



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

SOLUCIÓN DE LA ECUACIÓN DE
CONVECCIÓN-DIFUSIÓN MEDIANTE FUNCIONES
DE BASE RADIAL

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
Licenciado en Ciencias de la Computación

PRESENTA:

Jonnathan Daniel Rivera Ruiz

DIRECTOR DE TESIS:

Dr. Pedro Gonzalez Casanova Henríquez



México, D.F., 2015



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

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

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

Resumen

El presente trabajo tiene por objetivo el desarrollo e implementación en código de diversos esquemas de solución numérica de EDPs basados en FBRs, tanto globales como locales, para la solución de problemas de Poisson y Convección-Difusión en una formulación directa y de control óptimo.

Se considera un conjunto de problemas pertenecientes a éstas clases presentes en la literatura actual a manera de comparación. En cuanto a ésta, se detectaron muy pocos trabajos referentes a la solución numérica de ésta formulación del problema de control, notablemente (1, 5, 6, 7), de entre los cuales sólo (5) hace uso de FBRs vía un método global.

De esta manera éste trabajo viene a esclarecer el panorama de lo que es posible lograr con métodos radiales, particularmente bajo esquemas locales como la interpolación local de Hermite (LHI), que no se ha utilizado, hasta donde es de nuestro conocimiento, para problemas de control de ninguna clase.

Debido a ésto el presente trabajo también participa en expandir el terreno conocido en cuanto a los alcances de LHI, sobre el cual se han realizado pocos trabajos desde su introducción en (9) pese a que en todos ellos se ha alzado como una prometedora nueva opción.

Hacia el final del trabajo vemos que efectivamente LHI posee una ventaja significativa en todos los aspectos, por sobre la colocación asimétrica de Kansa, en problemas que inherentemente necesitan de un gran número de centros, concluyendo que, en el ámbito de la solución numérica del problema de control, los métodos de base radial tienen un alcance mucho mayor al mostrado en la literatura previa vía LHI, y desde el punto de vista del estudio de LHI mismo, que ésta clase de problemas posee una variedad de ejemplos donde se hace patente que su diseño local le da una ventaja con la que los métodos globales no pueden competir.

Índice general

1. Interpolación y solución numérica de EDPs mediante FBRs	1
1.1. Funciones de base radial	1
1.2. El enfoque Hilbertiano abstracto para el problema de interpolación. . .	2
1.3. Esquemas de colocación para solución de EDPs	5
1.4. Interpolación Local de Hermite	6
2. Solución de las ecuaciones de Poisson y convección-difusión mediante interpolación local de Hermite	9
2.1. Descripción del código	10
2.1.1. Colocación de los nodos	10
2.1.2. Determinación de los vecinos más cercanos	11
2.1.3. Solución de los sistemas locales	11
2.1.4. Solución del sistema global	12
2.2. Aspectos teóricos de la implementación	13
2.3. Casos de prueba para problemas de tipo Poisson	15
2.3.1. Condiciones Dirichlet	15
2.3.2. Condiciones mixtas	16
3. Solución del problema de control óptimo de tipo Poisson y convección-difusión mediante interpolación local de Hermite	17
3.1. El problema de minimización	17
3.2. La derivada de Fréchet	18
3.3. El enfoque del Lagrangiano	20
3.4. Solución numérica del problema de control	22
3.4.1. Solución con colocación de Kansa	22
3.4.2. Solución con interpolación local de Hermite	23
3.5. Casos de prueba para problemas de tipo Poisson	25
3.5.1. Problema 1	25
3.5.2. Problema 2	27
3.5.3. Problema 3	29
3.5.4. Problema 4	31
3.6. Casos de prueba para problemas de tipo Convección-Difusión.	35
3.6.1. Problema 5	35

ÍNDICE GENERAL

3.6.2. Problema 6	38
4. Conclusiones	41
A. Nociones básicas de análisis funcional	43
A.1. Espacios lineales	43
A.2. Espacios métricos, normados y con producto interior.	47
A.3. Espacios de Hilbert	54
B. Sobre el operador \mathcal{M}	59
Bibliografía	61

Interpolación y solución numérica de EDPs mediante FBRs

En este capítulo se plantean los problemas de interpolación y solución de ecuaciones en derivadas parciales y se presenta el enfoque Hilbertiano para obtener soluciones en términos de una función de base radial, que se presentarán a continuación.

1.1. Funciones de base radial

Una función de base radial Φ es una función real definida en un subconjunto $\Omega \subset \mathbb{R}^d$ de la forma $\Phi(x) = \phi(\|x - \mathbf{x}\|)$, donde $\mathbf{x} \in \mathbb{R}^d$ y $\phi: \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}$. A \mathbf{x} se le llama el centro -o base- de la misma, de forma que la función solamente depende de la distancia radial con respecto a un punto base, de donde se desprende el nombre FBR.

Ejemplo 1.1.1. Las siguientes son algunas de las funciones de base radial más utilizadas:

1. Funciones Gaussianas

$$\Phi(x) = \exp\left(-\frac{\|x - \mathbf{x}\|^2}{c}\right)$$

2. Muticuádrico

$$\Phi(x) = \sqrt{\|x - \mathbf{x}\|^2 + c}$$

3. Muticuádrico inverso

$$\Phi(x) = \frac{1}{\sqrt{\|x - \mathbf{x}\|^2 + c}}$$

4. Poliharmónicos

$$\Phi(x) = \begin{cases} \|x - \mathbf{x}\|^k & \text{si } k \text{ es impar} \\ \|x - \mathbf{x}\|^k \ln \|x - \mathbf{x}\| & \text{si } k \text{ es par} \end{cases}$$

Particularmente el spline placa delgada, correspondiente a $k = 2$ y dado por

$$\Phi(x) = \|x - \mathbf{x}\|^2 \ln \|x - \mathbf{x}\|$$

Como el lector seguramente habrá notado, en algunas de las FBR que acabamos de presentar aparece una constante (positiva) c , que se conoce como parámetro de forma. Éste permite, entre otras cosas, ajustar los valores de Φ a la escala de las distancias en el dominio que nos interesa, lo que contribuye a mejorar el condicionamiento de las matrices de Gramm que introduciremos en la siguiente sección. Por ahora nos limitamos a resaltar que la determinación del valor óptimo de éste es un problema difícil que se está estudiando activamente en éstos días.

1.2. El enfoque Hilbertiano abstracto para el problema de interpolación.

Un problema frecuente en muchas disciplinas es el de hacer inferencia sobre un conjunto de mediciones, en la forma de parejas $\{(\mathbf{x}_k, u_k) : \mathbf{x}_k \in \Omega \subset \mathbb{R}^d, u_k \in \mathbb{R}, 1 \leq k \leq n\}$. Una manera de lograrlo es obteniendo una función que se ajuste -en algún sentido apropiado- a los datos, y la forma más común de hacerlo buscar una función $u : \Omega \rightarrow \mathbb{R}$ tal que $u(\mathbf{x}_k) = u_k$. A tal función se le conoce como interpolante. Asimismo, es usual requerir también ciertas condiciones optimalidad para el interpolante, de suerte que no solo se busca una función que interpole sino que también minimice o maximice ciertos criterios, como puede ser la norma de la función en un espacio adecuado.

Utilizando la teoría de espacios de Hilbert el problema de interpolación se puede tratar de la siguiente manera: Sea $\mathbf{X} = \{\mathbf{x}_k\}_{k=1}^n \subset \Omega$ el conjunto de nodos y $\mathcal{H} = \mathcal{H}(\Omega)$ un espacio de Hilbert de funciones sobre Ω . Consideramos el subconjunto de funciones interpolantes dado por

$$V_I = \{u \in \mathcal{H} : u(\mathbf{x}_k) = u_k \text{ para toda } k = 1, \dots, n\}$$

Y el subespacio de \mathcal{H} asociado, $V_0 = \{u \in \mathcal{H} : u(\mathbf{x}_k) = 0 \text{ para toda } k = 1, \dots, n\}$. Entonces el interpolante que buscamos es una función $u \in V_I$ tal que

$$\|u\|_{\mathcal{H}} = \inf_{v \in V_I} \|v\|_{\mathcal{H}}$$

El teorema de proyección ortogonal implica entonces que u es el único elemento en $V_I \cap V_0^\perp$. Si el espacio \mathcal{H} posee un núcleo reproductor K , podemos escribir a V_I y V_0 como

$$\begin{aligned} V_I &= \{u \in \mathcal{H} : \langle u | K(\cdot, x_k) \rangle = u_k \text{ para toda } k = 1, \dots, n\} \\ V_0 &= \{u \in \mathcal{H} : \langle u | K(\cdot, x_k) \rangle = 0 \text{ para toda } k = 1, \dots, n\} \end{aligned}$$

y de esta forma, el complemento ortogonal del subespacio V_0 es un subespacio de \mathcal{H} de dimensión finita, dado por

$$V_0^\perp = \text{span} \{K(\cdot, x_k)\}_{k=1}^n$$

Así, el interpolante es de la forma

$$u(x) = \sum_{k=1}^n \lambda_k K(x, x_k)$$

En particular si se escoge como espacio de Hilbert el espacio nativo de una función radial Φ , el kernel reproductor será $K(x, y) = \Phi(x - y)$ y con ello el problema de interpolación se traduce en resolver el sistema de ecuaciones lineales $G\lambda = d$, donde

$$G = \begin{bmatrix} \Phi(x_1 - x_1) & \cdots & \Phi(x_1 - x_n) \\ \vdots & & \vdots \\ \Phi(x_n - x_1) & \cdots & \Phi(x_n - x_n) \end{bmatrix}$$

se conoce como la matriz de Gramm de Φ con respecto a X y $d = [u_1, \dots, u_n]^t$ es el vector de los datos. De esta manera la atención pasa a estudiar éste sistema lineal. Desafortunadamente sucede que para una buena cantidad de FBRs la matriz de Gramm no es invertible, pero su espacio nulo se ha determinado de manera precisa, como lo recoge la siguiente

Definición 1.2.1. Una función $\Phi : \Omega \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ se dice condicionalmente positiva definida de orden m si se tiene que

$$\sum_{j,k=1}^n c_j c_k \Phi(x_j - x_k) > 0$$

para cualquier $c \in \mathbb{R}^n \setminus \{0\}$ y cualquier colección de puntos $\{x_k\}_{k=1}^n \subset \Omega$ con la propiedad de que $\sum_{k=1}^n c_k p(x_k) = 0$ para todo $p \in P_m(\mathbb{R}^d)$ (el espacio de polinomios de grado menor a m en d variables). Ésta última condición se conoce como condición de momentos, y es útil hacer notar que basta con que $\sum_{k=1}^n c_k p_\ell(x_k) = 0$ para una base $\{p_\ell\}$ de $P_m(\mathbb{R}^d)$.

Ésta situación motiva la siguiente restricción sobre los nodos:

Definición 1.2.2. Decimos que un conjunto de puntos $\{x_k\}_{k=1}^n \subset \Omega$ es $P_m(\mathbb{R}^d)$ -unisolvente si el único elemento de dicho espacio de polinomios tal que $p(x_k) = 0$ para toda $k \in \{1, \dots, n\}$ es el polinomio cero.

Habiendo introducido éstas nociones es posible demostrar el siguiente resultado, fundamental para la obtención del interpolante en la práctica:

Teorema 1.2.3. Si Φ es una función condicionalmente positiva definida (en lo sucesivo abreviado como c.p.d) de orden m y $\{x_k\}_{k=1}^n \subset \Omega$ un conjunto de puntos $P_m(\mathbb{R}^d)$ -unisolvente, entonces la matriz de Gramm aumentada,

$$A = \begin{bmatrix} G & P \\ P^t & 0 \end{bmatrix}$$

donde $P_{k,\ell} = p_\ell(x_k)$, es positiva definida (en particular, invertible).

Demostración. Si λ, c son vectores -de dimensiones apropiadas- tales que

$$\begin{bmatrix} G & P \\ P^t & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ c \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

entonces λ verifica la condición de momentos, escrita como $P^t \lambda = 0$.

Así, $0 = \begin{bmatrix} \lambda^t & c^t \end{bmatrix} \begin{bmatrix} G & P \\ P^t & 0 \end{bmatrix} \begin{bmatrix} \lambda \\ c \end{bmatrix} = \lambda^t G \lambda + 2c^t P^t \lambda = \lambda^t G \lambda$ y como G es positiva definida para aquellos vectores que satisfacen $P^t \lambda = 0$, se obtiene que $\lambda = 0$. Luego se tiene que $Pc = 0$, lo que significaría que el polinomio $\sum_{\ell} c_\ell p_\ell(x)$ se anula en el conjunto de nodos. En virtud de la unisolvencia del mismo, $c = 0$ y con ello A es invertible. \square

Así queda demostrado que el problema de interpolación para una FBR Φ c.p.d. de orden m , tiene siempre una única solución, cuando el conjunto de nodos es $P_m(\mathbb{R}^d)$ -unisolvente, de la forma

$$u(x) = \lambda_1 \Phi(x - x_1) + \dots + \lambda_n \Phi(x - x_n) + p(x)$$

donde $p \in P_m(\mathbb{R}^d)$, lo que abre el camino a una buena cantidad de esquemas de interpolación al escoger Φ y, en particular, a esquemas de solución numérica de EDPs, como describiremos en la siguiente sección.

1.3. Esquemas de colocación para solución de EDPs

Como ya mencionamos, el objetivo central de este trabajo es resolver numéricamente problemas del tipo

$$\begin{aligned} \mathcal{L}u &= f & \text{en } \Omega \\ \mathcal{B}u &= g & \text{en } \partial\Omega \end{aligned} \quad (1.1)$$

donde \mathcal{L} , \mathcal{B} son operadores diferenciales en un espacio de funciones adecuado.

