODICION INICIAL
C  U1=VECTOR DE ESTADO ALCANZADO
C  DNG0=VECTOR DE DERIVADA NORMAL

```

```

C*****

```

```

C
C  DOUBLE PRECISION U0(2500),U1(2500),V(2500),b(2500),X(50),
1      W1(50),W2(50),W3(50),W4(50),CDV0(4000),
1      DNG0(4000),DDW0(4000),AUX1(4000),AUX2(2500)
C
C  DOUBLE PRECISION H,DT,TOL,EPS,PIG0,PIG1,PI1,PI2,PI3,PI4,
1      FCI3,F3,FUSOL3,CERO,RO,TOLGC,GAM,G0,CRI
C
C  INTEGER N,M,KMAX,CONT,MAXIT
C
C  EXTERNAL F3,CERO
C
C  WRITE(*,*) 'INGRESA:N,DT,M,KMAX,TOL,EPS,TOLGC,MAXIT'
C  READ(*,*) N,DT,M,KMAX,TOL,EPS,TOLGC,MAXIT
C  H=1.0D0/DBLE(N+1)
C
C  DO 10 I=1,N*M
10  CDV0(I)=0.0D0
C
C  DO 30 I=1,N
C  DO 20 J=1,N
20  U0((I-1)*N+J)=FCI3(I*H,J*H)
30  CONTINUE
C
C  CALL RECUCAL(U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1      F3,CDV0,DNG0)
C
C  DO 40 I=1,N*M
40  AUX1(I)=0.0D0
C
C  DO 60 I=1,N
C  DO 50 J=1,N
C  U0((I-1)*N+J)=(U1((I-1)*N+J)-FUSOL3(I*H,J*H,DT*M))/EPS
50  AUX2((I-1)*N+J)=U1((I-1)*N+J)-FUSOL3(I*H,J*H,DT*M)
60  CONTINUE
C
C  CALL RECUCAL (U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1      CERO,AUX1,DNG0)
C
C  DO 90 K=1,M-1
C  DO 80 I=1,N
80  AUX1((K-1)*N+I)=DNG0(((M-K)-1)*N+I)

```

```

DO 100 I=1,N
100 AUX1((M-1)*N+I)=(0.5D0*AUX2((I-1)*N+2)-
1      2.0D0*AUX2((I-1)*N+1))/(EPS*H)
C
DO 110 I=1,N*M
110 DNG0(I)=CDV0(I)-AUX1(I)
C
DO 120 I=1,N*M
120 DDW0(I)=DNG0(I)
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG0)
G0=PIG0
C
C
CONT=0
121 CONT=CONT+1
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG0)
C
DO 130 I=1,N*N
130 U0(I)=0.0D0
C
CALL RECUCAL(U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1      CERO,DDW0,DNG0)
C
CALL PIL20T(H,N,DT,M,CDV0,DDW0,PI1)
CALL PIL20T(H,N,DT,M,DDW0,DDW0,PI2)
CALL PIL2(H,N,AUX2,U1,PI3)
CALL PIL2(H,N,U1,U1,PI4)
RO=(PI1+PI3/EPS)/(PI2+PI4/EPS)
C
DO 140 I=1,N*M
140 CDV0(I)=CDV0(I)-RO*DDW0(I)
C
DO 160 I=1,N
DO 150 J=1,N
150 U0((I-1)*N+J)=FCI3(I*H,J*H)
160 CONTINUE
C
CALL RECUCAL(U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1      F3,CDV0,DNG0)
C
DO 170 I=1,N*M
170 AUX1(I)=0.0D0
C
DO 190 I=1,N
DO 180 J=1,N
U0((I-1)*N+J)=(U1((I-1)*N+J)-FUSOL3(I*H,J*H,DT*M))/EPS
180 AUX2((I-1)*N+J)=U1((I-1)*N+J)-FUSOL3(I*H,J*H,DT*M)
190 CONTINUE
C
CALL RECUCAL(U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1      CERO,AUX1,DNG0)
C
DO 210 K=1,M-1
DO 200 I=1,N
200 AUX1((K-1)*N+I)=DNG0(((M-K)-1)*N+I)
210 CONTINUE
DO 220 I=1,N
220 AUX1((M-1)*N+I)=(0.5D0*AUX2((I-1)*N+2)
1      -2.0D0*AUX2((I-1)*N+1))/(H*EPS)
C

```

```

230 DNG0(I)=CDV0(I)-AUX1(I)
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG1)
C
GAM=PIG1/PIG0
CRI=DSQRT(PIG1/G0)
WRITE(*,*) 'CRITERIO DE PARO=',CRI
C
IF (CRI.LE.TOLGC) THEN
WRITE(*,*) 'CONVERGENCIA EN GRADIENTE CONJUGADO'
WRITE(*,*) 'CONTROL OPTIMO'
DO 240 I=1,N*M
240 WRITE(*,*) CDV0(I)
C
DO 242 I=1,N
DO 241 J=1,N
241 U0((I-1)*N+J)=FCI3(I*H,J*H)
242 CONTINUE
CALL RECUCAL (U0,U1,V,b,X,W1,W2,W3,W4,N,DT,M,KMAX,TOL,
1 F3,CDV0,DNG0)
DO 243 I=1,N*N
243 WRITE(*,*) U1(I)
C
STOP
ELSE
DO 250 I=1,N*M
250 DDW0(I)=DNG0(I)+GAM*DDW0(I)
END IF
C
IF (CONT.EQ.MAXIT) THEN
WRITE(*,*) 'NO HUBO CONVERGENCIA EN GRADIENTE CONJUGADO'
STOP
ELSE
WRITE(*,*) 'ITERACION EN GRADIENTE CONJUGADO=',CONT
GO TO 121
END IF
END
C*****
C
C LA SIGUIENTE SUBROUTINA RESUELVE LOS PROBLEMAS DIRECTOS
C ASOCIADOS AL PROBLEMA DE CONTROL OPTIMO.
C*****
SUBROUTINE RECUCAL(U0,U1,V,b,X,W1,W2,W3,W4,
1 N,DT,M,KMAX,TOL,FLD,CD,DNG0)
C
DOUBLE PRECISION H,ALFA,BETA,GAMA,DT,TOL,E,U0(1),
1 U1(1),V(1),b(1),W1(1),W2(1),W3(1),W4(1),FLD,
1 X(1),R,S,T,P,CD(1),DNG0(1)
INTEGER N,M,KMAX
C
H=1.0D0/DBLE(N+1)
ALFA=3.0D0/(2.0D0*DT)+4.0D0/H**2
BETA=-1.0D0/H**2
GAMA=1.0D0/DT+4.0D0/H**2
C
DO 10 I=1,N**2
10 V(I)=U0(I)
C
DO 30 I=1,N

```

0 b((I-1)*N+J)=FLD(I*H,J*H,DT)+U0((I-1)*N+J)/DT

0 CONTINUE

DO 35 I=1,N

15 b((I-1)*N+1)=b((I-1)*N+1)-BETA*CD(I)

R=GAMA+(4.0D0/3.0D0)*BETA

S=BETA

T=GAMA+(8.0D0/3.0D0)*BETA

P=(2.0D0/3.0D0)*BETA

CALL FACTLU(R,S,T,P,W1,W2,N)

R=GAMA

S=BETA

T=GAMA+(4.0D0/3.0D0)*BETA

P=(2.0D0/3.0D0)*BETA

CALL FACTLU(R,S,T,P,W3,W4,N)

CALL GAUSEI(U0,U1,b,TOL,KMAX,N,E,BETA,X,W1,W2,W3,W4)

DO 36 I=1,N

36 DNG0(I)=(1.5D0*CD(I)+0.5D0*U1((I-1)*N+2)-2.0D0*

1 U1((I-1)*N+1))/H

DO 60 I=1,N**2

U0(I)=V(I)

50 V(I)=0.0D0

R=ALFA+(4.0D0/3.0D0)*BETA

S=BETA

T=ALFA+(8.0D0/3.0D0)*BETA

P=(2.0D0/3.0D0)*BETA

CALL FACTLU(R,S,T,P,W1,W2,N)

R=ALFA

S=BETA

T=ALFA+(4.0D0/3.0D0)*BETA

P=(2.0D0/3.0D0)*BETA

CALL FACTLU(R,S,T,P,W3,W4,N)

DO 150 K=2,M

DO 80 I=1,N

DO 70 J=1,N

70 b((I-1)*N+J)=FLD(I*H,J*H,K*DT)+

1 (4.0D0*U1((I-1)*N+J)-U0((I-1)*N+J))/(2.0D0*DT)

80 CONTINUE

DO 85 I=1,N

85 b((I-1)*N+1)=b((I-1)*N+1)-BETA*CD((K-1)*N+1)

DO 90 I=1,N**2

V(I)=U1(I)

U0(I)=U1(I)

90 U1(I)=0.0D0

CALL GAUSEI(U0,U1,b,TOL,KMAX,N,E,BETA,X,W1,W2,W3,W4)

DO 95 I=1,N

95 DNG0((K-1)*N+I)=(1.5D0*CD((K-1)*N+I)+0.5D0*U1((I-1)*N+2)-2.0D0*

1 U1((I-1)*N+1))/H

DO 100 I=1,N**2

U0(I)=V(I)