La manera más simple de obtener una solución es como lo propone Kansa en :

Se elige una FBR Φ c.p.d. de orden m , un conjunto $X \subset \bar{\Omega}$ de n nodos ($P_m(\mathbb{R}^d)$ -unisolvente), de entre los cuales n_b pertenecen a la frontera y n_i al interior de Ω ; y se considera un ansatz de la forma

$$\begin{aligned} u(x) &= \lambda_1 \Phi(x - x_1) + \cdots + \lambda_n \Phi(x - x_n) + \lambda_{n+1} p_1(x) + \cdots + \lambda_{n+n_p} p_{n_p}(x) \\ &= \sum_{k=1}^n \lambda_k \Phi(x - x_k) + \sum_{\ell=1}^{n_p} \lambda_{\ell+n} p_{\ell}(x) \end{aligned} \quad (1.2)$$

donde n_p es la dimensión de $P_m(\mathbb{R}^d)$ como \mathbb{R} espacio vectorial. Se exigen las siguientes condiciones, que se traducen en que u sea solución -aproximada- de el problema 2.1:

$$\begin{aligned} g(x_i) &= \mathcal{B}u(x_i) = \sum_{k=1}^n \lambda_k \mathcal{B}\Phi(x_i - x_k) + \sum_{\ell=1}^{n_p} \lambda_{\ell+n} \mathcal{B}p_{\ell}(x_i) & \text{para } x_i \text{ en la frontera de } \Omega \\ f(x_i) &= \mathcal{L}u(x_i) = \sum_{k=1}^n \lambda_k \mathcal{L}\Phi(x_i - x_k) + \sum_{\ell=1}^{n_p} \lambda_{\ell+n} \mathcal{L}p_{\ell}(x_i) & \text{para } x_i \text{ en el interior de } \Omega \\ 0 &= \sum_{k=1}^n \lambda_k p_{\ell}(x_k), & \text{las condiciones de momentos} \end{aligned}$$

Éste es un sistema algebraico en λ , que se puede reescribir como

$$\begin{bmatrix} \mathcal{B}\Phi_{ij} & \mathcal{B}P \\ \mathcal{L}\Phi_{ij} & \mathcal{L}P \\ P^t & 0 \end{bmatrix} \lambda = d$$

donde $d = \begin{bmatrix} g(x_j) & \dots & f(x_j) & \dots & 0 \\ 1 \leq j \leq n_b & & n_i < j \leq n & & 1 \leq \ell \leq n_p \end{bmatrix}^t \in \mathbb{R}^{n+n_p}$ es el vector de los datos.

Éste esquema de solución numérica, si bien es tan simple como efectivo, no se tiene hasta la fecha una demostración de que la matriz involucrada sea invertible. Esto aunado a que tampoco es simétrica y a que por lo general tiene un mal condicionamiento llevó al desarrollo de lo que se conoce como colocación simétrica de Wu, y que procede de la misma manera que la de Kansa pero considera un ansatz de la forma

$$u(x) = \sum_{k=1}^{n_b} \lambda_k \mathcal{B}_x \Phi(x - x_k) + \sum_{k=n_b+1}^n \lambda_k \mathcal{L}_x \Phi(x - x_k) + \sum_{\ell=1}^{n_p} \lambda_{\ell+n} p_{\ell}(x) \quad (1.3)$$

Éste da lugar al sistema lineal

$$\begin{bmatrix} \mathcal{B}_x \mathcal{B}_x \Phi_{ij} & \mathcal{B}_x \mathcal{L}_x \Phi_{ij} & \mathcal{B}P \\ \mathcal{L}_x \mathcal{B}_x \Phi_{ij} & \mathcal{L}_x \mathcal{L}_x \Phi_{ij} & \mathcal{L}P \\ \mathcal{B}P^t & \mathcal{L}P^t & 0 \end{bmatrix} \lambda = d$$

con el mismo vector de datos.

De esta forma se obtiene una matriz que sí es simétrica, y para la cual se conoce una prueba de su invertibilidad(2).

1.4. Interpolación Local de Hermite

En esta sección se describe en detalle el esquema de interpolación local de Hermite (9).

Consideramos un conjunto de n nodos $X \subset \bar{\Omega}$, de entre los cuales se selecciona un subconjunto X_c de tamaño n_c que jugarán el papel de los centros de ciertos discos con algún radio fijado de antemano. Resolveremos un problema de interpolación para cada disco, y es por ésto que el método se califica como local.

Estando en en k -ésimo disco D_k ,

$n^{(k)}$ es la cantidad de nodos en D_k , i.e. $\#(X \cap D_k)$.

$n_c^{(k)}$ es el número de centros en D_k , i.e. $\#(X_c \cap D_k)$.

$n_b^{(k)}$ es el número de puntos de D_k en $\partial\Omega$ que no son centros i.e. $\#(\partial\Omega \cap (D_k \setminus X_c))$.

$n_i^{(k)}$ es el número de puntos de D_k en el interior de Ω que no son centros, es decir, los restantes, de manera que $n^{(k)} = n_c^{(k)} + n_b^{(k)} + n_i^{(k)}$.

Enumeramos los nodos en D_k de forma que $\{x_1^{(k)}, \dots, x_{n_c^{(k)}}^{(k)}\}$ sean los centros y $x_1^{(k)}$ el centro de D_k , $\{x_{n_c^{(k)}+1}^{(k)}, \dots, x_{n_c^{(k)}+n_b^{(k)}}^{(k)}\}$ los nodos en la frontera y $\{x_{n_c^{(k)}+n_b^{(k)}+1}^{(k)}, \dots, x_{n^{(k)}}^{(k)}\}$ los nodos en el interior.

Proponemos una solución de la forma

$$\begin{aligned} u^{(k)}(x) &= \sum_{j=1}^{n_c^{(k)}} \lambda_j^{(k)} \Phi(x - x_j^{(k)}) + \sum_{j=n_c^{(k)}+1}^{n_c^{(k)}+n_b^{(k)}} \lambda_j^{(k)} \mathcal{B}_x \Phi(x - x_j^{(k)}) \\ &+ \sum_{j=n_c^{(k)}+n_b^{(k)}+1}^{n_c^{(k)}+n_b^{(k)}+n_i^{(k)}} \lambda_j^{(k)} \mathcal{L}_x \Phi(x - x_j^{(k)}) + \sum_{j=n^{(k)}+1}^{n^{(k)}+n_p} \lambda_j^{(k)} p_{j-n^{(k)}}(x) \end{aligned} \quad (1.4)$$

en donde Φ es una función radial condicionalmente positiva definida de orden m y $\{p_\ell\}_{\ell=1}^{n_p}$ una base del espacio de polinomios de grado menor que m .
Notemos que podemos escribir

$$u^{(k)}(x) = H^{(k)}(x)\lambda^{(k)}$$

con $\lambda^{(k)} \in \mathbb{R}^{n^{(k)}+n_p}$ el vector columna de los coeficientes que buscamos determinar y

$$H^{(k)}(x) = \begin{bmatrix} \Phi(x - x_j^{(k)}) \dots \mathcal{B}_x \Phi(x - x_j^{(k)}) \dots \mathcal{L}_x \Phi(x - x_j^{(k)}) \dots p_\ell(x) \\ 1 \leq j \leq n_c^{(k)} \quad n_c^{(k)} < j \leq n_c^{(k)} + n_b^{(k)} \quad n_c^{(k)} + n_b^{(k)} < j \leq n^{(k)} \quad 1 \leq \ell \leq n_p \end{bmatrix} \in \mathbb{R}^{n^{(k)}+n_p}$$

es el vector renglón conocido como vector de reconstrucción.

Ahora, como ilustramos anteriormente, en la mayoría de los esquemas de interpolación y solución numérica de EDPs el siguiente paso es obtener un sistema algebraico en el que la matriz de coeficientes sea conocida y luego pasar a resolverlo. Éste es el punto que distingue a LHI, plantearemos una familia de sistemas algebraicos locales en donde haga falta la información en los centros, y luego a través de éstos construiremos un sistema global cuya solución nos dará el valor de la solución en los centros.

Para ello consideramos el siguiente sistema de ecuaciones:

- $u(x_j^{(k)}) = H^{(k)}(x_j^{(k)})\lambda^{(k)}$ para $j = 1, \dots, n_c^{(k)}$. Ésta es la información pendiente.
- $g(x_j^{(k)}) = \mathcal{B}H^{(k)}(x_j^{(k)})\lambda^{(k)}$ para $j = n_c^{(k)} + 1, \dots, n_c^{(k)} + n_b^{(k)}$, de la condición de frontera.
- $f(x_j^{(k)}) = \mathcal{L}H^{(k)}(x_j^{(k)})\lambda^{(k)}$ para $j = n_c^{(k)} + n_b^{(k)} + 1, \dots, n^{(k)}$, de la EDP.
- $0 = \sum_{j=1}^{n_c^{(k)}} \lambda_j^{(k)} p_\ell(x_j^{(k)}) + \sum_{j=n_c^{(k)}+1}^{n_b^{(k)}+n_c^{(k)}} \lambda_j^{(k)} \mathcal{B}p_\ell(x_j^{(k)}) + \sum_{j=n_b^{(k)}+n_c^{(k)}+1}^{n_b^{(k)}+n_c^{(k)}+n_l^{(k)}} \lambda_j^{(k)} \mathcal{L}p_\ell(x_j^{(k)})$,
para $\ell = 1, \dots, n_p$, las conocidas condiciones de momentos.

en donde $\lambda^{(k)}$ es la incógnita y salvo los valores de u en $\{x_1^{(k)}, \dots, x_{n_c^{(k)}}^{(k)}\}$ -que aparecen en las primeras $n_c^{(k)}$ ecuaciones- todos los demás términos involucrados son conocidos. Escribamos este sistema algebraico como

$$A^{(k)}\lambda^{(k)} = d^{(k)} \tag{1.5}$$

con

$$d^{(k)}(x) = \begin{bmatrix} u(x_j^{(k)}) \dots g(x_j^{(k)}) \dots f(x_j^{(k)}) \dots 0 \\ 1 \leq j \leq n_c^{(k)} \quad n_c^{(k)} < j \leq n_c^{(k)} + n_b^{(k)} \quad n_c^{(k)} + n_b^{(k)} < j \leq n^{(k)} \quad 1 \leq \ell \leq n_p \end{bmatrix}^t \in \mathbb{R}^{n^{(k)}+n_p}$$

1. INTERPOLACIÓN Y SOLUCIÓN NUMÉRICA DE EDPS MEDIANTE FBRs

y $A^{(k)} \in \mathbb{R}^{(n^{(k)}+n_p) \times (n^{(k)}+n_p)}$ la matriz dada por

$$\begin{array}{c}
 i = 1 \\
 \vdots \\
 n_c^{(k)} \\
 \hline
 \vdots \\
 n_c^{(k)} + n_b^{(k)} \\
 \hline
 \vdots \\
 n^{(k)} \\
 \hline
 \vdots \\
 n^{(k)} + n_p
 \end{array}
 \left(
 \begin{array}{c|c|c|c}
 j = 1, \dots, n_c^{(k)} & \dots, n_c^{(k)} + n_b^{(k)} & \dots, n^{(k)} & \dots, n^{(k)} + n_p \\
 \hline
 \Phi_{ij} & \mathcal{B}_x \Phi_{ij} & \mathcal{L}_x \Phi_{ij} & P \\
 \hline
 \mathcal{B}_x \Phi_{ij} & \mathcal{B}_x \mathcal{B}_x \Phi_{ij} & \mathcal{B}_x \mathcal{L}_x \Phi_{ij} & \mathcal{B}P \\
 \hline
 \mathcal{L}_x \Phi_{ij} & \mathcal{L}_x \mathcal{B}_x \Phi_{ij} & \mathcal{L}_x \mathcal{L}_x \Phi_{ij} & \mathcal{L}P \\
 \hline
 P^t & \mathcal{B}P^t & \mathcal{L}P^t & 0
 \end{array}
 \right) \quad (1.6)$$

donde $P = [p_\ell(x_i^{(k)})] \in \mathbb{R}^{n_c^{(k)} \times n_p}$, $\Phi_{ij} = [\Phi(x_i^{(k)} - x_j^{(k)})] \in \mathbb{R}^{n_c^{(k)} \times n^{(k)}}$ y los subíndices x y x indican la aplicación del operador respecto a la primera y segunda variable. Así, como $A^{(k)}$ es una matriz simétrica y Φ una FBR c.p.d de orden m , $A^{(k)}$ resulta invertible. Despejando obtenemos que $\lambda^{(k)} = (A^{(k)})^{-1} d^{(k)}$ y con ello que

$$u^{(k)}(x) = H^{(k)}(x) (A^{(k)})^{-1} d^{(k)} = W^{(k)}(x) d^{(k)}$$

a $W^{(k)}$ se le conoce como el vector de pesos. Ésta última expresión nos permite calcular también $\mathcal{Q}u^{(k)}$ para cualquier operador diferencial \mathcal{Q} vía $\mathcal{Q}u^{(k)}(x) = (\mathcal{Q}W^{(k)})(x) d^{(k)}$. Llamemos $u_c = [u(x_1^{(k)})]_{k=1}^{n_c} \in \mathbb{R}^{n_c}$ el vector de los valores de u en cada uno de los centros (recordemos que $x_1^{(k)}$ es el centro de D_k). Entonces, para cada k , las entradas que no conocemos en $d^{(k)}$ son parte de u_c y basta determinar a éste último. Para esto consideremos el sistema de ecuaciones

$$h(x_1^{(k)}) = W_{\mathcal{Q}}^{(k)}(x_1^{(k)}) d^{(k)} \quad (1.7)$$

para $k = 1, \dots, n_c$, en donde $W_{\mathcal{Q}}^{(k)} = \mathcal{Q}W^{(k)}$ y la pareja (h, \mathcal{Q}) es (f, \mathcal{L}) si $x_1^{(k)}$ es un nodo interior y (g, \mathcal{B}) si es un nodo frontera.

Éste es un sistema lineal en u_c , que escribiremos como $Su_c = b$ y como en cada $d^{(k)}$ solo aparecen pocos centros -i.e. $n_c^{(k)}$ es relativamente pequeña- la matriz S es una matriz dispersa. Ésto permite el uso de técnicas específicamente desarrolladas para sistemas lineales de este tipo, que en general presentan mucha menor dificultad computacional que una matriz llena.

Solución de las ecuaciones de Poisson y convección-difusión mediante interpolación local de Hermite

En éste capítulo se describirá detalladamente la forma en la que se realizó la implementación del algoritmo de interpolación local de Hermite en C++ para obtener numéricamente la solución de ecuaciones de tipo Poisson y convección-difusión.

El programa que se describe a continuación obtiene la solución de un problema elíptico de la forma

$$\begin{aligned}\mathcal{L}u &= f && \text{en } \Omega \\ \mathcal{B}u &= g && \text{en } \partial\Omega\end{aligned}\tag{2.1}$$

con operadores \mathcal{L} y \mathcal{B} de la forma

$$\begin{aligned}\mathcal{L} &= \Delta - \omega \cdot \nabla \\ \mathcal{B} &= pI + q \cdot \nabla\end{aligned}\tag{2.2}$$

donde p es una función real, ω, q funciones vectoriales definidas en Ω que, junto con f y g , forman parte de los datos que proporciona el usuario.

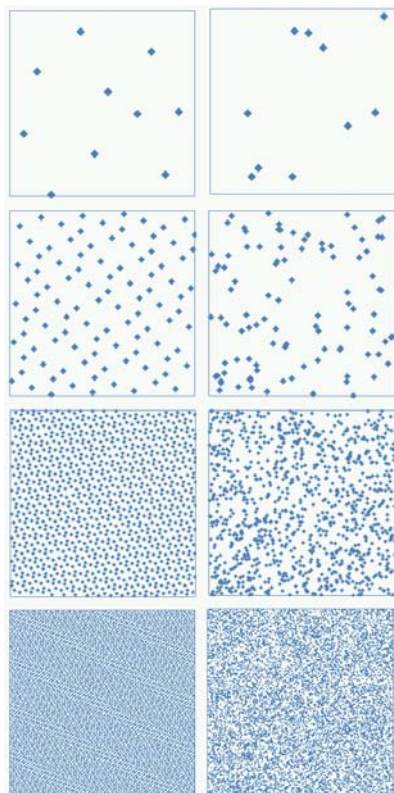
Ésta elección en la forma de \mathcal{L} y \mathcal{B} es lo suficientemente general como para permitir fácilmente el cambio entre condiciones de frontera de Dirichlet, Neumann y Robin, así como combinaciones de las mismas en secciones de $\partial\Omega$; así como pasar de la ecuación de Poisson a la de convección-difusión vía ω .

2.1. Descripción del código

2.1.1. Colocación de los nodos

El programa puede utilizar un conjunto de nodos prescrito por el usuario en forma de un archivo de texto donde cada línea es un punto y sus coordenadas se separan por un espacio en blanco.

Alternativamente, el programa puede generar nodos aleatorios en un dominio rectangular $\Omega = [x_{min}, x_{min}] = \{x \in \mathbb{R}^d : x_{min} \leq x \leq x_{min}\}$, donde el orden es lexicográfico. Si se elige ésta ruta se deben proporcionar los puntos x_{min}, x_{min} , la dimensión del espacio, d , y la cantidad de centros n_c en los que se va a obtener la solución.



Estos nodos se colocan de acuerdo a la sucesión de Sobol, que es sucesión pseudoaleatoria de baja discrepancia diseñada para obtener la convergencia más rápida de la aproximación del promedio

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N f(x_k) = \int_{[0,1]^d} f(x) dx$$

a la integral de una función $f \in L_1([0,1]^d)$.

Como se puede apreciar en la imagen, la sucesión de puntos Sobol, a la izquierda, llena de manera mucho más uniforme el cuadrado unitario que una sucesión de puntos aleatorios con distribución uniforme en el mismo dominio.

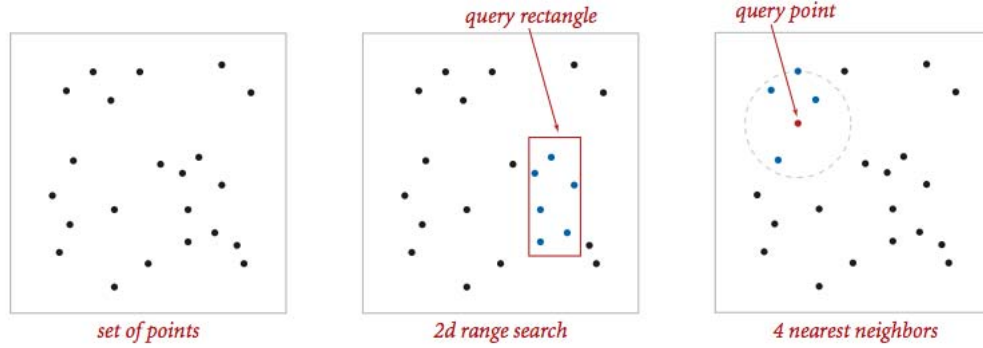
Esto resulta muy conveniente tanto para el número de condición de los sistemas locales, como para la visualización de la solución obtenida numéricamente. En el programa éstos se generan a través de la GNU Scientific Library (GSL).

2.1.2. Determinación de los vecinos más cercanos

El siguiente paso en el algoritmo LHI consiste en construir los sistemas locales, para lo cual es necesario determinar los n_v vecinos más cercanos a cada nodo de tipo centro en el conjunto de nodos del paso anterior.

Esto se realiza con un árbol kd, que es una estructura de datos diseñada para almacenar un conjunto de puntos d -dimensionales de manera que, entre otras cosas, el problema de determinar los puntos más cercanos a otro tenga complejidad logarítmica.

Se utiliza la implementación contenida en la biblioteca ANN: Approximate Nearest Neighbor, de David M. Mount y Sunil Arya, realizada en C++ y distribuida libremente.



De esta manera para cada uno de los n_c sistemas locales la determinación de los n_v nodos más cercanos a cada centro toma un tiempo $O(n_v \log n_c) = O(\log n_c)$, esto último debido a que n_v es constante respecto a la entrada.

2.1.3. Solución de los sistemas locales

Con el conocimiento de los vecinos más cercanos se construye la matriz del sistema local $A^{(k)}$ descrita en 1.6 y el vector de reconstrucción $\mathbf{H}_{\mathcal{L}}^{(k)} \left(\mathbf{x}_1^{(k)} \right)$ y se resuelve para el vector de pesos en el sistema

$$A^{(k)} \mathbf{W}_{\mathcal{L}}^{(k)} \left(\mathbf{x}_1^{(k)} \right) = \mathbf{H}_{\mathcal{L}}^{(k)} \left(\mathbf{x}_1^{(k)} \right)$$

mediante la rutina `lu_solve()` de la biblioteca GMM++, que por supuesto se basa en una factorización LU para resolver el sistema de ecuaciones y está optimizada para matrices llenas, como es el caso de $A^{(k)}$.

Si bien la solución de sistemas lineales mediante factorización LU toma un tiempo de orden cúbico, cabe notar que el tamaño de $A^{(k)}$ es constante e igual a $(n_v + n_p) \times (n_v + n_p)$, de suerte que podemos considerarlo constante con respecto a n_c .

Con el vector de pesos formamos el k -ésimo renglón de la matriz global S y el vector del lado derecho b . En cada iteración se realizan entonces $O((n_v + n_p)^3)$ operaciones, que en términos de el tamaño de la entrada, n_c , es $O(1)$.

2. SOLUCIÓN DE LAS ECUACIONES DE POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

2.1.4. Solución del sistema global

Alcanzado este punto se tiene construido el sistema global $Su_c = b$. Para explotar el que la que la matriz S que se obtiene es dispersa, se almacena en el formato de compresión CSR que explicamos a continuación:

En lugar de almacenar un arreglo de $n_c \times n_c$ de tipo double (flotante) se almacenan tres arreglos, que denominaremos Srow, Scol, Sval como sigue

	tipo	tamaño
Srow	int	$n_c + 1$
Scol	int	nnz
Sval	double	nnz

en donde nnz es la cantidad de entradas de S distintas de cero, que se va actualizando en cada iteración del paso anterior. En la k -ésima entrada de los arreglos Scol y Sval está almacenado el índice de la columna y el valor de la k -ésima entrada distinta de cero de S , mientras que en la j -ésima entrada de Srow se almacena la posición en el que comienzan los índices correspondientes al j -ésimo renglon de S en el arreglo Scol. Por ejemplo, la matriz

$$A = \begin{bmatrix} 1 & 3.8 & 0 & 0 \\ 0 & 4.2 & 6 & 0 \\ 2 & 5 & 7 & 8.6 \\ 0 & 0 & 0 & 9 \end{bmatrix}$$

se guarda en formato CSR como

$$\text{Arow} = [0, 2, 4, 8, 9]$$

$$\text{Acol} = [0, 1, 1, 2, 0, 1, 2, 3, 3]$$

$$\text{Aval} = [1, 3.8, 4.2, 6, 2, 5, 7, 8.6, 9]$$

El sistema global se resuelve entonces con la rutina `cusolverSp<double>csrslvlu()` de la biblioteca CUSOLVER de CUDA. Ésta está también basada en factorización LU, pero fuertemente optimizada para matrices con alto grado de dispersidad, a través de la cual se obtiene un rendimiento muy superior comparado con otras implementaciones basadas en factorización LU para matrices sparse, como la de la biblioteca CXSPARSE. Cabe destacar que de acuerdo con la documentación oficial de la versión 7.0 del SDK de CUDA, que es la que más reciente y con la que se llevó a cabo éste trabajo, la rutina `csrslvlu` ni siquiera hace uso de la tarjeta de video; la implementación que se ofrece -planeando cambiar para la próxima versión- lleva a cabo todos sus cálculos en un solo hilo del CPU.

2.2. Aspectos teóricos de la implementación

En esta sección comentamos las consideraciones restantes referentes a la implementación del esquema de solución mediante interpolación local de Hermite, que se encuentran más ligados con la teoría que con la forma en la que se le programó.

Como hemos venido mencionando, el usuario fija la cantidad de nodos de tipo centro, n_c al ejecutar el programa, a partir de la cual se genera una cantidad total de $n = 2n_c$ nodos, de los cuales el 50 % son, por supuesto, de tipo centro, 10 % de tipo frontera y el 40 % restante, nodos interiores. Éstas proporciones se pueden modificar fácilmente por el usuario pero no recomendamos destinar más de la mitad de los nodos a nodos centro, pues es a través de los otros dos tipos de nodos que se incorporan los datos del problema al sistema de ecuaciones que produce la solución.

El código está diseñado para permitir el uso de cualquier función de base radial arbitraria, ya sea estrictamente positiva definida o condicionalmente positiva definida de algún orden m . Se trabajó principalmente con el muticuadrado tradicional $\sqrt{r^2 + c}$, aunque también se probaron poliharmónicos y gaussianas. Ésta tendencia se debe en parte a las estimaciones del tamaño de los sistemas locales y -sobre todo- a la del parámetro de forma. Si bien la manera de estimarlos que se describe en esta sección es válida en general, es necesario hacer el análisis por separado.

Para la estimación automática de un tamaño adecuado de los sistemas locales fijamos un intervalo para n_c , digamos, $N_{\text{mín}} \leq n_c \leq N_{\text{máx}}$. En los extremos se determina experimentalmente un valor que consideremos óptimo para n_v , que denotaremos por $n_v(N_{\text{mín}})$, $n_v(N_{\text{máx}})$ respectivamente.

La aproximación que se propone entonces es de la forma

$$n_v(N) = \lfloor \gamma \log N + \sigma \rfloor$$

donde γ y σ se obtienen resolviendo el sistema de ecuaciones.

$$\begin{aligned} n_v(N_{\text{mín}}) &= \gamma \log N_{\text{mín}} + \sigma \\ n_v(N_{\text{máx}}) &= \gamma \log N_{\text{máx}} + \sigma \end{aligned}$$

Se ha encontrado que ésta elección del tamaño de los sistemas locales conserva las propiedades que se impusieron en los extremos. Por ejemplo, si lo que se buscó fue el tamaño mínimo tal que el error entre la solución y una solución analítica fuese menor que alguna cantidad, se tiene lo mismo en casi todo el interior del rango seleccionado para n_c ; las posibles excepciones ocurren cuando $\gamma \log N + \sigma$ está mucho más cerca de $\lfloor \gamma \log N + \sigma \rfloor$. Lo mismo se tiene si en lugar de optimizar para conseguir el mejor tiempo de ejecución se hace para disminuir el error en la solución.

En cuanto al parámetro de forma, se ha encontrado que para el multicuadrado una muy buena estimación constante -i.e. independiente de en qué sistema local nos encontremos- está dada por $c(n_c) = \frac{1}{n_c}$. Ésta resulta particularmente óptima si se hace uso de secuencias de baja discrepancia para la colocación de los nodos, como hacemos con la sucesión de puntos Sobol. Para otras FBR se utilizaron también estimaciones proporcionales al

2. SOLUCIÓN DE LAS ECUACIONES DE POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

inverso de la cantidad de nodos, pero los resultados muestran que un mejor balance entre tiempo de ejecución y error de la solución necesita de análisis más profundo.

Cabe aclarar, sin embargo, que éstas estimaciones son solamente una propuesta en este trabajo, pues como mencionamos al principio del mismo éstos son temas que se encuentran bajo una investigación muy activa en estos momentos. Debido a esto el programa está diseñado para recibir éstos dos parámetros del usuario de forma manual, como usualmente se hace, y solo recurre a los valores proporcionados por éstas estimaciones si no se le proporcionan.

2.3. Casos de prueba para problemas de tipo Poisson

2.3.1. Condiciones Dirichlet

Comenzamos las pruebas con siguiente función:

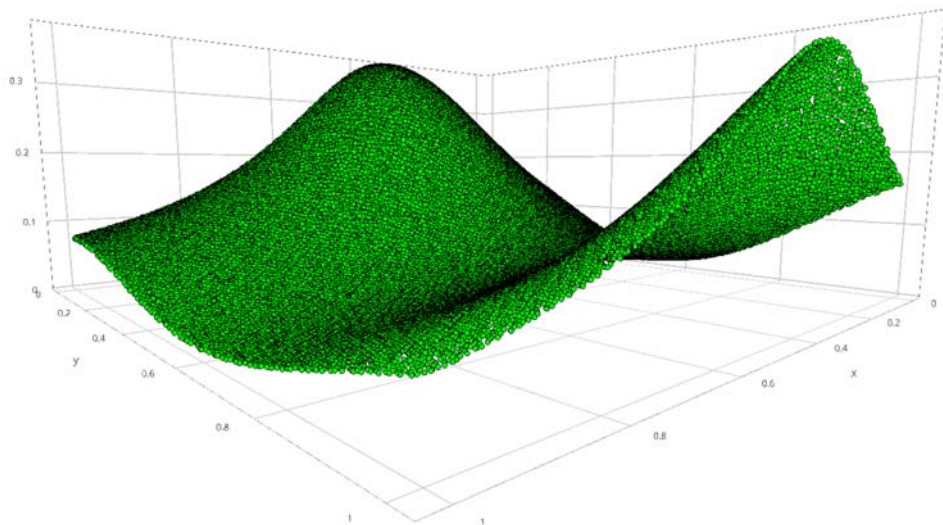
$$u(x, y) = \frac{1.25 + \cos(5.4y)}{6 + 6(3x - 1)^2}$$