```

C
150 CONTINUE
C
RETURN
END
C*****
C
C LA SIGUIENTE SUBROUTINA CALCULA LOS PRODUCTOS INTERIORES
C EN [0,1] X [0,T].
C
C*****
SUBROUTINE PIL20T (H,N,DT,M,XI,ETA,PI)
DOUBLE PRECISION H,DT,XI(1),ETA(1),PI,SUM
INTEGER N,M
PI=0.0D0
C
SUM=0.0D0
DO 10 I=1,N
10 SUM=SUM+2.0D0*XI(I)*ETA(I)
SUM=SUM+XI(1)*ETA(1)
SUM=(SUM+XI(N)*ETA(N))*H/2.0D0
C
PI=PI+SUM
C
DO 30 K=1,M-1
SUM=0.0D0
DO 20 I=1,N
20 SUM=SUM+2.0D0*XI((K-1)*N+I)*ETA((K-1)*N+I)
SUM=SUM+XI((K-1)*N+1)*ETA((K-1)*N+1)
SUM=(SUM+XI(K*N)*ETA(K*N))*H/2.0D0
PI=PI+2.0D0*SUM
30 CONTINUE
C
SUM=0.0D0
DO 40 I=1,N
40 SUM=SUM+2.0D0*XI((M-1)*N+I)*ETA((M-1)*N+I)
SUM=SUM+XI((M-1)*N+1)*ETA((M-1)*N+1)
SUM=(SUM+XI(M*N)*ETA(M*N))*H/2.0D0
C
PI=(PI+SUM)*DT/2.0D0
RETURN
END
C*****
C
C LA SIGUIENTE SUBROUTINA CALCULA LOS PRODUCTOS INTERIORES
C EN [0,1] X [0,1].
C
C*****
SUBROUTINE PIL2 (H,N,XI,ETA,PI)
DOUBLE PRECISION H,XI(1),ETA(1),PI,SUM
INTEGER N
PI=0.0D0
C
SUM=0.0D0
SUM=SUM*XI(1)*ETA(1)
DO 10 J=1,N
10 SUM=SUM+2.0D0*XI(J)*ETA(J)
SUM=(SUM+XI(N)*ETA(N))*H/2.0D0
PI=PI+SUM
C
DO 30 I=1,N
SUM=0.0D0

```

```

DO 20 J=1,N
SUM=SUM+2.0D0*X1((I-1)*N+J)*ETA((I-1)*N+J)
SUM=(SUM+X1(I*N)*ETA(I*N))*H/2.0D0
PI=PI+2.0D0*SUM
CONTINUE

```

```

PI=(PI+SUM)*H/2.0D0
RETURN
END

```

```

*****
SUBROUTINE GAUSEI(X0,X1,b,TOL,KMAX,N,E,BETA,X,V1,V2,V3,V4)
DOUBLE PRECISION X0(1),X1(1),b(1),TOL,E,BETA,
i      X(1),V1(1),V2(1),V3(1),V4(1)
INTEGER N,KMAX

```

SOLUCION DEL SISTEMA POR BLOQUES CON GAUSS-SEIDEL Y EN CADA BLOQUE SE RESUELVE DIRECTAMENTE CON UNA FACTORIZACION LU.

```

*****
K=0
0 K=K+1
DO 20 I=1,N
0 X(I)=b(I)-(2.0D0*BETA*X0(N+I))/3.0D0
CALL RESUELVE (N,V1,V2,X,BETA,2.0D0*BETA/3.0D0)
DO 30 I=1,N
0 X1(I)=X(I)
DO 70 J=2,N-1
DO 50 I=1,N
i0 X(I)=b((J-1)*N+I)-BETA*(X1((J-2)*N+I)+X0(J*N+I))
CALL RESUELVE (N,V3,V4,X,BETA,2.0D0*BETA/3.0D0)
DO 60 I=1,N
j0 X1((J-1)*N+I)=X(I)
70 CONTINUE
DO 80 I=1,N
30 X(I)=b((N-1)*N+I)-(2.0D0*BETA*X1((N-2)*N+I))/3.0D0
CALL RESUELVE (N,V1,V2,X,BETA,2.0D0*BETA/3.0D0)
DO 90 I=1,N
30 X1((N-1)*N+I)=X(I)
CALL ERROR (X0,X1,E,N)
IF (E.GT.TOL.AND.K.LT.KMAX) THEN
DO 100 I=1,N**2
X0(I)=X1(I)
X1(I)=0.0D0
100 CONTINUE
GO TO 10
END IF
IF (E.LE.TOL) THEN
WRITE(*,*) 'CONVERGENCIA EN EL PROCESO DE GAUSS-SEIDEL'
WRITE(*,*) 'ITERACIONES= ',K
ELSE
WRITE(*,*) 'NO HUBO CONVERGENCIA EN GAUSS-SEIDEL'
STOP
END IF
RETURN
END

```

```

C*****
C
C LA SIGUIENTE SUBROUTINA CALCULA LA FACTORIZACION LU DE
C CADA BLOQUE DE LA DIAGONAL.
C
C*****

```

```

DOUBLE PRECISION R,S,T,P,L(1),U(1)
INTEGER N
----INSTRUCCIONES PARA FACTORIZAR EN LU CADA BLOQUE-----
DO 5 I=1,N
L(I)=0.0D0
; U(I)=0.0D0
L(1)=R
U(1)=S/L(1)
DO 10 I=2,N-1
L(I)=R-S*U(I-1)
U(I)=S/L(I)
10 CONTINUE
L(N)=T-P*U(N-1)
U(N)=0.0D0
RETURN
END
)*****
)
) LA SIGUIENTE SUBROUTINA RESUELVE EL SISTEMA PARA CADA BLOQUE
) UTILIZANDO LA FACTORIZACION LU.
)
)*****
SUBROUTINE RESUELVE (N,L,U,X,S,P)
DOUBLE PRECISION L(1),U(1),X(1),S,P
INTEGER N
X(1)=X(1)/L(1)
DO 10 I=2,N-1
X(I)=(X(I)-S*X(I-1))/L(I)
10 CONTINUE
X(N)=(X(N)-P*X(N-1))/L(N)
DO 20 I=N-1,1,-1
X(I)=X(I)-U(I)*X(I+1)
20 CONTINUE
RETURN
END
C*****
C
SUBROUTINE ERROR (X0,X1,E,N)
DOUBLE PRECISION X0(1),X1(1),E
INTEGER N
C----CALCULO DE LA NORMA INFINITO DE LOS VECTORES X0 y X1----
E=0.0D0
E=DMAX1(DABS(X0(1)-X1(1)),DABS(X0(2)-X1(2)))
DO 10 I=3,N**2
E=DMAX1(E,DABS(X0(I)-X1(I)))
10 CONTINUE
RETURN
END
C*****

```

PROGRAM GRATRA

```

C*****
C
C ESTE PROGRAMA RESUELVE EL PROBLEMA DE CONTROL OPTIMO
C ASOCIADO A LA ECUACION DE TRANSPORTE USANDO EL METODO
C DE GRADIENTE CONJUGADO.
C
C*****
DOUBLE PRECISION U0(2500),U1(2500),V(2500),bLD(2500),X(50),
1      W1(50),W2(50),W3(50),W4(50),W5(50),W6(50),
1      CDV0(4000),DNG0(4000),DDW0(4000),AUX(2500)
C
DOUBLE PRECISION H,DT,TOL,DELTA,Z1,Z2,EPS,RO,FTRA,FCI3,CERO,
1      PI1,PI2,PI3,PI4,PIG0,PIG1,G0,TOLGC,GAMA,
1      FUSOL3
C
INTEGER N,M,KMAX,CONT,MAXIT
C
EXTERNAL FTRA,CERO
C
WRITE(*,*) 'INGRESA:N,DT,M,KMAX,TOL,DELTA,Z1,Z2,EPS,TOLGC,MAXIT'
READ(*,*) N,DT,M,KMAX,TOL,DELTA,Z1,Z2,EPS,TOLGC,MAXIT
H=1.0D0/DBLE(N+1)
C
DO 10 I=1,N*M
10 CDV0(I)=0.0D0
C
DO 30 I=1,N
DO 20 J=1,N
20 U0((I-1)*N+J)=FCI3(I*H,J*H)
30 CONTINUE
C
CALL RECUTRA(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,
1      TOL,DELTA,Z1,Z2,FTRA,CDV0)
C
DO 50 I=1,N
DO 40 J=1,N
AUX((I-1)*N+J)=U1((I-1)*N+J)-FUSOL3(I*H,J*H,M*DT)
40 U0((I-1)*N+J)=AUX((I-1)*N+J)/EPS
50 CONTINUE
C
CALL RECUAD(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,
1      TOL,DELTA,Z1,Z2,DNG0)
C
DO 60 I=1,N*M
DNG0(I)=CDV0(I)-DELTA*DNG0(I)
60 DDW0(I)=DNG0(I)
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG0)
G0=DSQRT(PIG0)
C
C
CONT=0
65 CONT=CONT+1
WRITE(*,*) 'ITERACION=',CONT,'DE GRADIENTE CONJUGADO'
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG0)
C
DO 70 I=1,N*N
70 U0(I)=0.0D0
C
CALL RECUTRA(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,

```

```

C
CALL PIL20T(H,N,DT,M,CDV0,DDW0,PI1)
CALL PIL20T(H,N,DT,M,DDW0,DDW0,PI2)
CALL PIL2(H,N,AUX,U1,PI3)
CALL PIL2(H,N,U1,U1,PI4)
RO=(PI1 + PI3/EPS)/(PI2 + PI4/EPS)
C
DO 80 I=1,N*M
80 CDV0(I)=CDV0(I)-RO*DDW0(I)
C
DO 100 I=1,N
DO 90 J=1,N
90 U0((I-1)*N+J)=FCI3(I*H,J*H)
100 CONTINUE
C
CALL RECUTRA(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,
1 TOL,DELTA,Z1,Z2,FTRA,CDV0)
C
DO 120 I=1,N
DO 110 J=1,N
AUX((I-1)*N+J)=U1((I-1)*N+J)-FUSOL3(I*H,J*H,M*DT)
110 U0((I-1)*N+J)=AUX((I-1)*N+J)/EPS
120 CONTINUE
C
CALL RECUAD(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,
1 TOL,DELTA,Z1,Z2,DNG0)
C
DO 130 I=1,N*M
130 DNG0(I)=CDV0(I)-DELTA*DNG0(I)
C
CALL PIL20T(H,N,DT,M,DNG0,DNG0,PIG1)
C
WRITE(*,*) 'CRITERIO DE PARO=',DSQRT(PIG1)/G0
C
IF (DSQRT(PIG1)/G0.LE.TOLGC) THEN
WRITE(*,*) 'CONTROL OPTIMO TIPO DIRICHLET'
DO 140 I=1,N*M
140 WRITE(*,*) CDV0(I)
WRITE(*,*) 'SOLUCION DE LA ECUACION DE TRANSPORTE CON EL COPT'
READ(*,*)
DO 142 I=1,N
DO 141 J=1,N
141 U0((I-1)*N+J)=FCI3(I*H,J*H)
142 CONTINUE
CALL RECUTRA(U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,N,DT,M,KMAX,
1 TOL,DELTA,Z1,Z2,FTRA,CDV0)
DO 145 I=1,N
DO 144 J=1,N
144 WRITE(*,*) U1((I-1)*N+J),' ',FUSOL3(I*H,J*H,M*DT)
145 CONTINUE
STOP
END IF
C
GAMA=PIG1/PIG0
DO 150 I=1,N*M
150 DDW0(I)=DNG0(I)+GAMA*DDW0(I)
C
IF (CONT.LT.MAXIT) THEN
GO TO 65
ELSE
WRITE(*,*) 'NO HUBO CONVERGENCIA EN GRADIENTE CONJUGADO'
STOP

```

END

C*****

C

C LA SIGUIENTE SUBROUTINA RESUELVE LOS PROBLEMAS DIRECTOS
C ASOCIADOS AL METODO DE GRADIENTE CONJUGADO.

C

C*****

SUBROUTINE RECUTRA (U0,U1,V,bLD,X,W1,W2,W3,W4,W5,W6,
1 N,DT,M,KMAX,TOL,DELTA,Z1,Z2,FLD,CD)
DOUBLE PRECISION H,ALFA,BETA,A,B,C,D,DT,TOL,E,U0(1),U1(1),
1 V(1),bLD(1),W1(1),W2(1),W3(1),W4(1),W5(1),
1 W6(1),X(1),Z1,Z2,FLD,CD(1),R,S,T,P,Q,DELTA

C

C---SOLUCION DE LA ECUACION DE TRANSPORTE POR DIFERENCIAS FINITAS---
C---DE ORDEN DOS. DISCRETIZACION EN EL TIEMPO CON GEAR-MODIFICADO --
C---Y EN ESPACIO CON DIFERENCIAS CENTRADAS ---

C

H=1.0D0/DBLE(N+1)
ALFA=3.0D0/(2.0D0*DT) + 4.0D0*DELTA/H**2
BETA=1.0D0/DT + 4.0D0*DELTA/H**2
A=-(DELTA/H**2 + Z1/(2.0D0*H))
B=Z1/(2.0D0*H)-DELTA/H**2
C=-(DELTA/H**2 + Z2/(2.0D0*H))
D=Z2/(2.0D0*H) - DELTA/H**2

C

R=BETA+4.0D0*A/3.0D0
S=BETA+4.0D0*(A+D)/3.0D0
T=D
P=C
Q=C-D/3.0D0
CALL LUFACT(R,S,T,P,Q,W1,W2,N)

C

R=BETA
S=BETA+4.0D0*D/3.0D0
CALL LUFACT(R,S,T,P,Q,W3,W4,N)

C

R=BETA + 4.0D0*B/3.0D0
S=BETA + 4.0D0*(B+D)/3.0D0
CALL LUFACT(R,S,T,P,Q,W5,W6,N)

C

DO 10 I=1,N**2

10 V(I)=U0(I)

C

DO 30 I=1,N
DO 20 J=1,N

20 bLD((I-1)*N+J)=FLD(I*H,J*H,DT,DELTA) + U0((I-1)*N+J)/DT

30 CONTINUE

DO 40 I=1,N

40 bLD((I-1)*N+1)=bLD((I-1)*N+1)-C*CD(I)

C

CALL SEIDEL(U0,U1,bLD,TOL,KMAX,N,E,A,B,T,W1,W2,W3,W4,
1 W5,W6,X)

C

DO 50 I=1,N**2
U0(I)=V(I)

50 V(I)=0.0D0

C

R=ALFA+4.0D0*A/3.0D0
S=ALFA+4.0D0*(A+D)/3.0D0
CALL LUFACT(R,S,T,P,Q,W1,W2,N)

C

R=ALFA

CALL LUFACT(R,S,T,P,Q,W3,W4,N)

R=ALFA + 4.0D0*B/3.0D0

S=ALFA + 4.0D0*(B+D)/3.0D0

CALL LUFACT(R,S,T,P,Q,W5,W6,N)

DO 110 K=2,M

DO 70 I=1,N

DO 60 J=1,N

0 bLD((I-1)*N+J)=FLD(I*H,J*H,K*DT,DELTA)+(4.0D0*U1((I-1)*N+J)-
1 U0((I-1)*N+J))/(2.0D0*DT)

0 CONTINUE

DO 80 I=1,N

0 bLD((I-1)*N+1)=bLD((I-1)*N+1)-C*CD((K-1)*N+I)

DO 90 I=1,N**2

V(I)=U1(I)

U0(I)=U1(I)

00 U1(I)=0.0D0

CALL SEIDEL(U0,U1,bLD,TOL,KMAX,N,E,A,B,T,W1,W2,W3,W4,

1 W5,W6,X)

DO 100 I=1,N**2

U0(I)=V(I)

00 V(I)=0.0D0

10 CONTINUE

RETURN

END

:

: LA SIGUIENTE SUBROUTINA RESUELVE LOS SISTEMAS LINEALES EN
: CADA NIVEL DE TIEMPO UTILIZANDO EL METODO DE GAUS-SEIDEL.

:

SUBROUTINE SEIDEL (X0,X1,bLD,TOL,KMAX,N,E,A,B,T,W1,W2,W3,W4,

1 W5,W6,X)

DOUBLE PRECISION X0(1),X1(1),bLD(1),TOL,E,A,B,T,W1(1),W2(1),

1 W3(1),W4(1),W5(1),W6(1),X(1)

INTEGER KMAX,N

)-----SOLUCION DEL SISTEMA EN BLOQUES USANDO GAUSS-SEIDEL-----

:

K=0

10 K=K+1

DO 20 I=1,N

20 X(I)=bLD(I)-(B-A/3.0D0)*X0(N+I)

CALL SOLVE (W1,W2,T,X,N)

DO 30 I=1,N

30 X1(I)=X(I)

DO 60 J=2,N-1

DO 40 I=1,N

40 X(I)=bLD((J-1)*N+I)-B*X0(J*N+I)-A*X1((J-2)*N+I)

CALL SOLVE (W3,W4,T,X,N)

DO 50 I=1,N

50 X1((J-1)*N+I)=X(I)

60 CONTINUE

DO 70 I=1,N

70 X(I)=bLD((N-1)*N+I)-(A-B/3.0D0)*X1((N-2)*N+I)

CALL SOLVE (W5,W6,T,X,N)

DO 80 I=1,N

80 X1((N-1)*N+I)=X(I)

CALL ERROR(X0,X1,E,N)

IF (E.GT.TOL.AND.K.LT.KMAX) THEN

DO 90 I=1,N*N