Resolviendo la ecuación de Poisson que se obtiene de calcular el laplaciano de ésta función, con condiciones de Dirichlet en la frontera, en el dominio $\Omega = [0, 1]^2$ se obtienen los siguientes resultados al variar el número de nodos centro en donde vamos a recibir los valores de la solución:

n_c	n_v	c	Tiempo	Error L_∞	Error L_2	Error relativo
5000	12	10^{-3}	0.880s	3.407×10^{-2}	1.560×10^{-2}	1.065×10^{-1}
10000	25	5×10^{-5}	5.435s	1.481×10^{-2}	6.131×10^{-3}	4.185×10^{-2}
15000	32	3.33×10^{-5}	13.554s	5.259×10^{-3}	2.159×10^{-3}	1.474×10^{-2}
20000	37	2.5×10^{-5}	24.954s	5.761×10^{-3}	2.046×10^{-3}	1.396×10^{-2}
25000	42	2×10^{-5}	41.263s	2.804×10^{-3}	8.854×10^{-4}	6.045×10^{-3}
30000	48	1.67×10^{-5}	1m06.5s	1.853×10^{-3}	4.859×10^{-4}	3.341×10^{-3}

En esto se utilizó la estimación propuesta en la sección anterior, buscando el menor tiempo manteniendo un error aceptable para 18mil y 96mil nodos para los cuales se encontró experimentalmente que los valores óptimos eran 36 y 61, respectivamente. La siguiente figura muestra la solución correspondiente a la ultima entrada de la tabla:

Solución con 30 000 nodos



2. SOLUCIÓN DE LAS ECUACIONES DE POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

2.3.2. Condiciones mixtas

El siguiente ejemplo que se consideró fue una ecuación de Poisson con condiciones mixtas en la frontera de tipo Dirichlet y Robin como sigue:

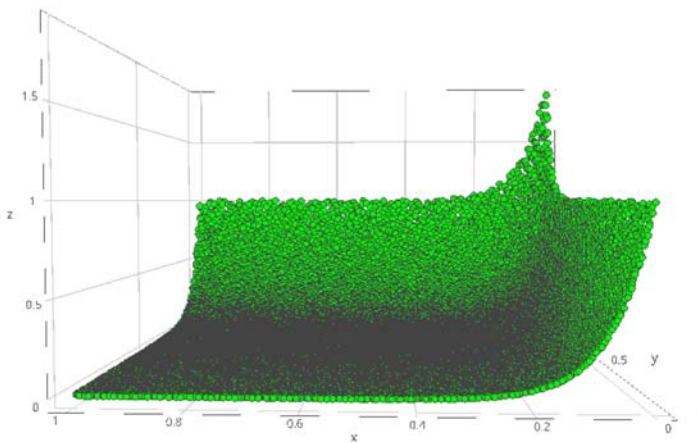
$$\begin{aligned}\Delta u &= 225e^{-15(x+y)} \text{ en } \Omega \\ \mathcal{B}u &= -15e^{-15(x+y)} \text{ en } \Gamma \\ u &= e^{-15(x+y)} \text{ en } \partial\Omega \setminus \Gamma\end{aligned}$$

donde Γ es el segmento de la recta $x = 1$ contenido en $\partial\Omega$ y $\mathcal{B} = \partial_x + \partial_y$.

Con la misma estimación del tamaño de los sistemas locales de la primera tabla del caso anterior se obtuvieron éstos resultados:

n_c	n_v	c	Tiempo	Error L_∞	Error L_2	Error relativo
10000	25	5×10^{-5}	5.415s	1.468×10^{-1}	2.718×10^{-2}	9.875×10^{-2}
15000	32	3.33×10^{-5}	13.579s	9.607×10^{-2}	1.711×10^{-2}	6.222×10^{-2}
20000	37	2.5×10^{-5}	25.034s	6.770×10^{-2}	2.549×10^{-2}	9.270×10^{-2}
25000	42	2×10^{-5}	41.629s	9.880×10^{-2}	1.775×10^{-2}	6.455×10^{-2}
30000	45	1.67×10^{-5}	1m0.432s	8.660×10^{-2}	9.322×10^{-3}	3.389×10^{-2}

Solución con 30 00 nodos



Solución del problema de control óptimo de tipo Poisson y convección-difusión mediante interpolación local de Hermite

3.1. El problema de minimización

Uno de los problemas más importantes en teoría de control se conoce como problema de control óptimo, que en el caso siguiente se traduce en la minimización de un funcional. Consideremos una ecuación diferencial parcial sobre una región $\Omega \subset \mathbb{R}^d$

$$\mathcal{L}y = u$$

Éste es un modelo de una situación en la que no podemos modificar directamente el estado del sistema -representado por la variable de estado y - y tenemos que hacerlo mediante un mecanismo externo, correspondiente a la variable u -conocida como variable de control- cuya acción sobre el estado del sistema está plasmada en una ecuación diferencial parcial $\mathcal{L}y = u$.

Por ejemplo pensemos en un sistema de ventilación que regula la distribución de temperatura en una cierta región del espacio, proporcionándonos control sobre ésta pero no de manera directa, sino a través de algunos parámetros en el aparato ventilador que sí podemos modificar, como puede ser el voltaje que se le suministra. El conocimiento de una ley que relaciona el voltaje con la capacidad del ventilador para regular la temperatura nos permitirá determinar en qué forma debemos suministrarlo de suerte que la distribución de temperatura sea lo más cercana a una distribución ideal fijada de antemano.

Una manera natural de hacerlo es buscar de entre las soluciones de $\mathcal{L}y = u$ aquella para la que $\|y - \hat{y}\|_{L^2(\Omega)}^2$ sea mínimo. Desafortunadamente tratar de minimizar solo la diferencia entre y y \hat{y} es un problema mal planteado (7) que se puede transformar en uno

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

bien planteado a través de una regularización de Tikhonov; esto es, si simultáneamente mantenemos una cota sobre la norma del control u y lo incorporamos al funcional de costo, cambiándolo por:

$$\frac{1}{2}\|y - \hat{y}\|_{L^2(\Omega)}^2 + \frac{\beta}{2}\|u\|_{L^2(\Omega)}^2$$

donde $\beta > 0$ es una constante de penalización que mantiene acotado al control; lo que también resulta de gran utilidad y aparece de manera natural en aplicaciones a problemas reales donde además se busca un control con norma acotada.

Finalmente introducimos algunos términos de uso común en la teoría de control: Las funciones y , u para las que el funcional en cuestión alcanza su mínimo se conocen como estado y control óptimos, respectivamente, y el objetivo de ésta teoría es no sólo estudiar las condiciones que garantizan la existencia de éstas, sino también proporcionar formas de obtenerlas explícitamente.

El problema que hemos descrito anteriormente es uno de la clase conocida como problemas de control distribuido, pues se tiene control por medio de u a lo largo de todo el dominio Ω . Cabe mencionar que también existe el caso con control sólo en la frontera, pero éste trabajo nos concentraremos en el caso distribuido.

3.2. La derivada de Fréchet

De la misma forma que en el análisis real, se puede mostrar que un funcional \mathcal{F} toma alguno de sus valores extremos en puntos donde su derivada de Fréchet es cero. Por lo tanto la ésta derivada jugará un papel fundamental en lo que nos ocupa y vale la pena recordar la definición y establecer los resultados que utilizaremos posteriormente.

Definición 3.2.1. Sean V, W espacios de Banach. Decimos que $\mathcal{F} : V \rightarrow W$ es Fréchet diferenciable en un punto x de su dominio si existe un operador lineal y continuo $A_x : V \rightarrow W$ tal que

$$\lim_{\|h\| \rightarrow 0} \frac{\|\mathcal{F}(x+h) - \mathcal{F}(x) - A_x h\|}{\|h\|} = 0$$

al cual llamaremos la derivada de Fréchet de \mathcal{F} en x y lo denotaremos como $\mathcal{D}\mathcal{F}(x) = A_x$.

Proposición 3.2.2. *En particular utilizaremos estos dos resultados:*

1. Si \mathcal{F} es lineal y acotado, $\mathcal{D}\mathcal{F}(x) = \mathcal{F}$ para todo x .
2. Si $\mathcal{F} : L_\infty(\Omega) \rightarrow \mathbb{R}$, con $\Omega \subset \mathbb{R}^d$ acotado, es un funcional de la forma $\mathcal{F}g = \int_{\Omega} F \circ g$, donde $F \in C^1(\mathbb{R})$, entonces $\mathcal{D}\mathcal{F}(g)$ está dado por $\mathcal{D}\mathcal{F}(g)h = \int_{\Omega} F'(g(x))h(x)dx$.

Demostración. .

1. Como \mathcal{F} es lineal, $\mathcal{F}(x + h) - \mathcal{F}(x) - Ah = \mathcal{F}(h) - Ah$ y como la derivada de Fréchet es única y \mathcal{F} es continuo, se sigue de la definición que $A_x = \mathcal{F}$.
2. Sean $g \in L_\infty(\Omega)$, A_g el funcional dado por $A_g h = \int_\Omega F'(g(x))h(x)dx$, $\epsilon > 0$ y $h \in L_\infty(\Omega)$ con $\|h\| \leq \delta$ para alguna $\delta > 0$ que fijaremos adelante. Elijamos $\rho > 0$ arbitrario, como F es de clase $C^1(\mathbb{R})$, es uniformemente diferenciable en el intervalo $[-\|g\| - \rho, \|g\| + \rho]$ y por lo tanto existe $\delta_0 > 0$ tal que

$$|F(a + \theta) - F(a) - F'(a)\theta| < |\theta| \frac{\epsilon}{2} \text{ para todo } a \in [-\|g\| - \rho, \|g\| + \rho] \text{ y } |\theta| < \delta_0$$

de manera que si $\delta < \min\{\rho, \delta_0\}$, entonces $\|F \circ (g + h) - F \circ g - (F' \circ g)h\| < \epsilon \|h\|$.

Así,

$$\begin{aligned} |\mathcal{F}(g + h) - \mathcal{F}(g) - A_g h| &= \left| \int_\Omega (F(g(x) + h(x)) - F(g(x)) - F'(g(x))h(x)) dx \right| \\ &\leq \int_\Omega |F(g(x) + h(x)) - F(g(x)) - F'(g(x))h(x)| dx \\ &< \epsilon \|h\| m(\Omega) \end{aligned}$$

como Ω es acotado su medida, $m(\Omega)$, es finita y se sigue la afirmación.

□

3.3. El enfoque del Lagrangiano

A continuación detallaremos el método que se conoce como de optimización-discretización para el caso de un operador elíptico. El objetivo es determinar el control óptimo para el problema de control distribuido

$$\min_{y,u} \frac{1}{2} \|y - \hat{y}\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \|u\|_{L^2(\Omega)}^2 \quad (3.1)$$

sueto a

$$\begin{aligned} \mathcal{L}y &= u & \text{en } \Omega \\ y &= g & \text{en } \partial\Omega \end{aligned} \quad (3.2)$$

Donde $\Omega \subset \mathbb{R}^d$ es un conjunto acotado, g es una función conocida y \mathcal{L} es una expresión diferencial de la forma

$$\mathcal{L}u = \sum_{j,k=1}^d \frac{\partial}{\partial x_k} \left(a_{k,j} \frac{\partial u}{\partial x_j} \right) + \sum_{j=1}^d b_j \frac{\partial u}{\partial x_j} + cu$$

con $a_{k,j}(x)$ una matriz simétrica y positiva (o negativa) definida. Definamos el adjunto formal de \mathcal{L} como

$$\mathcal{L}^*v = \sum_{j,k=1}^d \frac{\partial}{\partial x_k} \left(a_{k,j} \frac{\partial v}{\partial x_j} \right) - \sum_{j=1}^d \frac{\partial}{\partial x_j} (b_j v) + cv$$

Entonces tiene lugar una fórmula análoga a la segunda identidad de Green:

$$\int_{\Omega} (v\mathcal{L}u - u\mathcal{L}^*v) = \int_{\partial\Omega} \left(v \frac{\partial u}{\partial m} - u \frac{\partial v}{\partial m} + \varpi uv \right) \quad (3.3)$$

en donde n es el vector normal unitario exterior en la frontera de Ω , $\varpi = n \cdot b$ y m es el vector co-normal asociado a \mathcal{L} , dado por $m = [a_{k,j}]n$, de manera que

$$\frac{\partial}{\partial m} = \sum_{j,k=1}^d a_{k,j} n_j \frac{\partial}{\partial x_k}$$

Introduzcamos entonces un lagrangiano que incorpore el funcional de costo (3.1) y la EDP elíptica y condiciones de frontera como sigue:

$$\mathcal{L}(y, u, p_1, p_2) = \frac{1}{2} \|y - \hat{y}\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \|u\|_{L^2(\Omega)}^2 + \int_{\Omega} (\mathcal{L}y - u) p_1 + \int_{\partial\Omega} (y - g) p_2 \quad (3.4)$$

Donde p_1, p_2 son funciones conocidas como multiplicadores de Lagrange y fungen como un parámetro auxiliar para determinar el mínimo del lagrangiano.

Para tal efecto, tomemos la derivada de Fréchet de \mathcal{L} con respecto a y :

Notemos que los primeros dos sumandos -correspondientes al funcional de costo- son como en la proposición 3.2.2.2, mientras que los que involucran a los multiplicadores son lineales. Luego

$$\mathcal{D}_y \mathcal{L}(h) = \int_{\Omega} (y - \hat{y}) h + \int_{\Omega} (\mathcal{L}h) p_1 + \int_{\partial\Omega} h p_2$$

utilizando la identidad (3.3),

$$\int_{\Omega} (\mathcal{L}h) p_1 = \int_{\Omega} (\mathcal{L}^* p_1) h + \int_{\partial\Omega} \left(p_1 \frac{\partial h}{\partial m} - h \frac{\partial p_1}{\partial m} + \varpi h p_1 \right)$$

Así, si buscamos que $\mathcal{D}_y \mathcal{L}$ sea el operador 0, debe tenerse que $\mathcal{D}_y \mathcal{L}(h) = 0$ para toda función h ; en particular para aquellas con $h = \frac{\partial h}{\partial m} = 0$ en $\partial\Omega$. Al sustituir se obtiene que

$$0 = \mathcal{D}_y \mathcal{L}(h) = \int_{\Omega} (y - \hat{y}) h + \int_{\Omega} (\mathcal{L}^* p_1) h$$

esto satisface si

$$\mathcal{L}^* p_1 = \hat{y} - y \text{ en } \Omega \quad (3.5)$$

Con esta nueva información y suponiendo ahora solamente que $h = 0$ en $\partial\Omega$ obtenemos que

$$0 = \mathcal{D}_y \mathcal{L}(h) = \int_{\partial\Omega} p_1 \frac{\partial h}{\partial m}$$

que a su vez se cumple si exigimos que $p_1 = 0$ en $\partial\Omega$.