```

X1(I)=0.0D0
90 CONTINUE
GO TO 10
END IF
IF (E.LE.TOL) THEN
WRITE(*,*) 'CONVERGENCIA EN EL PROCESO DE GAUSS-SEIDEL'
WRITE(*,*) 'ITERACIONES=',K
ELSE
WRITE(*,*) 'NO HUBO CONVERGENCIA EN GAUSS-SEIDEL'
STOP
END IF
RETURN
END
C*****
C
C LA SIGUIENTE SUBROUTINA CALCULA LA FACTORIZACION LU EN CADA
C BLOQUE DE LA DIAGONAL.
C
C*****
SUBROUTINE LUFAC (R,S,T,P,Q,L,U,N)
DOUBLE PRECISION R,S,T,P,Q,L(1),U(1)
INTEGER N
DO 10 I=1,N
L(I)=0.0D0
U(I)=0.0D0
10 CONTINUE
U(1)=R
DO 20 I=1,N-2
L(I)=P/U(I)
U(I+1)=R-L(I)*T
20 CONTINUE
L(N-1)=Q/U(N-1)
U(N)=S-L(N-1)*T
L(N)=0.0D0
RETURN
END
C*****
C
C LA SIGUIENTE SUBROUTINA UTILIZA LA FACTORIZACION LU DE CADA
C BLOQUE PARA RESOLVER EL SISTEMA POR BLOQUES.
C
C*****+*****
SUBROUTINE SOLVE (L,U,T,X,N)
DOUBLE PRECISION T,L(1),U(1),X(1)
INTEGER N
C----SUSTITUCION HACIA ADELANTE-----
C
DO 10 I=2,N
10 X(I)=X(I)-L(I-1)*X(I-1)
C
C----SUSTITUCION HACIA ATRAS-----
C
X(N)=X(N)/U(N)
DO 20 N-1,1,-1
20 X(I)=(X(I)-T*X(I+1))/U(I)
C
RETURN
END

```

EL SIGUIENTE PROGRAMA RESUELVE EL PROBLEMA DE CONTROL OPTIMO
ASOCIADO A LA ECUACION DE CALOR USANDO EL METODO DE GRADIENTE
CONJUGADO. EL CONTROL EN ESTE CASO ES DE TIPO NEUMANN.

IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION CD(101,1001),U0(101,101),U1(101,101),P(101,1001),
1 G(101,1001),W(101,1001),AUX1(101,101)
EXTERNAL FLD1,CERO

C INGRESO DE DATOS Y CALCULO DE CONSTANTES

WRITE(*,*) 'INGRESAR N,MT,T,EPS,TOLGC'
READ(*,*) N,MT,T,EPS,TOLGC
H=1.0D0/N
DT=T/MT
IPDIM=101

C PASO #1 Y #2 DEL ALGORITMO DE GRADIENTE CONJUGADO

DO 20 K=1,MT+1
DO 10 I=1,N+1
P(I,K)=0.0D0
10 CD(I,K)=0.0D0
20 CONTINUE

DO 40 I=1,N+1
DO 30 J=1,N+1
U1(I,J)=0.0D0
30 U0(I,J)=FCI1((I-1)*H,(J-1)*H)
40 CONTINUE

C CALL ECALFN(FLD1,CD,U0,U1,IPDIM,N,MT,T,P)

DO 60 I=1,N+1
DO 50 J=1,N+1
AUX1(I,J)=U1(I,J)-FSOL1((I-1)*H,(J-1)*H,T)
U0(I,J)=AUX1(I,J)/EPS
50 U1(I,J)=0.0D0
60 CONTINUE

DO 80 K=1,MT+1
DO 70 I=1,N+1
70 P(I,K)=0.0D0
80 CONTINUE

C CALL ECALFN(CERO,CD,U0,U1,IPDIM,N,MT,T,P)

DO 99 K=1,MT+1
DO 90 I=1,N+1
G(I,K)=CD(I,K)+P(I,K)
90 W(I,K)=G(I,K)
99 CONTINUE

C CALCULO DE LA NORMA DE G0

CALL PROINT(G,G,H,N,DT,MT,PINT,IPDIM)
GON=DSQRT(PINT)

C PASO #3 Y #4 DEL ALGORITMO DE GRADIENTE CONJUGADO

```

ICONT=0
100 ICONT=ICONT+1
C
DO 120 I=1,N+1
DO 110 J=1,N+1
U0(I,J)=0.0D0
110 U1(I,J)=0.0D0
120 CONTINUE
C
CALL ECALFN(CERO,W,U0,U1,IPDIM,N,MT,T,P)
C
CALL PROINT(CD,W,H,N,DT,MT,PINT,IPDIM)
P1=PINT
CALL PROINT(AUX1,U1,H,N,H,N,PINT,IPDIM)
P2=PINT/EPS
CALL PROINT(W,W,H,N,DT,MT,PINT,IPDIM)
P3=PINT
CALL PROINT(U1,U1,H,N,H,N,PINT,IPDIM)
P4=PINT/EPS
RO=(P1+P2)/(P3+P4)
C
DO 140 K=1,MT+1
DO 130 I=1,N+1
130 CD(I,K)=CD(I,K)-RO*W(I,K)
140 CONTINUE
C-----
C PASO #5 DEL ALGORITMO DE GRADIENTE CONJUGADO
C
DO 160 I=1,N+1
DO 150 J=1,N+1
U0(I,J)=FCI1((I-1)*H,(J-1)*H)
150 U1(I,J)=0.0D0
160 CONTINUE
C
CALL ECALFN(FLD1,CD,U0,U1,IPDIM,N,MT,T,P)
C
DO 180 I=1,N+1
DO 170 J=1,N+1
AUX1(I,J)=U1(I,J)-FSOL1((I-1)*H,(J-1)*H,T)
U0(I,J)=AUX1(I,J)/EPS
170 U1(I,J)=0.0D0
180 CONTINUE
C
C CALCULO DEL PRODUCTO INTERIOR (GN,GN)
C
CALL PROINT(G,G,H,N,DT,MT,PINT,IPDIM)
GNN=PINT
C
DO 200 K=1,MT+1
DO 190 I=1,N+1
P(I,K)=0.0D0
190 G(I,K)=0.0D0
200 CONTINUE
C
CALL ECALFN(CERO,G,U0,U1,IPDIM,N,MT,T,P)
C
DO 220 K=1,MT+1
DO 210 I=1,N+1
210 G(I,K)=CD(I,K)+P(I,K)
220 CONTINUE
C-----
C PASO #6 Y #7 DE GRADIENTE CONJUGADO

```

```
ICONT=0
0 ICONT=ICONT+1
```

```
DO 120 I=1,N+1
DO 110 J=1,N+1
U0(I,J)=0.0D0
0 U1(I,J)=0.0D0
0 CONTINUE
```

```
CALL ECALFN(CERO,W,U0,U1,IPDIM,N,MT,T,P)
```

```
CALL PROINT(CD,W,H,N,DT,MT,PINT,IPDIM)
P1=PINT
CALL PROINT(AUX1,U1,H,N,H,N,PINT,IPDIM)
P2=PINT/EPS
CALL PROINT(W,W,H,N,DT,MT,PINT,IPDIM)
P3=PINT
CALL PROINT(U1,U1,H,N,H,N,PINT,IPDIM)
P4=PINT/EPS
RO=(P1+P2)/(P3+P4)
```

```
DO 140 K=1,MT+1
DO 130 I=1,N+1
30 CD(I,K)=CD(I,K)-RO*W(I,K)
40 CONTINUE
```

PASO #5 DEL ALGORITMO DE GRADIENTE CONJUGADO

```
DO 160 I=1,N+1
DO 150 J=1,N+1
U0(I,J)=FCI1((I-1)*H,(J-1)*H)
50 U1(I,J)=0.0D0
60 CONTINUE
```

```
CALL ECALFN(FLD1,CD,U0,U1,IPDIM,N,MT,T,P)
```

```
DO 180 I=1,N+1
DO 170 J=1,N+1
AUX1(I,J)=U1(I,J)-FSOL1((I-1)*H,(J-1)*H,T)
U0(I,J)=AUX1(I,J)/EPS
70 U1(I,J)=0.0D0
80 CONTINUE
```

CALCULO DEL PRODUCTO INTERIOR (GN,GN)

```
CALL PROINT(G,G,H,N,DT,MT,PINT,IPDIM)
GNN=PINT
```

```
DO 200 K=1,MT+1
DO 190 I=1,N+1
P(I,K)=0.0D0
190 G(I,K)=0.0D0
200 CONTINUE
```

```
CALL ECALFN(CERO,G,U0,U1,IPDIM,N,MT,T,P)
```

```
DO 220 K=1,MT+1
DO 210 I=1,N+1
210 G(I,K)=CD(I,K)+P(I,K)
220 CONTINUE
```

PASO #6 Y #7 DE GRADIENTE CONJUGADO

```

CALL PROINT(G,G,H,N,DT,MT,PINT,IPDIM)
CRIT=DSQRT(PINT)/GON
WRITE(*,*) 'GRADIENTE=',CRIT
WRITE(*,*) 'ITERACION DE GRADIENTE=',ICONT

IF (CRIT.LE.TOLGC) THEN
GO TO 250
ELSE
GAMA=PINT/GNN
DO 240 K=1,MT+1
DO 230 I=1,N+1
30 W(I,K)=G(I,K)+GAMA*W(I,K)
40 CONTINUE
GO TO 100
END IF
50 CONTINUE

```

RESULTADOS

```

DO 280 K=1,MT+1
DO 270 I=1,N+1
70 WRITE(*,*) CD(I,K), ' ',FCD1((I-1)*H,(K-1)*DT)
80 CONTINUE
READ(*,*)

DO 290 I=1,N+1
DO 285 J=1,N+1
U1(I,J)=0.0D0
85 U0(I,J)=FCH1((I-1)*H,(J-1)*H)
90 CONTINUE

CALL ECALFN(FLD1,CD,U0,U1,IPDIM,N,MT,T,P)

DO 310 I=1,N+1
DO 300 J=1,N+1
300 U0(I,J)=FSOL1((I-1)*H,(J-1)*H,T)
310 CONTINUE

CALL ERROR(U0,U1,E,N,IPDIM)

WRITE(*,*) 'ERROR DE APROXIMACION AL ESTADO DESEADO ES=',E
STOP
END

```

LA SIGUIENTE SUBROUTINA CALCULA LOS PRODUCTOS INTERIORES

```

SUBROUTINE PROINT(U,V,H1,N1,H2,N2,PINT,IPDIM)
IMPLICIT DOUBLE PRECISION (A-H,O-Z)
DIMENSION U(IPDIM,1),V(IPDIM,1)

PINT=0.0D0
DO 20 I=1,N1+1
AUX=0.0D0
DO 10 J=2,N2
10 AUX=AUX+2.0D0*U(I,J)*V(I,J)
AUX=H2*(U(I,1)*V(I,1)+U(I,N2+1)*V(I,N2+1)+AUX)/2.0D0
IF (I.EQ.1.OR.I.EQ.N1+1) THEN
PINT=PINT+AUX
ELSE
PINT=PINT+2.0D0*AUX

```

```
CONTINUE
PINT=H1*PINT/2.0D0
RETURN
END
```

LA SIGUIENTE SUBROUTINA CALCULA LA NORMA INFINITO DE DOS
VECTORES.

```
SUBROUTINE ERROR(X0,X1,E,N,IPDIM)
IMPLICIT DOUBLE PRECISION(A-H,O-Z)
DIMENSION X0(IPDIM,1),X1(IPDIM,1)
```

```
E=0.0D0
```

```
DO 20 I=1,N+1
```

```
DO 10 J=1,N+1
```

```
0 E=DMAX1(E,DABS(X0(I,J)-X1(I,J)))
```

```
0 CONTINUE
```

```
RETURN
```

```
END
```

LA SIGUIENTE SUBROUTINA RESUELVE LOS PROBLEMAS DIRECTOS
UTILIZANDO LA SUBROUTINA FISHPAK EN LOS PROBLEMAS ELIPTICOS.

SUBROUTINE ECALFN(FLD,CD,U0,U1,IPDIM,N,MT,T)

IMPLICIT DOUBLE PRECISION(A-H,O-Z)

DIMENSION CD(IPDIM,1),U0(IPDIM,1),U1(IPDIM,1)

DIMENSION BDA(101),BDB(101),BDC(101),BDD(101),GRHS(101,101),

1 USOL(101,101),W(4000)

EXTERNAL FLD

DT=T/MT

H=1.0D0/N

IORDER=4

A=0.0D0

B=1.0D0

M=N

MBDCDN=3

DO 10 J=1,N+1

BDA(J)=0.0D0

BDB(J)=0.0D0

BDD(J)=0.0D0

0 CONTINUE

ALPHA=0.0D0

BETA=0.0D0

C=0.0D0

D=1.0D0

NBDCND=3

DO 20 I=1,N+1

10 BDC(I)=-CD(I,2)

DO 40 I=1,N+1

DO 30 J=1,N+1

30 GRHS(I,J)=-FLD((I-1)*H,(J-1)*H,DT)-U0(I,J)/DT

40 CONTINUE

IDMN=101

IBANDERA=0

CALL DEPX4(IORDER,A,B,M,MBDCND,BDA,ALPHA,BDB,BETA,C,D,N,NBDCND,BDC,

1 BDD,COFX,GRHS,USOL,IDMN,W,PERTRB,IERROR)

DO 60 I=1,N+1

DO 50 J=1,N+1

U1(I,J)=USOL(I,J)

50 USOL(I,J)=0.0D0

60 CONTINUE

WRITE(*,*) 'SE CALCULO LA SOLUCION AL TIEMPO=1'

7
7
7
7

DO 200 K=3,MT+1

IORDER=4

A=0.0D0

B=1.0D0

M=N

MBDCND=3

DO 70 J=1,N+1

BDA(J)=0.0D0

BDB(J)=0.0D0

BDD(J)=0.0D0

70 CONTINUE

ALPHA=0.0D0