Finalmente, ahora para h arbitraria vemos que la ecuación $\mathcal{D}_y \mathcal{L}(h) = 0$ se reduce a

$$0 = \mathcal{D}_y \mathcal{L}(h) = \int_{\partial\Omega} \left(p_2 - \frac{\partial p_1}{\partial m} \right) h$$

de manera que p_2 queda determinado por p_1 a través de la relación $p_2 = \frac{\partial p_1}{\partial m}$ y por lo tanto podemos omitirlo y escribir $p = p_1$ para abreviar. Reescribamos el lagrangiano como

$$\mathcal{L}(y, u, p) = \frac{1}{2} \|y - \hat{y}\|_{L^2(\Omega)}^2 + \frac{\beta}{2} \|u\|_{L^2(\Omega)}^2 + \int_{\Omega} (\mathcal{L}y - u) p + \int_{\partial\Omega} (y - g) \frac{\partial p}{\partial m} \quad (3.6)$$

La ecuación 3.5 que acabamos de obtener se conoce como ecuación adjunta.

Diferenciando ahora con respecto a u e igualando a 0 obtenemos que:

$$0 = \mathcal{D}_u \mathcal{L}(h) = \int_{\Omega} (\beta u - p) h$$

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

que proporciona la ecuación gradiente

$$\begin{aligned} \beta u - p &= 0 \text{ en } \Omega \\ u &= 0 \text{ en } \partial\Omega \end{aligned} \quad (3.7)$$

Finalmente diferenciando con respecto a p recupera la ecuación de estado (3.2). De esta manera obtenemos el sistema

$$\begin{array}{l|l|l} \mathcal{L}y = u & \text{en } \Omega & \beta u - p = 0 & \text{en } \Omega & \mathcal{L}^*p = \hat{y} - y & \text{en } \Omega \\ y = g & \text{en } \partial\Omega & u = 0 & \text{en } \partial\Omega & p = 0 & \text{en } \partial\Omega \end{array} \quad (3.8)$$

cuyas soluciones son, por construcción, mínimos del funcional de costo. Resolveremos este sistema numéricamente a través de interpolación local de Hermite, como en el caso del problema directo del capítulo anterior.

3.4. Solución numérica del problema de control

En éste trabajo abordamos la solución numérica del problema de control óptimo descrito anteriormente utilizando tanto el esquema de colocación asimétrica de Kansa, como con Interpolación local de Hermite.

En ambos casos se observa que el sistema de EDPs 3.8 se puede reducir a

$$\begin{array}{l|l} \mathcal{L}y = u & \text{en } \Omega \\ y = g & \text{en } \partial\Omega \end{array} \quad \left| \quad \begin{array}{l} \beta \mathcal{L}^*u = \hat{y} - y & \text{en } \Omega \\ u = 0 & \text{en } \partial\Omega \end{array} \right. \quad (3.9)$$

3.4.1. Solución con colocación de Kansa

Para extender el esquema de Kansa a un sistema de EDPs consideraremos soluciones de la forma

$$y(x) = H(x)\lambda \quad (3.10)$$

$$u(x) = H(x)\mu \quad (3.11)$$

donde el vector de reconstrucción se toma como en el caso de una sola ecuación cuando se utiliza colocación de Kansa, esto es:

$$H^{(k)}(x) = \left[\begin{array}{c} \Phi(x - x_j) \dots p_\ell(x) \\ 1 \leq j \leq n \quad 1 \leq \ell \leq n_p \end{array} \right] \in \mathbb{R}^{n+n_p}$$

El sistema de algebraico que resulta es de la forma

$$\begin{bmatrix} A & \beta E_* \\ -E & A \end{bmatrix} \begin{bmatrix} \lambda \\ \mu \end{bmatrix} = \begin{bmatrix} d \\ 0 \end{bmatrix} \quad (3.12)$$

En donde $E = \begin{bmatrix} 0 \\ \mathcal{L}H(x) \\ 0 \end{bmatrix}$, $E_* = \begin{bmatrix} 0 \\ \mathcal{L}^*H(x) \\ 0 \end{bmatrix}$ son matrices cuadradas cuyas primeras n_b filas, correspondientes a los nodos en la frontera, son 0; A es la matriz de Gramm aumentada con respecto al conjunto de nodos en cuestión y $d = \begin{bmatrix} g(x_j) & \dots & \hat{y}(x_j) & \dots & 0 \\ 1 \leq j \leq n_b & & n_b+1 \leq j \leq n & & 1 \leq \ell \leq n_p \end{bmatrix}^t$.

Resolvemos éste sistema factorizando la matriz como

$$\begin{bmatrix} A & \beta E_* \\ -E & A \end{bmatrix} = \begin{bmatrix} I_{n+n_p} & 0 \\ -EA^{-1} & I_{n+n_p} \end{bmatrix} \begin{bmatrix} A & \beta E_* \\ 0 & R \end{bmatrix}$$

en donde $R = A + \beta EA^{-1}E_*$ es el complemento de Schur de A .

3.4.2. Solución con interpolación local de Hermite

La solución del problema 3.9 con LHI se llevó a cabo de manera esencialmente distinta a la extensión obvia (que es de la forma 3.10, cambiando H), debido a que la matriz que se obtiene de ésta es singular.

En lugar de considerar ese sistema, resolveremos el problema

$$\begin{aligned} \mathcal{M}y &= \hat{y} & \text{en } \Omega \\ \mathcal{L}y &= 0 & \text{en } \partial\Omega \\ y &= g & \text{en } \partial\Omega \end{aligned} \quad (3.13)$$

en donde el operador \mathcal{M} está dado por $\mathcal{M} = I + \beta \mathcal{L}^* \mathcal{L}$.

Para obtener la solución procedemos de la misma forma que como en LHI pero cambiando el vector de reconstrucción por

$$H^{(k)}(x) = \begin{bmatrix} \Phi \left(x - x_j^{(k)} \right) \dots p_\ell(x) \\ 1 \leq j \leq n^{(k)} & 1 \leq \ell \leq n_p \end{bmatrix} \in \mathbb{R}^{n^{(k)}+n_p}$$

con $n_b^{(k)} = n_{b1}^{(k)} + n_{b2}^{(k)}$ de forma que la matriz de los sistemas locales ya no será simétrica y éstos estarán dados por

$$A^{(k)} \lambda = \begin{bmatrix} \Phi_{i,j} & P \\ \mathcal{B}_x \Phi_{i,j} & \mathcal{B}P \\ \mathcal{L}_x \Phi_{i,j} & \mathcal{L}P \\ \mathcal{M}_x \Phi_{i,j} & \mathcal{M}P \\ P^t & 0 \end{bmatrix} \lambda = d$$

en donde $\mathcal{B} = I$, es el operador identidad, y el vector de los datos está dado por

$$d^{(k)}(x) = \begin{bmatrix} y \left(x_j^{(k)} \right) \dots g \left(x_j^{(k)} \right) \dots 0 \dots \hat{y} \left(x_j^{(k)} \right) \dots 0 \\ 1 \leq j \leq n_c^{(k)} & n_c^{(k)} < j \leq n_c^{(k)} + n_{b1}^{(k)} & n_c^{(k)} + n_{b1}^{(k)} < j \leq n_c^{(k)} + n_b^{(k)} & n_c^{(k)} + n_b^{(k)} < j \leq n^{(k)} & 1 \leq \ell \leq n_p \end{bmatrix}^t$$

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

Se construye de la misma manera un sistema global sparse y resolviéndolo se consigue el vector solución y_c .

Vale la pena aclarar que, si bien la pérdida de la simetría de las matrices locales respecto al esquema original de LHI nos deja sin un resultado general que demuestre su invertibilidad, de acuerdo con (3) en la práctica no se observa un deterioro significativo de la solución; mientras que supone una tarea de programación más llevadera pues siguiendo el ansatz simétrico habría que calcular \mathcal{M}^2 e involucraría derivadas de orden 8 sin una ganancia para compensar tal esfuerzo. De esta manera, llegado este punto ya se cuenta con los valores de el estado en los centros.

Se proponen las siguientes formas de proceder:

1. Resolver el sistema restante para u , que es

$$\begin{aligned} \beta \mathcal{L}^* u &= \hat{y} - y && \text{en } \Omega \\ u &= 0 && \text{en } \partial\Omega \end{aligned}$$

recordando que sólo se tienen los valores de y en ciertos nodos.

Para ésto uno puede cambiar los papeles de los nodos centro e interiores que se utilizaron cuando se resolvió para y , de forma que se obtendrán los valores de u en los que habíamos denominado nodos interiores.

Ésto tiene el inconveniente de calcular el control óptimo en una cantidad menor de nodos -pues solo 30 % son nodos interiores- así como el de resolver dos problemas elípticos y no uno solo.

2. Realizar la evaluación de

$$u = \mathcal{L}y$$

utilizando cuadratura diferencial (8) para discretizar el operador \mathcal{L} . Para ésto utilizaremos otra FBR que llamaremos Ψ con parámetro de forma τ . El inconveniente de éste enfoque es que cuadratura diferencial requiere resolver un sistema lineal de tamaño global por cada nodo (es decir es, en principio, de orden cúbico en n_c). Ésto se suele remediar considerando solamente éstos sistemas lineales para una vecindad alrededor de cada nodo. Si se utiliza una FBR de soporte compacto que esté contenido en dicha vecindad, no hay pérdida al restringirse al sistema local, mientras que si se utilizan FBRs como, por ejemplo, gaussianas con parámetro de forma adecuado, lo que se deja fuera del sistema local es suficientemente chico como para desdeñarlo sin pérdida significativa. De esta manera calcular el control u con cuadratura diferencial es lineal en n_c .

En ésta implementación hemos usado el mismo valor de n_v para la primera etapa en la que se obtiene y con LHI y la segunda de cuadratura diferencial para los sistemas locales.

En adelante nos referiremos al esquema 1 como LHI-LHI y al 2 como LHI-DQ, en alusión a cómo se obtiene el estado y el control en la primera y segunda etapa, respectivamente.

Todas las tablas siguientes se realizaron con precisión cuádruple (32 dígitos decimales).

3.5. Casos de prueba para problemas de tipo Poisson

Los siguientes 3 problemas que se presentan fueron estudiados en (7) con elemento finito y en (5) con radiales. Cabe destacar que este último es el único trabajo que aborda problemas de control con FBRs de el que se tuvo conocimiento durante la realización de el presente. Debido a ésto incluiremos sus resultados, obtenidos con Kansa acoplado y nodos equiespaciados, en la comparación entre Kansa acoplado, desacoplado, LHI-LHI, LHI-DQ.

3.5.1. Problema 1

$$\begin{aligned}
 -\Delta y &= u, & -\beta \Delta u &= y - \hat{y} & \text{en } \Omega \\
 y &= g, & u &= 0 & \text{en } \partial\Omega \\
 \hat{y} &= \sin \pi x_1 \sin \pi x_2 \\
 g &= 0
 \end{aligned} \tag{3.14}$$

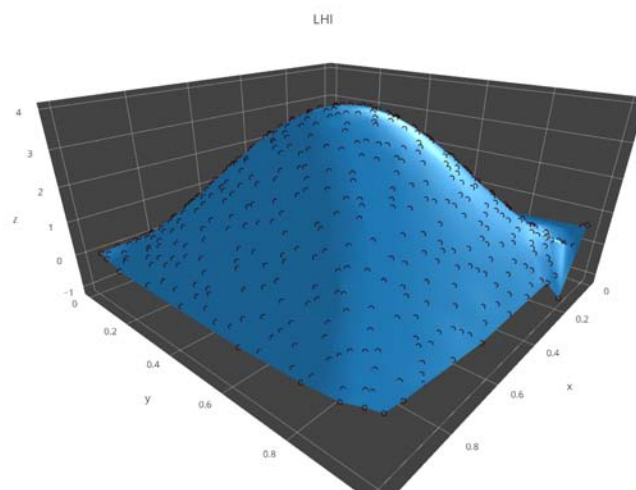
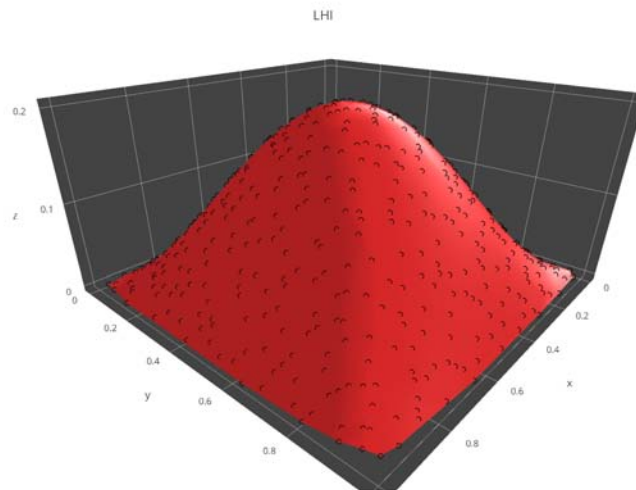
En las tablas correspondientes a este problema se reportan los valores de $\frac{\|y-y_\beta\|}{\|y_\beta\|}$ para y y $\frac{\|u-u_\beta\|}{\|u_\beta\|}$ para u , donde y_β, u_β son las soluciones exactas para cada valor de β . Para 625 nodos, $\beta = 10^{-2}$, $n_v = 50$ para LHI y gaussianas para DQ.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Ga	$c = 40$ $y : 3.24 \times 10^{-4}$ $u : 1.52 \times 10^{-4}$ no reportado.	$c = 40$ $y : 5.63 \times 10^{-5}$ $u : 5.52 \times 10^{-5}$ tiempo: 1m23s	$c = 4$ $y : 1.59 \times 10^{-5}$ $u : 7.52 \times 10^{-5}$ tiempo: 15s	$c = 4$ $y : 3.75 \times 10^{-3}$ $u : 1.54 \times 10^{-2}$ tiempo: 7s	$c = 4$ $y : 2.84 \times 10^{-3}$ $u : 9.01 \times 10^{-1}$ tiempo: 7s
	no reportado.	$c = 0.5$ $y : 5.64 \times 10^{-6}$ $u : 4.54 \times 10^{-6}$ tiempo: 1m25s	$c = 0.5$ $y : 4.85 \times 10^{-5}$ $u : 8.72 \times 10^{-5}$ tiempo: 15s	$c = 0.5$ $y : 5.58 \times 10^{-3}$ $u : 3.75 \times 10^{-2}$ tiempo: 7s	$c = 0.5$ $y : 5.65 \times 10^{-3}$ $u : 5.75 \times 10^{-2}$ tiempo: 7s