```

C=0.0D0
D=1.0D0
NBDCND=3
DO 80 I=1,N+1
0 BDC(I)=-CD(I,K)
DO 100 I=1,N+1
DO 90 J=1,N+1
0 GRHS(I,J)=-FLD((I-1)*H,(J-1)*H,(K-1)*DT)-2.0D0*U1(I,J)/DT+
1 U0(I,J)/(2.0D0*DT)
00 CONTINUE
IDMN=101
IBANDERA=1
CALL DEPX4(IORDER,A,B,M,MBDCND,BDA,ALPHA,BDB,BETA,C,D,N,NBDCND,BDC,
1 BDD,COFX,GRHS,USOL,IDMN,W,PERTRB,IERROR)
DO 120 I=1,N+1
DO 110 J=1,N+1
U0(I,J)=U1(I,J)
U1(I,J)=USOL(I,J)
110 USOL(I,J)=0.0D0
120 CONTINUE
WRITE(*,*) 'SE CALCULO LA SOLUCION AL TIEMPO=',K-1
200 CONTINUE
RETURN
END

```

```

SUBROUTINE DEPX4(IORDER,A,B,M,MBDCND,BDA,ALPHA,BDB,BETA,C,D,N,
1NBDCND,BDC,BDD,COFX,GRHS,USOL,IDMN,W,PERTRB,IERROR)

```

```

*****

```

```

*
*           *
*       F I S H P A K           *
*           *
*           *
* A PACKAGE OF FORTRAN SUBPROGRAMS FOR THE SOLUTION OF *
*           *
* SEPARABLE ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS *
*           *
* (VERSION 3.1 , OCTOBER 1980) *
*           *
*           *
*       B Y           *
*           *
* JOHN ADAMS, PAUL SWARZTRAUBER AND ROLAND SWEET *
*           *
*       O F           *
*           *
* THE NATIONAL CENTER FOR ATMOSPHERIC RESEARCH *
*           *
* BOULDER, COLORADO (80307) U.S.A. *
*           *
*       W H I C H   I S   S P O N S O R E D   B Y *
*           *
* THE NATIONAL SCIENCE FOUNDATION *
*           *

```

```

*****

```

ARGUMENTS USOL(IDMN,N+1), GRHS(IDMN,N+1),
W (SEE ARGUMENT LIST)

TEST REVISION OCTOBER 1980

PURPOSE SEPX4 SOLVES FOR EITHER THE SECOND-ORDER
FINITE DIFFERENCE APPROXIMATION OR A
A FOURTH-ORDER APPROXIMATION TO THE
SOLUTION OF A SEPARABLE ELLIPTIC EQUATION
 $AF(X)*UXX + BF(X)*UX + CF(X)*U + UYY = G(X, Y)$

ON A RECTANGLE (X GREATER THAN OR EQUAL TO A
AND LESS THAN OR EQUAL TO B; Y GREATER THAN
OR EQUAL TO C AND LESS THAN OR EQUAL TO D).
ANY COMBINATION OF PERIODIC OR MIXED BOUNDARY
CONDITIONS IS ALLOWED.

IF BOUNDARY CONDITIONS IN THE X DIRECTION
ARE PERIODIC (SEE MBDCND=0 BELOW) THEN THE
COEFFICIENTS MUST SATISFY
 $AF(X)=C1, BF(X)=0, CF(X)=C2$ FOR ALL X.
HERE C1,C2 ARE CONSTANTS, C1.GT.0.

THE POSSIBLE BOUNDARY CONDITIONS ARE
IN THE X-DIRECTION:

- (0) PERIODIC, $U(X+B-A, Y)=U(X, Y)$ FOR ALL Y,X
- (1) $U(A, Y), U(B, Y)$ ARE SPECIFIED FOR ALL Y
- (2) $U(A, Y), DU(B, Y)/DX + BETA*U(B, Y)$ ARE
SPECIFIED FOR ALL Y
- (3) $DU(A, Y)/DX + ALPHA*U(A, Y), DU(B, Y)/DX +$
 $BETA*U(B, Y)$ ARE SPECIFIED FOR ALL Y
- (4) $DU(A, Y)/DX + ALPHA*U(A, Y), U(B, Y)$ ARE
SPECIFIED FOR ALL Y

IN THE Y-DIRECTION:

- (0) PERIODIC, $U(X, Y+D-C)=U(X, Y)$ FOR ALL X,Y
- (1) $U(X, C), U(X, D)$ ARE SPECIFIED FOR ALL X
- (2) $U(X, C), DU(X, D)/DY$ ARE SPECIFIED FOR ALL X
- (3) $DU(X, C)/DY, DU(X, D)/DY$ ARE SPECIFIED FOR
ALL X
- (4) $DU(X, C)/DY, U(X, D)$ ARE SPECIFIED FOR ALL X

USAGE CALL SEPX4(IORDER, A, B, M, MBDCND, BDA, ALPHA, BDB,
BETA, C, D, N, NBDCND, BDC, BDD, COFX,
GRHS, USOL, IDMN, W, PERTRB, IERROR)

ARGUMENTS

IORDER

= 2 IF A SECOND-ORDER APPROXIMATION IS SOUGHT
= 4 IF A FOURTH-ORDER APPROXIMATION IS SOUGHT

A, B

THE RANGE OF THE X-INDEPENDENT VARIABLE;
I.E., X IS GREATER THAN OR EQUAL TO A AND
LESS THAN OR EQUAL TO B. A MUST BE LESS THAN
B.

M

THE NUMBER OF PANELS INTO WHICH THE INTERVAL
[A, B] IS SUBDIVIDED. HENCE, THERE WILL BE

$XI = A + (I-1) * DLX$ FOR $I=1,2,\dots,M+1$ WHERE
 $DLX = (B-A)/M$ IS THE PANEL WIDTH. M MUST BE
LESS THAN $IDMN$ AND GREATER THAN 5.

MBDCND

INDICATES THE TYPE OF BOUNDARY CONDITION AT
 $X=A$ AND $X=B$

- = 0 IF THE SOLUTION IS PERIODIC IN X ; I.E.,
 $U(X+B-A, Y) = U(X, Y)$ FOR ALL Y, X
- = 1 IF THE SOLUTION IS SPECIFIED AT $X=A$ AND
 $X=B$; I.E., $U(A, Y)$ AND $U(B, Y)$ ARE
SPECIFIED FOR ALL Y
- = 2 IF THE SOLUTION IS SPECIFIED AT $X=A$ AND
THE BOUNDARY CONDITION IS MIXED AT $X=B$;
I.E., $U(A, Y)$ AND $DU(B, Y)/DX + BETA * U(B, Y)$
ARE SPECIFIED FOR ALL Y
- = 3 IF THE BOUNDARY CONDITIONS AT $X=A$ AND $X=B$
ARE MIXED; I.E., $DU(A, Y)/DX + ALPHA * U(A, Y)$
AND $DU(B, Y)/DX + BETA * U(B, Y)$ ARE SPECIFIED
FOR ALL Y
- = 4 IF THE BOUNDARY CONDITION AT $X=A$ IS MIXED
AND THE SOLUTION IS SPECIFIED AT $X=B$;
I.E., $DU(A, Y)/DX + ALPHA * U(A, Y)$ AND $U(B, Y)$
ARE SPECIFIED FOR ALL Y

BDA

A ONE-DIMENSIONAL ARRAY OF LENGTH $N+1$ THAT
SPECIFIES THE VALUES OF $DU(A, Y)/DX +$
 $ALPHA * U(A, Y)$ AT $X=A$, WHEN $MBDCND=3$ OR 4.

$$BDA(J) = DU(A, YJ)/DX + ALPHA * U(A, YJ);$$

$$J=1,2,\dots,N+1$$

WHEN $MBDCND$ HAS ANY OTHER VALUE, BDA IS A
DUMMY PARAMETER.

N INPUT

ALPHA

THE SCALAR MULTIPLYING THE SOLUTION IN CASE
OF A MIXED BOUNDARY CONDITION AT $X=A$ (SEE
ARGUMENT BDA). IF $MBDCND \neq 3,4$ THEN $ALPHA$ IS
A DUMMY PARAMETER.

BDB

A ONE-DIMENSIONAL ARRAY OF LENGTH $N+1$ THAT
SPECIFIES THE VALUES OF $DU(B, Y)/DX +$
 $BETA * U(B, Y)$ AT $X=B$. WHEN $MBDCND=2$ OR 3

$$BDB(J) = DU(B, YJ)/DX + BETA * U(B, YJ);$$

$$J=1,2,\dots,N+1$$

WHEN $MBDCND$ HAS ANY OTHER VALUE, BDB IS A
DUMMY PARAMETER.

BETA

THE SCALAR MULTIPLYING THE SOLUTION IN CASE
OF A MIXED BOUNDARY CONDITION AT $X=B$ (SEE
ARGUMENT BDB). IF $MBDCND \neq 2,3$ THEN $BETA$ IS A
DUMMY PARAMETER.

C,D

THE RANGE OF THE Y -INDEPENDENT VARIABLE;
I.E., Y IS GREATER THAN OR EQUAL TO C AND
LESS THAN OR EQUAL TO D . C MUST BE LESS THAN
 D .

N

[C,D] IS SUBDIVIDED. HENCE, THERE WILL BE N+1 GRID POINTS IN THE Y-DIRECTION GIVEN BY $YJ=C+(J-1)*DLY$ FOR $J=1,2,\dots,N+1$ WHERE $DLY=(D-C)/N$ IS THE PANEL WIDTH. IN ADDITION, N MUST BE GREATER THAN 4.

NBDCND

INDICATES THE TYPES OF BOUNDARY CONDITIONS AT $Y=C$ AND $Y=D$

- = 0 IF THE SOLUTION IS PERIODIC IN Y; I.E.,
 $U(X,Y+D-C)=U(X,Y)$ FOR ALL X,Y
- = 1 IF THE SOLUTION IS SPECIFIED AT $Y=C$ AND $Y=D$, I.E., $U(X,C)$ AND $U(X,D)$ ARE SPECIFIED FOR ALL X
- = 2 IF THE SOLUTION IS SPECIFIED AT $Y=C$ AND THE BOUNDARY CONDITION IS MIXED AT $Y=D$; I.E., $DU(X,C)/DY$ AND $U(X,D)$ ARE SPECIFIED FOR ALL X
- = 3 IF THE BOUNDARY CONDITIONS ARE MIXED AT $Y=C$ AND $Y=D$ I.E., $DU(X,C)/DY$ AND $DU(X,D)/DY$ ARE SPECIFIED FOR ALL X
- = 4 IF THE BOUNDARY CONDITION IS MIXED AT $Y=C$ AND THE SOLUTION IS SPECIFIED AT $Y=D$; I.E. $DU(X,C)/DY+GAMA*U(X,C)$ AND $U(X,D)$ ARE SPECIFIED FOR ALL X

BDC

A ONE-DIMENSIONAL ARRAY OF LENGTH M+1 THAT SPECIFIES THE VALUE $DU(X,C)/DY$ AT $Y=C$. WHEN NBDCND=3 OR 4

$$BDC(I) = DU(XI,C)/DY$$
$$I=1,2,\dots,M+1.$$

WHEN NBDCND HAS ANY OTHER VALUE, BDC IS A DUMMY PARAMETER.

BDD

A ONE-DIMENSIONAL ARRAY OF LENGTH M+1 THAT SPECIFIES THE VALUE OF $DU(X,D)/DY$ AT $Y=D$. WHEN NBDCND=2 OR 3

$$BDD(I)=DU(XI,D)/DY$$
$$I=1,2,\dots,M+1.$$

WHEN NBDCND HAS ANY OTHER VALUE, BDD IS A DUMMY PARAMETER.

COFX

A USER-SUPPLIED SUBPROGRAM WITH PARAMETERS X, AFUN, BFUN, CFUN WHICH RETURNS THE VALUES OF THE X-DEPENDENT COEFFICIENTS $AF(X)$, $BF(X)$, $CF(X)$ IN THE ELLIPTIC EQUATION AT X.

IF BOUNDARY CONDITIONS IN THE X DIRECTION ARE PERIODIC THEN THE COEFFICIENTS MUST SATISFY $AF(X)=C1, BF(X)=0, CF(X)=C2$ FOR ALL X. HERE $C1.GT.0$ AND C2 ARE CONSTANTS.

NOTE THAT COFX MUST BE DECLARED EXTERNAL IN THE CALLING ROUTINE.

GRHS

VALUES OF THE RIGHT-HAND SIDE OF THE ELLIPTIC EQUATION; I.E., $GRHS(I,J)=G(XI,YI)$, FOR $I=2,\dots,M$; $J=2,\dots,N$. AT THE BOUNDARIES, GRHS IS DEFINED BY

MBDCND GRHS(1,J) GRHS(M+1,J)