Para 625 nodos, $\beta = 10^{-5}$, $n_v = 50$ para LHI.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Ga	$c = 30$ $y : 9.21 \times 10^{-6}$ $u : 1.43 \times 10^{-4}$ no reportado.	$c = 30$ $y : 3.87 \times 10^{-6}$ $u : 5.55 \times 10^{-5}$ tiempo: 1m23s	$c = 4$ $y : 3.40 \times 10^{-6}$ $u : 2.94 \times 10^{-3}$ tiempo: 15s	$c = 4$ $y : 1.65 \times 10^{-5}$ $u : 2.83 \times 10^{+1}$ tiempo: 7s	$c = 4$ $y : 1.74 \times 10^{-5}$ $u : 3.83 \times 10^{-3}$ tiempo: 7s
	no reportado.	$c = 0.5$ $y : 4.73 \times 10^{-5}$ $u : 8.40 \times 10^{-4}$ tiempo: 1m24s	$c = 0.5$ $y : 2.77 \times 10^{-6}$ $u : 4.97 \times 10^{-3}$ tiempo: 15s	$c = 0.5$ $y : 1.52 \times 10^{-4}$ $u : 3.37 \times 10^{+2}$ tiempo: 7s	$c = 0.5$ $y : 1.40 \times 10^{-4}$ $u : 1.94 \times 10^{-3}$ tiempo: 7s

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



Solución obtenida vía LHI-DQ para $\beta = 10^{-2}$, 625 nodos, 50 en cada sistema local y utilizando multicuadrado para y y gaussianas para u . No se observa ningún inconveniente en resolver utilizando Kansa en este ejemplo

3.5.2. Problema 2

$$\begin{aligned}
 -\Delta y &= u, & -\beta \Delta u &= y - \hat{y} & \text{en } \Omega \\
 y &= g, & u &= 0 & \text{en } \partial\Omega \\
 \hat{y} &= \begin{cases} 1 & \text{en } [0, \frac{1}{2}]^2 \cap \Omega \\ 0 & \text{en otro caso} \end{cases} \\
 g &= \begin{cases} 1 & \text{en } [0, \frac{1}{2}]^2 \cap \partial\Omega \\ 0 & \text{en otro caso} \end{cases}
 \end{aligned} \tag{3.15}$$

En las tablas correspondientes a este problema se reportan los valores de $\|y - \hat{y}\|_{l_2}$ para y , $\|u\|_{l_2}$ para u y el funcional de costo, $\frac{1}{2}\|y - \hat{y}\|_{l_2}^2 + \frac{\beta}{2}\|u\|_{l_2}^2$, donde la norma $\|\cdot\|_{l_2}$ se ha calculado como

$$\|u\|_{l_2}^2 = \sum_{k=n_b+1}^n |u(x_k)|^2$$

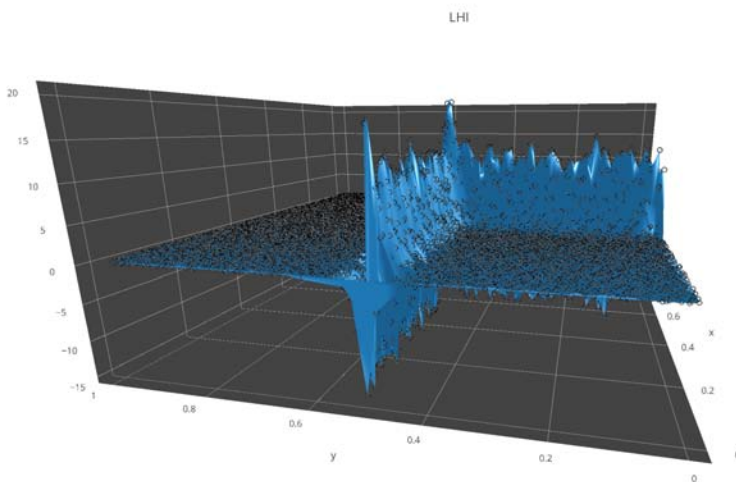
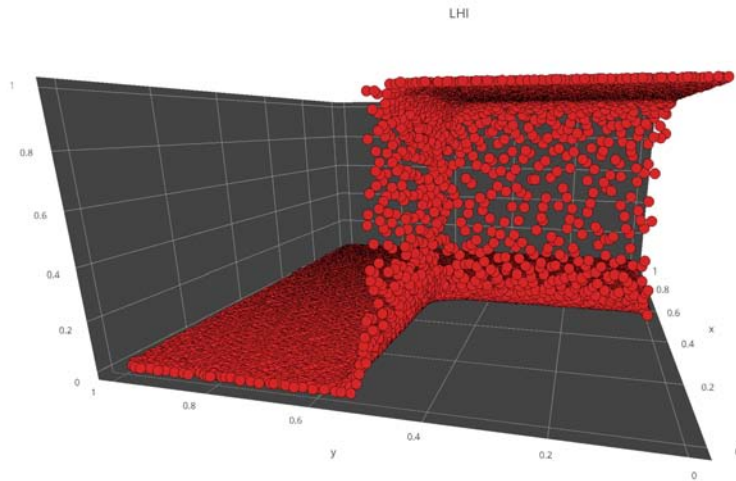
esto es, la norma l_2 del vector de los valores de la función en los nodos interiores. Para 289 nodos, $\beta = 10^{-2}$, $n_v = 50$ para LHI y gaussianas para DQ.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Gaussiana	$c = 50$ $y : 5.48 \times 10^0$ $u : 4.54 \times 10^0$ costo: 15.23 no reportado.	$c = 50$ $y : 3.51 \times 10^0$ $u : 4.68 \times 10^0$ costo: 6.29 tiempo: 8s	$c = 50$ $y : 2.46 \times 10^1$ $u : 8.46 \times 10^1$ costo: 340.09 tiempo: 2s	$c = 50$ $y : 7.41 \times 10^1$ $u : 2.04 \times 10^2$ costo: 2960.14 tiempo: 3s	$c = 50$ $y : 1.80 \times 10^1$ $u : 3.59 \times 10^3$ costo: 64770.83 tiempo: 3s
MQ	no reportado.	$c = 5 \times 10^{-5}$ $y : 3.32 \times 10^0$ $u : 1.63 \times 10^0$ costo: 5.54 tiempo: 8s	$c = 5 \times 10^{-5}$ $y : 3.18 \times 10^0$ $u : 1.58 \times 10^0$ costo: 5.08 tiempo: 2s	$c = 5 \times 10^{-5}$ $y : 3.52 \times 10^0$ $u : 2.19 \times 10^0$ costo: 6.22 tiempo: 3s	$c = 5 \times 10^{-5}$ $y : 3.79 \times 10^0$ $u : 2.54 \times 10^2$ costo: 330.91 tiempo: 3s

Para 289 nodos, $\beta = 10^{-5}$, $n_v = 50$ para LHI.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Gaussiana	$c = 50$ $y : 4.48 \times 10^0$ $u : 3.54 \times 10^2$ costo: 10.67 no reportado.	$c = 50$ $y : 1.98 \times 10^0$ $u : 4.01 \times 10^2$ costo: 2.77 tiempo: 8s	$c = 50$ $y : 1.22 \times 10^1$ $u : 4.49 \times 10^3$ costo: 176.02 tiempo: 2s	$c = 50$ $y : 3.18 \times 10^1$ $u : 1.19 \times 10^5$ costo: 71368.61 tiempo: 3s	$c = 50$ $y : 2.06 \times 10^1$ $u : 2.38 \times 10^4$ costo: 3045.78 tiempo: 3s
MQ	no reportado.	$c = 5 \times 10^{-5}$ $y : 2.52 \times 10^0$ $u : 3.84 \times 10^2$ costo: 3.92 tiempo: 8s	$c = 5 \times 10^{-5}$ $y : 3.23 \times 10^0$ $u : 2.01 \times 10^3$ costo: 25.42 tiempo: 2s	$c = 5 \times 10^{-5}$ $y : 3.61 \times 10^0$ $u : 2.27 \times 10^3$ costo: 30.46 tiempo: 3s	$c = 5 \times 10^{-5}$ $y : 3.80 \times 10^0$ $u : 1.57 \times 10^3$ costo: 15.74 tiempo: 3s

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



Solución obtenida vía LHI-DQ para $\beta = 10^{-5}$, 20mil nodos, 50 en cada sistema local y utilizando multicuadrado para y y para u . Se utilizó una cantidad grande de nodos en estas figuras para mostrar con todo detalle la naturaleza casi discontinua de las soluciones de éste problema.

3.5.3. Problema 3

$$\begin{aligned}
 -\Delta y = u, \quad -\beta \Delta u = y - \hat{y} & \quad \text{en } \Omega \\
 y = g, \quad u = 0 & \quad \text{en } \partial\Omega \\
 \hat{y} = \begin{cases} 1 & \text{en } [0, \frac{1}{2}]^2 \cap \Omega \\ 0 & \text{en otro caso} \end{cases} & \quad (3.16) \\
 g = 0 &
 \end{aligned}$$

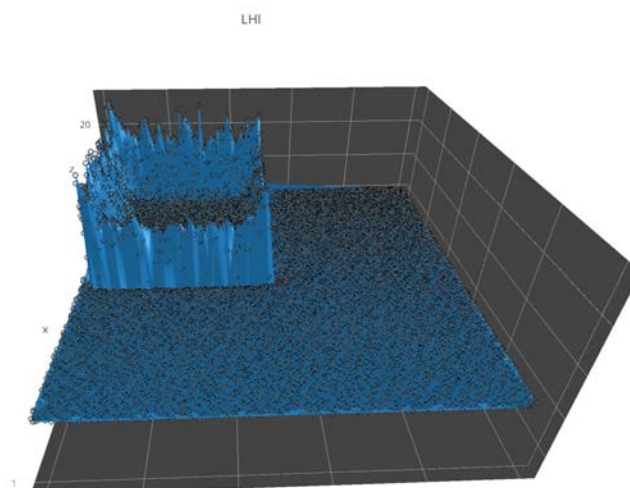
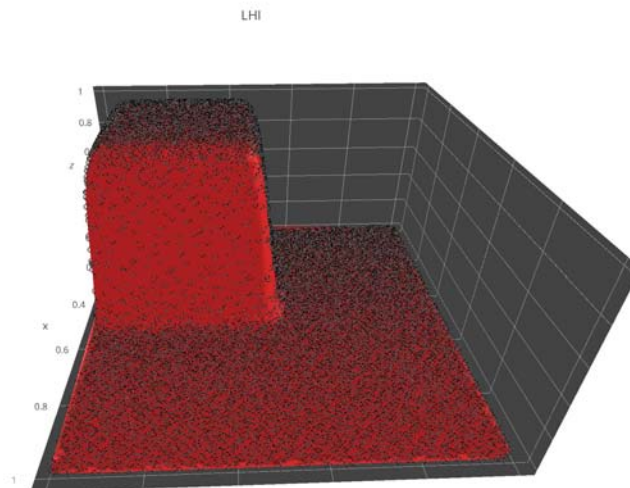
En este problema se reportan los mismos valores que en el problema 2.
 Para 289 nodos, $\beta = 10^{-2}$, $n_v = 50$ para LHI y gaussianas para DQ.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Gaussiana	$c = 50$ $y : 7.50 \times 10^0$ $u : 1.87 \times 10^1$ costo: 29.91 no reportado.	$c = 50$ $y : 7.49 \times 10^0$ $u : 1.70 \times 10^1$ costo: 29.52 tiempo: 8s	$c = 50$ $y : 8.68 \times 10^0$ $u : 1.33 \times 10^1$ costo: 38.61 tiempo: 2s	$c = 50$ $y : 8.66 \times 10^0$ $u : 2.46 \times 10^1$ costo: 40.55 tiempo: 3s	$c = 50$ $y : 8.53 \times 10^0$ $u : 3.79 \times 10^1$ costo: 37.98 tiempo: 3s
MQ	no reportado.	$c = 5 \times 10^{-5}$ $y : 7.74 \times 10^0$ $u : 9.91 \times 10^0$ costo: 30.49 tiempo: 8s	$c = 5 \times 10^{-5}$ $y : 7.68 \times 10^0$ $u : 9.01 \times 10^0$ costo: 29.90 tiempo: 2s	$c = 5 \times 10^{-5}$ $y : 8.54 \times 10^0$ $u : 1.14 \times 10^1$ costo: 37.15 tiempo: 3s	$c = 5 \times 10^{-5}$ $y : 8.54 \times 10^0$ $u : 1.40 \times 10^{-2}$ costo: 36.49 tiempo: 3s

Para 289 nodos, $\beta = 10^{-5}$, $n_v = 50$ para LHI.

FBR	Pearson (K.Ac.)	Kansa Ac.	Kansa Des.	LHI-LHI	LHI-DQ
Gaussiana	$c = 50$ $y : 2.40 \times 10^0$ $u : 6.06 \times 10^2$ costo: 4.71 no reportado.	$c = 50$ $y : 3.72 \times 10^0$ $u : 6.41 \times 10^2$ costo: 9.00 tiempo: 8s	$c = 50$ $y : 5.40 \times 10^0$ $u : 2.15 \times 10^3$ costo: 37.80 tiempo: 2s	$c = 50$ $y : 7.92 \times 10^0$ $u : 2.55 \times 10^4$ costo: 3284.13 tiempo: 3s	$c = 50$ $y : 2.48 \times 10^1$ $u : 1.71 \times 10^4$ costo: 1784.11 tiempo: 3s
MQ	no reportado.	$c = 5 \times 10^{-5}$ $y : 4.16 \times 10^0$ $u : 7.65 \times 10^2$ costo: 11.61 tiempo: 8s	$c = 5 \times 10^{-5}$ $y : 7.56 \times 10^0$ $u : 8.57 \times 10^3$ costo: 396.15 tiempo: 2s	$c = 5 \times 10^{-5}$ $y : 8.36 \times 10^0$ $u : 1.09 \times 10^4$ costo: 629.75 tiempo: 3s	$c = 5 \times 10^{-5}$ $y : 8.36 \times 10^0$ $u : 1.53 \times 10^1$ costo: 35.01 tiempo: 3s

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



Solución obtenida vía LHI-DQ para $\beta = 10^{-5}$, 20mil nodos, 50 en cada sistema local y utilizando multicuádrico para y y para u , tal como en las figuras del problema anterior. Estos dos últimos problemas muestran que LHI es capaz de reconstruir correctamente la solución aún cuando β se toma muy chico y la solución exacta, aunque suave, se apega más a las discontinuidades de los datos.

3.5.4. Problema 4

$$\begin{aligned}
 -\Delta y &= u, & -\beta \Delta u &= y - \hat{y} & \text{en } \Omega \\
 y &= g, & u &= 0 & \text{en } \partial\Omega \\
 \hat{y} &= \sin 16\pi x_1 \sin 16\pi x_2 \\
 g &= 0
 \end{aligned} \tag{3.17}$$

Éste es un problema análogo al problema 1 que se incluyó en éste trabajo con la intención de exhibir un caso de prueba donde la calidad de la solución demande un número elevado de nodos.

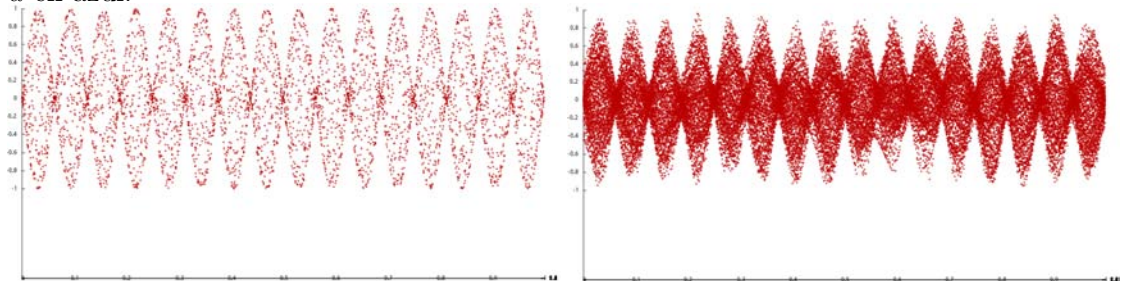
Se reportan los errores relativos de y y u respecto a la solución analítica, dada por

$$\begin{aligned}
 y_\beta(x_1, x_2) &= \frac{1}{1 + 4\beta(16\pi)^4} \sin 16\pi x_1 \sin 16\pi x_2 \\
 u_\beta(x_1, x_2) &= \frac{2(16\pi)^2}{1 + 4\beta(16\pi)^4} \sin 16\pi x_1 \sin 16\pi x_2
 \end{aligned}$$

Los resultados siguientes se obtuvieron con 5mil nodos para LHI, 50 en los sistemas locales y con mil nodos para Kansa. Se empleó multicuádrico y gaussiana para LHI-DQ y multicuádrico para Kansa.

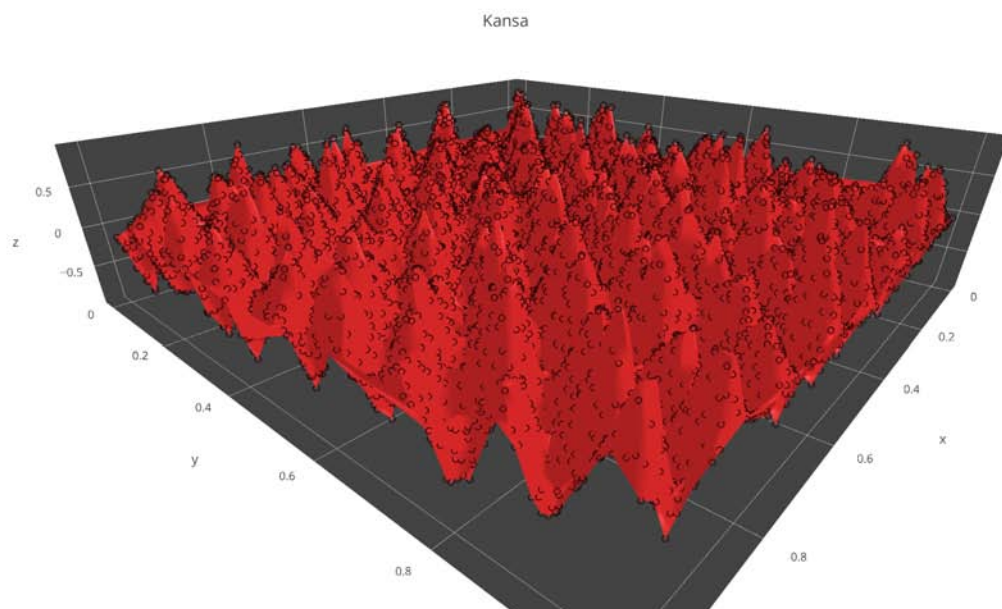
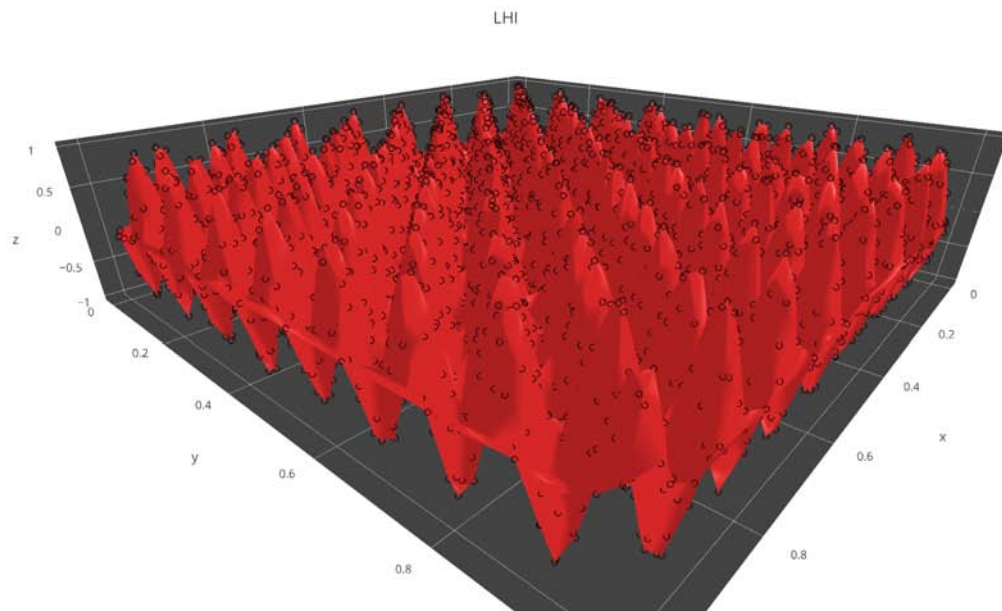
β	LHI-DQ	Kansa
10^{-8}	$y : 6.48 \times 10^{-4}$ $u : 9.09 \times 10^{-3}$ tiempo: 59s	$y : 5.41 \times 10^{-1}$ $u : 8.95 \times 10^0$ tiempo: 6min12s
10^{-10}	$y : 6.85 \times 10^{-6}$ $u : 1.71 \times 10^{-3}$ tiempo: 59s	$y : 5.31 \times 10^{-1}$ $u : 8.89 \times 10^0$ tiempo: 6min12s

Las siguientes imágenes muestran la solución correspondiente a $\beta = 10^{-8}$, y en rojo y u en azul.

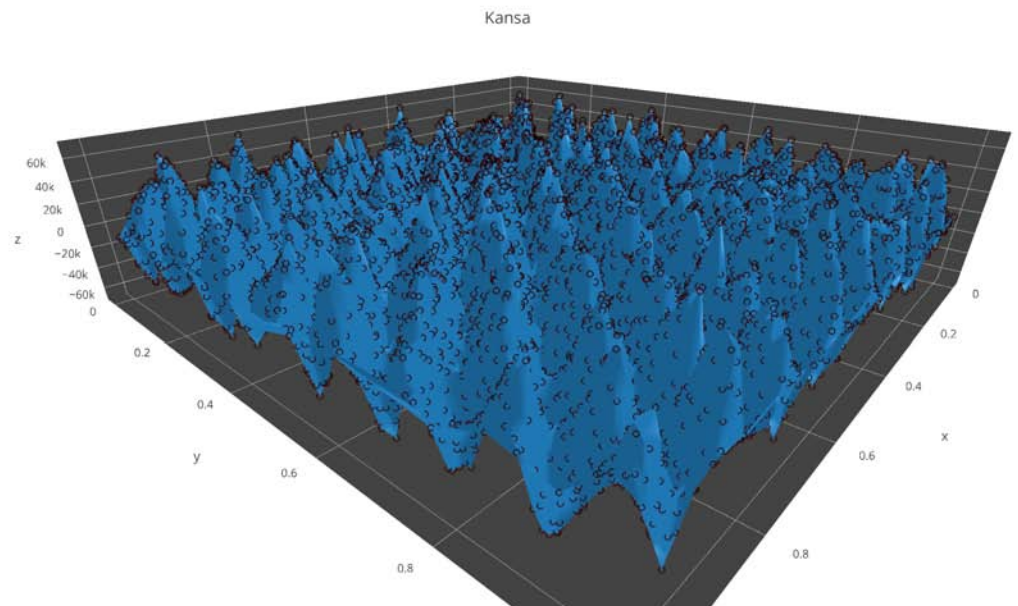
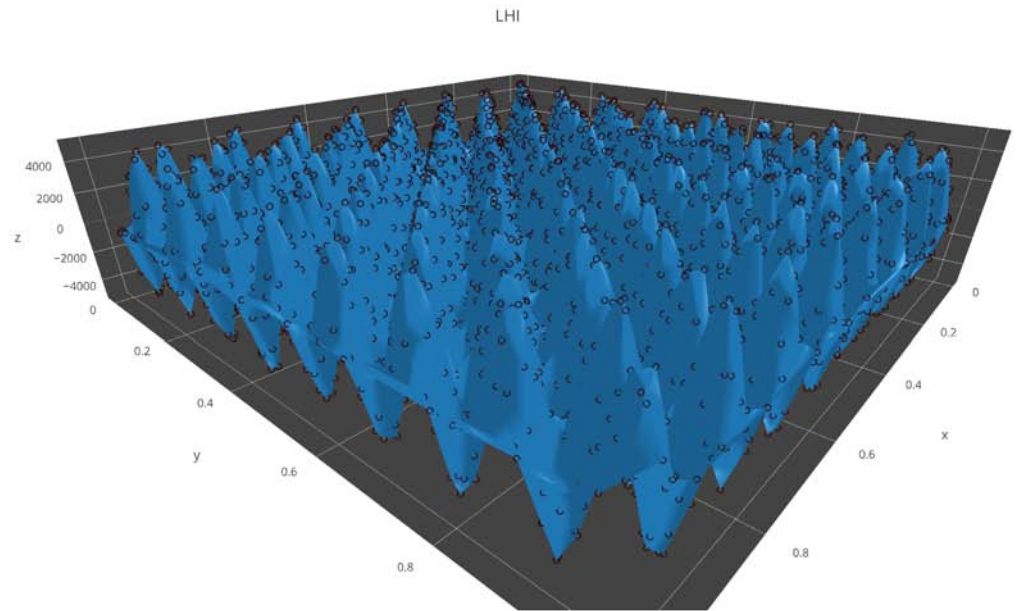


Vista de perfil del estado.

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



3.5 Casos de prueba para problemas de tipo Poisson



3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE

Como se puede observar en las imágenes anteriores, la naturaleza oscilatoria de la solución exacta se manifiesta en que la capacidad de reconstrucción que se obtiene a través de Kansa con 1000 nodos no es suficiente para obtener la solución que se espera. Si bien lo que se obtiene es razonablemente parecido a la solución exacta (al menos visualmente), presenta irregularidades y oscilaciones más pronunciadas comparada con la solución exacta. En principio estos inconvenientes deberían ser subsanados al incrementar el número de nodos. Ésto se ve imposibilitado tanto por el incremento del número de condición de la matriz global de Kansa como por el rápido crecimiento del costo computacional que conlleva: vale la pena mencionar que se realizó la prueba para 5mil nodos y luego de haber transcurrido 32min y aún no terminaba su ejecución.

De ésta manera se hace patente la necesidad de utilizar LHI que, gracias a su formulación local, es capaz de manejar cantidades de nodos mucho más grandes, con una demanda de recursos computacionales considerablemente menor, mejores cifras de error y la capacidad de reconstruir de manera fidedigna la solución.

Cabe resaltar que la frecuencia de la solución, que en éste problema es de 16, se escogió relativamente pequeña para posibilitar la comparación visual de las soluciones, y que es ésta la que determina esencialmente la cantidad de nodos necesarios para reconstruir la solución correctamente. Ésto significa que para frecuencias más altas la solución obtenida con Kansa será aún más pobre, mientras que LHI seguirá obteniéndola gracias a su capacidad de manejar cantidades de nodos muy superiores en un tiempo mucho más razonable y por lo tanto será la única opción viable de entre éstas dos, siendo tal vez otro candidato el esquema de descomposición de dominios, como se presenta en (4), y que no se analiza en el presente trabajo.

Asimismo es importante hacer notar que el número de condicionamiento de las matrices locales de LHI es de orden 10^5 , mientras que el de la matriz global de Kansa se encuentra alrededor de 10^{22} . Además, se espera que éste se incremente junto con el número de nodos, mientras que el de LHI, que mantiene el tamaño de los sistemas locales, no se aleja más de dos órdenes de magnitud de la cifra mencionada al aumentar el número de nodos. En cuanto al condicionamiento de la matriz del sistema global de LHI, la forma en que se implementó el código impide su calculo exacto, pero los experimentos numéricos que se llevaron a cabo a lo largo de éste trabajo, así como la baja densidad de elementos distintos de cero, que para 5mil nodos y 50 en sistemas locales es de aproximadamente 0.5% sugieren que éste número de condición se encuentra en un rango aceptable.

Así, podemos concluir que, para problemas semejantes a éste, LHI es una opción mucho más sólida que la colocación usual de Kansa.

3.6. Casos de prueba para problemas de tipo Convección-Difusión.

Los siguientes problemas de convección-difusión se pueden encontrar en (7) y (6), respectivamente, aunque abordados a través de elemento finito.