```

-----
0  G(A,YJ)  G(B,YJ)
1  *        *
2  *        G(B,YJ) J=1,2,...,N+1
3  G(A,YJ)  G(B,YJ)
4  G(A,YJ)  *
```

NBDCND GRHS(I,1) GRHS(I,N+1)

```

-----
0  G(XI,C)  G(XI,D)
1  *        *
2  *        G(XI,D) I=1,2,...,M+1
3  G(XI,C)  G(XI,D)
4  G(XI,C)  *
```

WHERE * MEANS THESE QUANTITIES ARE NOT USED. GRHS SHOULD BE DIMENSIONED IDMN BY AT LEAST N+1 IN THE CALLING ROUTINE.

USOL

A TWO-DIMENSIONAL ARRAY THAT SPECIFIES THE VALUES OF THE SOLUTION ALONG THE BOUNDARIES. AT THE BOUNDARIES, USOL IS DEFINED BY

MBDCND USOL(1,J) USOL(M+1,J)

```

-----
0  *        *
1  U(A,YJ)  U(B,YJ)
2  U(A,YJ)  *   J=1,2,...,N+1
3  *        *
4  *        U(B,YJ)
```

NBDCND USOL(I,1) USOL(I,N+1)

```

-----
0  *        *
1  U(XI,C)  U(XI,D)
2  U(XI,C)  *   I=1,2,...,M+1
3  *        *
4  *        U(XI,D)
```

WHERE * MEANS THE QUANTITIES ARE NOT USED IN THE SOLUTION.

IF IORDER=2, THE USER MAY EQUIVALENCE GRHS AND USOL TO SAVE SPACE. NOTE THAT IN THIS CASE THE TABLES SPECIFYING THE BOUNDARIES OF THE GRHS AND USOL ARRAYS DETERMINE THE BOUNDARIES UNIQUELY EXCEPT AT THE CORNERS. IF THE TABLES CALL FOR BOTH $G(X,Y)$ AND $U(X,Y)$ AT A CORNER THEN THE SOLUTION MUST BE

CHOSEN. FOR EXAMPLE, IF MBDCND=2 AND NBDCND=4, THEN $U(A,C)$, $U(A,D)$, $U(B,D)$ MUST BE CHOSEN AT THE CORNERS IN ADDITION TO $G(B,C)$.

IF IORDER=4, THEN THE TWO ARRAYS, USOL AND

USOL SHOULD BE DIMENSIONED IDMN BY AT LEAST N+1 IN THE CALLING ROUTINE.

IDMN

THE ROW (OR FIRST) DIMENSION OF THE ARRAYS GRHS AND USOL AS IT APPEARS IN THE PROGRAM CALLING SEPX4. THIS PARAMETER IS USED TO SPECIFY THE VARIABLE DIMENSION OF GRHS AND USOL. IDMN MUST BE AT LEAST 7 AND GREATER THAN OR EQUAL TO M+1.

W

A ONE-DIMENSIONAL ARRAY THAT MUST BE PROVIDED BY THE USER FOR WORK SPACE.

$10*N+(16+INT(\log_2(N)))*(M+1)+23$ WILL SUFFICE AS A LENGTH FOR W. THE ACTUAL LENGTH OF W IN THE CALLING ROUTINE MUST BE SET IN W(1) (SEE IERROR=11).

OUTPUT

USOL

CONTAINS THE APPROXIMATE SOLUTION TO THE ELLIPTIC EQUATION. USOL(I,J) IS THE APPROXIMATION TO $U(XI,YJ)$ FOR $I=1,2,\dots,M+1$ AND $J=1,2,\dots,N+1$. THE APPROXIMATION HAS ERROR $O(DLX^{**2}+DLY^{**2})$ IF CALLED WITH IORDER=2 AND $O(DLX^{**4}+DLY^{**4})$ IF CALLED WITH IORDER=4.

W

CONTAINS INTERMEDIATE VALUES THAT MUST NOT BE DESTROYED IF SEPX4 IS CALLED AGAIN WITH INTL=1. IN ADDITION W(1) CONTAINS THE EXACT MINIMAL LENGTH (IN FLOATING POINT) REQUIRED FOR THE WORK SPACE (SEE IERROR=11).

PERTRB

IF A COMBINATION OF PERIODIC OR DERIVATIVE BOUNDARY CONDITIONS (I.E., ALPHA=BETA=0 IF MBDCND=3) IS SPECIFIED AND IF $CF(X)=0$ FOR ALL X THEN A SOLUTION TO THE DISCRETIZED MATRIX EQUATION MAY NOT EXIST (REFLECTING THE NON-UNIQUENESS OF SOLUTIONS TO THE PDE). PERTRB IS A CONSTANT CALCULATED AND SUBTRACTED FROM THE RIGHT HAND SIDE OF THE MATRIX EQUATION INSURING THE EXISTENCE OF A SOLUTION. SEPX4 COMPUTES THIS SOLUTION WHICH IS A WEIGHTED MINIMAL LEAST SQUARES SOLUTION TO THE ORIGINAL PROBLEM. IF SINGULARITY IS NOT DETECTED PERTRB=0.0 IS RETURNED BY SEPX4.

IERROR

AN ERROR FLAG THAT INDICATES INVALID INPUT PARAMETERS OR FAILURE TO FIND A SOLUTION

- = 0 NO ERROR
- = 1 IF A GREATER THAN B OR C GREATER THAN D
- = 2 IF MBDCND LESS THAN 0 OR MBDCND GREATER THAN 4
- = 3 IF NBDCND LESS THAN 0 OR NBDCND GREATER THAN 4
- = 4 IF ATTEMPT TO FIND A SOLUTION FAILS.
(THE LINEAR SYSTEM GENERATED IS NOT

- = 5 IF IDMN IS TOO SMALL (SEE DISCUSSION OF IDMN)
- = 6 IF M IS TOO SMALL OR TOO LARGE (SEE DISCUSSION OF M)
- = 7 IF N IS TOO SMALL (SEE DISCUSSION OF N)
- = 8 IF IORDER IS NOT 2 OR 4
- = 9 IF INTL IS NOT 0 OR 1
- = 10 IF AFUN IS LESS THAN OR EQUAL TO ZERO FOR SOME INTERIOR MESH POINT XI
SOME INTERIOR MESH POINT (XI, YJ)
- = 11 IF THE WORK SPACE LENGTH INPUT IN W(1) IS LESS THAN THE EXACT MINIMAL WORK SPACE LENGTH REQUIRED OUTPUT IN W(1).
- = 12 IF MBDCND=0 AND AF(X)=CF(X)=CONSTANT OR BF(X)=0 FOR ALL X IS NOT TRUE.

SPECIAL CONDITIONS NONE

COMMON BLOCKS SPL4

DETAILED NONE

PRECISION SINGLE

REQUIRED LIBRARY NONE
FILES

SPECIALIST JOHN C. ADAMS, NCAR, BOULDER, COLORADO 80307

LANGUAGE FORTRAN

ENTRY POINTS SEPX4, SPELI4, CHKPR4, CHKSN4, ORTHO4, MINSO4, TRIS4,
DEFE4, DX4, DY4

HISTORY SEPX4 WAS DEVELOPED BY MODIFYING THE ULIB ROUTINE SEPELI DURING OCTOBER 1978. IT SHOULD BE USED INSTEAD OF SEPELI WHENEVER POSSIBLE. THE INCREASE IN SPEED IS AT LEAST A FACTOR OF THREE.

ALGORITHM SEPX4 AUTOMATICALLY DISCRETIZES THE SEPARABLE ELLIPTIC EQUATION WHICH IS THEN SOLVED BY A GENERALIZED CYCLIC REDUCTION ALGORITHM IN THE SUBROUTINE POIS. THE FOURTH ORDER SOLUTION IS OBTAINED USING THE TECHNIQUE OF DEFERRED CORRECTIONS REFERENCED BELOW.

REFERENCES KELLER, H.B., NUMERICAL METHODS FOR TWO-POINT BOUNDARY-VALUE PROBLEMS, BLAISDEL (1968), WALTHAM, MASS.

SWARZTRAUBER, P., AND R. SWEET (1975): EFFICIENT FORTRAN SUBPROGRAMS FOR THE SOLUTION OF ELLIPTIC PARTIAL DIFFERENTIAL EQUATIONS. NCAR TECHNICAL NOTE NCAR-TN/IA-109, PP. 135-137.

Bibliografía

- [1] J.L. Lions, "*Optimal Control of Systems Governed by Partial Differential Equations*". Springer, 1971.
- [2] J.L. Lions, "*Exact Controllability, Stabilization and Perturbation for Distributed Systems*". SIAM Review, 30 (1): 1-68, 1988.
- [3] R. Glowinski, C.H. Li, J.L. Lions, "*A Numerical Approach to the Exact Controllability of the wave Equation (I) Dirichlet Controls: Description of the Numerical Methods*". Japan J. Appl. Math., 7, (1990), 1, pp. 1-76.
- [4] Nicolás Carrizosa Alfredo, "*Numerical Aspects of some time dependent Partial Differential Equations Problems*" Thesis presented to the Faculty of the Department of Mathematics, University of Houston, 1988.
- [5] R. Glowinski, W. Kinton and M.F. Wheeler, "*Mixed Finite Element Methods for Times Dependent Problems: Application to Control.*", Int. J. Num., Meth. Eng., 27, 623 (1989).
- [6] Martin Berggren, "*Control and Simulation of Advection-Diffusion Systems*". Thesis presented to the Faculty of the Department of Mechanical Engineering. University of Houston, 1992.
- [7] R. Glowinski, and J.L. Lions "*Exact and approximate controllability for distributed parameter systems*". Acta Numérica (1994), pp. 269-378.
- [8] A.N. Kolmogorov, S.V. Fomin, "*Elementos de la teoría de Funciones y del Análisis Funcional*". Mir Moscú, 1978.
- [9] C. Baiocchi, A. Capelo, "*Variational and Quasivariational Inequalities*". John Wiley and Sons, 1984.
- [10] Reed, Simon, "*Methods of Modern Mathematical Physics*" Vol. I, Functional Analysis. Academic Press, 1972.
- [11] Keller, H.B., "*Numerical Methods for two-point Boundary-Value Problems*". Blaisdel, 1968.

- [12] M. Al-Baali, "*Descent Property and Global Convergence of the Fletcher-Reeves Method with Inexact Line Search*". IMA Journal of Numerical Analysis, 5: 121-124, '985.
- [13] J.E. Dennis and R.B. Schnabel, "*Numerical Methods for Unconstrained Optimization and Non-linear Equations*". Prentice-Hall, Englewood Cliffs, N.J., 1983.
- [14] J. Douglas and H. Rachford, "*The Numerical solution of parabolic and elliptic differential equations*". J. SIAM, 3:28-41, 1955.
- [15] R. Glowinski, "*Numerical Methods for Nonlinear Variational Problems*". Springer-Verlag, New York, 1984.
- [16] J.L. Lions and E. Magens, "*Non-Homogeneous Boundary Value Problems and Applications*". Springer, Berlin and New York, 1972.
- [17] J. Lions, "*Some Aspects of the Optimal Control of Distributed Parameter Systems*". Regional Conference Series in Applied Mathematics. SIAM, Philadelphia, Pennsylvania, 1972.
- [18] J. Douglas and H. Rachford. "*The Numerical Solution of Parabolic and Elliptic Differential Equations*". J. SIAM, 3:28-41, 1955.
- [19] H.W. Schey, "*Div, Grad, Curl and all that; Vector Calculus*". W.W. Norton and Company, 1992.
- [20] F. Abergel and R. Teman. "*On Some Control Problems in Fluid Mechanics*". Theoretical and Computational Fluid Dynamics, pp. 303-325, 1990.