3.6.1. Problema 5

$$\begin{aligned}
 (-\epsilon \Delta + \omega \cdot \nabla)y = u, \quad \beta(-\epsilon \Delta - \omega \cdot \nabla)u = y - \hat{y} & \quad \text{en } \Omega \\
 y = g, \quad u = 0 & \quad \text{en } \partial\Omega \\
 \hat{y} = \begin{cases} (2x_1 - 1)^2(2x_2 - 1)^2 & \text{en } [0, \frac{1}{2}]^2 \cap \Omega \\ 0 & \text{en otro caso} \end{cases} \\
 g = \begin{cases} (2x_1 - 1)^2(2x_2 - 1)^2 & \text{en } [0, \frac{1}{2}]^2 \cap \partial\Omega \\ 0 & \text{en otro caso} \end{cases} \\
 \omega = (\cos \theta, \sin \theta), \quad \text{con } \theta = 2.4 \\
 \epsilon = \frac{1}{200}
 \end{aligned} \tag{3.18}$$

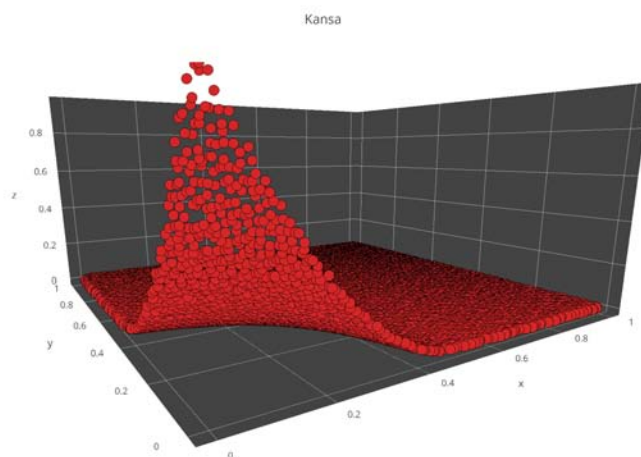
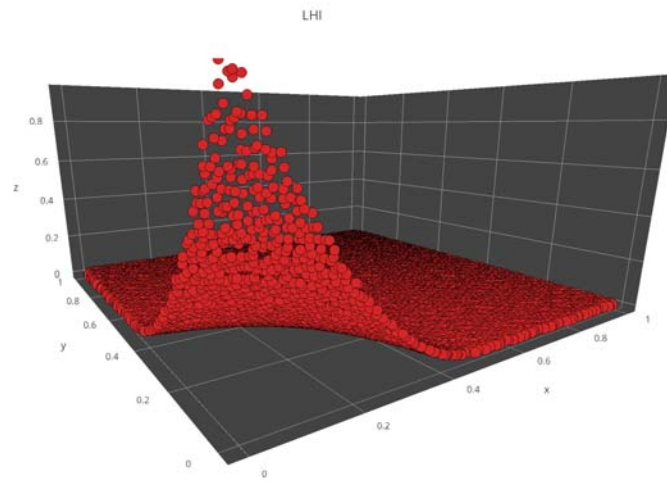
Este problema carece de solución analítica y por lo tanto reportaremos los mismos valores que en los problemas 2 y 3, a saber, $\|y - \hat{y}\|_{l_2}$ para y , $\|u\|_{l_2}$ para u y el funcional de costo, $\frac{1}{2}\|y - \hat{y}\|_{l_2}^2 + \frac{\beta}{2}\|u\|_{l_2}^2$, pero tomaremos la norma como sigue:

$$\|u\|_{l_2}^2 = \frac{1}{n} \sum_{k=n_b+1}^n |u(x_k)|^2$$

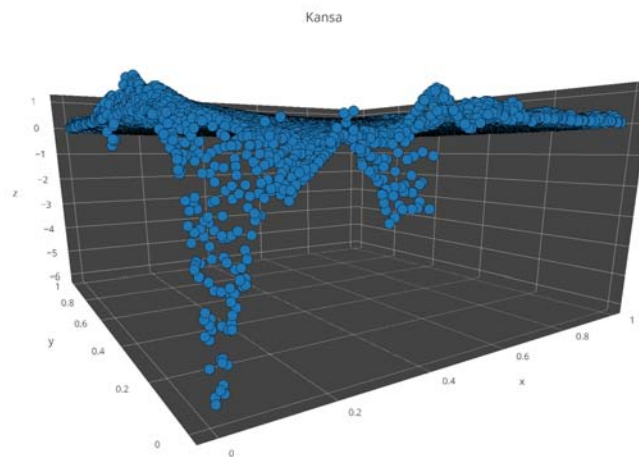
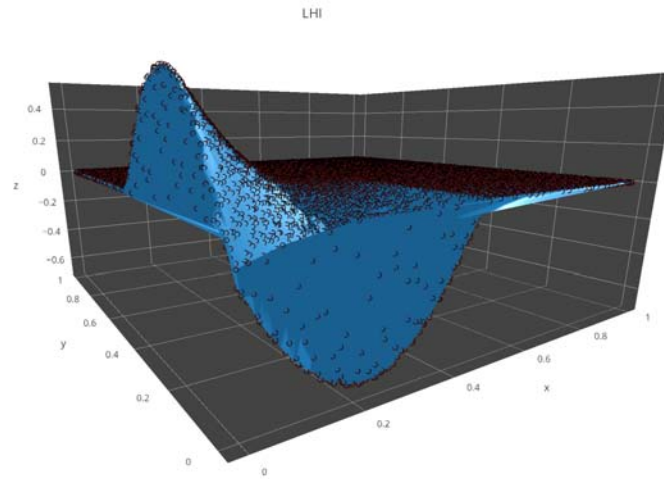
La siguiente tabla se construyó con 5000 nodos y 50 en cada sistema local para LHI-DQ -con MQ para DQ- y con 1000 para Kansa. En se utilizó multicuádrico con parámetro de forma $c = 10^{-6}$.

β	LHI-DQ	Kansa
1	$y : 5.52 \times 10^{-4}$ $u : 1.38 \times 10^{-1}$ costo: 1.62×10^{-7} tiempo: 1min12s	$y : 5.17 \times 10^{-4}$ $u : 2.71 \times 10^{-1}$ costo: 1.70×10^{-7} tiempo: 5min50s
10^{-2}	$y : 2.58 \times 10^{-5}$ $u : 1.33 \times 10^{-1}$ costo: 4.22×10^{-10} tiempo: 1min08s	$y : 4.06 \times 10^{-4}$ $u : 3.78 \times 10^{-1}$ costo: 8.32×10^{-8} tiempo: 5min50s
10^{-5}	$y : 4.64 \times 10^{-8}$ $u : 1.33 \times 10^{-1}$ costo: 8.98×10^{-14} tiempo: 1min09s	$y : 4.53 \times 10^{-4}$ $u : 3.43 \times 10^{-1}$ costo: 1.02×10^{-7} tiempo: 5min48s

3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



3.6 Casos de prueba para problemas de tipo Convección-Difusión.



Destacamos que, si bien la solución para y obtenida con Kansa es en todos los casos razonablemente buena, el control que arroja no corresponde al que da LHI. Éste último coincide con el que se obtiene con elemento finito (7) y por lo tanto concluimos una vez más que LHI tiene una ventaja significativa por sobre la colocación de Kansa.

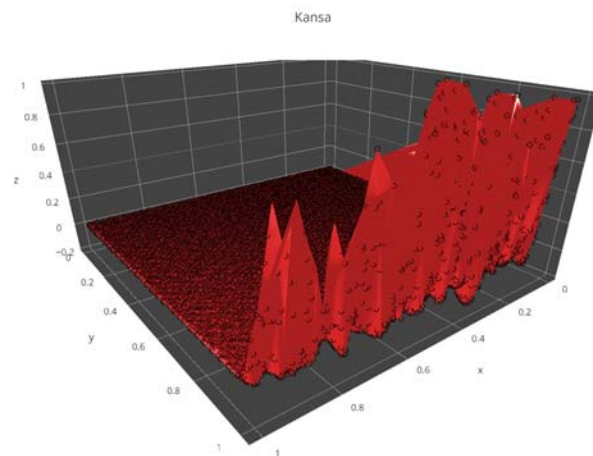
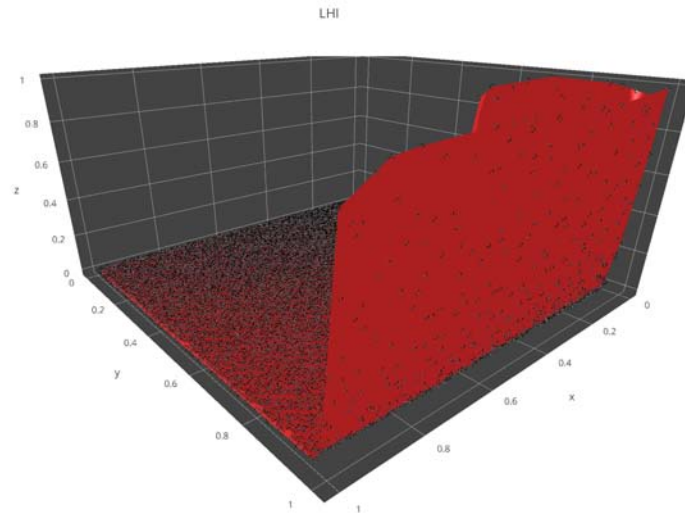
3.6.2. Problema 6

$$\begin{aligned}
 (-\epsilon \Delta + \omega \cdot \nabla)y &= u, & \beta(-\epsilon \Delta - \omega \cdot \nabla)u &= y - \hat{y} & \text{en } \Omega \\
 y &= g, & u &= 0 & \text{en } \partial\Omega \\
 \hat{y} &= 0 \\
 g &= \begin{cases} 1 & \text{en } \{0\} \times [\frac{1}{2}, 1] \cup [0, 1] \times \{1\} \\ 0 & \text{en otro caso} \end{cases} & (3.19) \\
 \omega &= (\cos \theta, \sin \theta), & \text{con } \theta &= \frac{\pi}{6} \\
 \epsilon &= \frac{1}{200}
 \end{aligned}$$

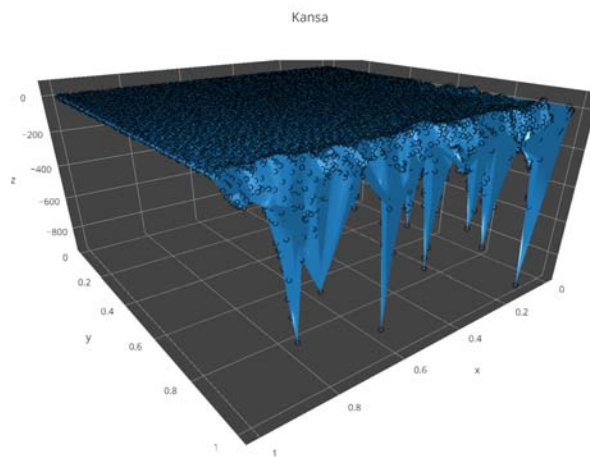
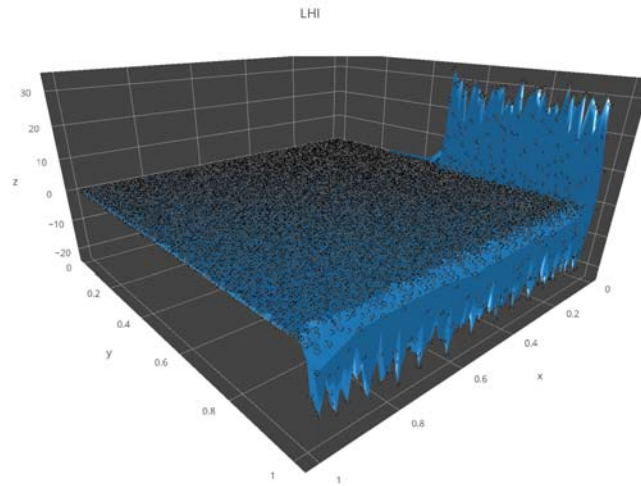
En este problema se reportan exactamente los mismos valores que en el anterior. Se utilizaron también los mismos parámetros en todos los casos.

β	LHI-DQ	Kansa
1	$y : 1.18 \times 10^{-1}$ $u : 4.12 \times 10^0$ costo: 7.07×10^{-3} tiempo: 1min07s	$y : 9.06 \times 10^{-2}$ $u : 3.02 \times 10^{+1}$ costo: 4.56×10^{-3} tiempo: 5min50s
10^{-1}	$y : 7.39 \times 10^{-2}$ $u : 5.20 \times 10^{+1}$ costo: 2.73×10^{-3} tiempo: 1min08s	$y : 8.65 \times 10^{-2}$ $u : 4.39 \times 10^{+1}$ costo: 3.84×10^{-3} tiempo: 5min50s
10^{-2}	$y : 3.75 \times 10^{-2}$ $u : 4.17 \times 10^0$ costo: 7.09×10^{-4} tiempo: 1min07s	$y : 9.00 \times 10^{-2}$ $u : 5.93 \times 10^{+1}$ costo: 4.07×10^{-3} tiempo: 5min48s

3.6 Casos de prueba para problemas de tipo Convección-Difusión.



3. SOLUCIÓN DEL PROBLEMA DE CONTROL ÓPTIMO DE TIPO POISSON Y CONVECCIÓN-DIFUSIÓN MEDIANTE INTERPOLACIÓN LOCAL DE HERMITE



Una vez más observamos que para obtener una solución más cercana a \hat{y} con colocación de Kansa requeriríamos un número mayor de nodos, aumentando al mismo tiempo y rápidamente la cantidad de recursos computacionales; mientras que LHI puede llegar sin problema a esas cantidades de nodos y producir una solución más próxima al estado objetivo \hat{y} .

Conclusiones

Éste trabajo concluye con un breve recuento de los avances en el área que motivaron el enfoque con el que se elaboró. Por una parte se tiene la formulación de la interpolación local de Hermite en la vanguardia de los esquemas de interpolación y solución numérica de *EDPs*, fruto del primer trabajo al respecto (9). Al mismo tiempo se ha despertado más interés en la solución numérica de problemas de control por autores como (1, 5, 7), tanto en el caso de Poisson como en el de Convección-Difusión. y, en uno de sus más recientes alcances, han empezado a considerar esquemas de solución basados en FBRs; en él se utilizó colocación de Kansa y concluye que la solución mediante FBRs es una alternativa sólida a elemento finito.

Con éste panorama en mente, el objetivo que se fijó para éste trabajo fue el de avanzar en ambas direcciones, resaltando las virtudes de LHI en términos computacionales, y simultáneamente realizar un estudio más extensivo del caso de control.

De esta manera decidimos ampliar lo que se conocía previamente en el caso de problemas de control de Poisson a problemas de convección-difusión, y al mismo tiempo estudiar el comportamiento de Kansa y LHI en ambos.

Como se muestra en algunos ejemplos de el capítulo de control, existen datos, tanto en Poisson como en Convección-Difusión, que requieren una cantidad elevada de nodos y que eventualmente excede el máximo manejable por la colocación asimétrica de Kansa. Así, en éste trabajo no solo se exhiben ejemplos donde el esquema de solución con FBRs con el que se contaba anteriormente deja de ser una opción viable, sino que también se propone, implementa, compara y valida una alternativa, LHI, con la que no solamente se obtiene la solución correctamente, sino que también mejora en otros aspectos como la cantidad de recursos computacionales que se requieren.

Por lo pronto concluimos aquí el presente trabajo, esperando que éste siembre un antecedente de la posibilidad de obtener resultados a la altura de otros esquemas de solución numérica en el ámbito de problemas de control donde hasta hace poco ni siquiera se consideraban radiales, y con ello despierte interés en continuar explorando a mayor profundidad lo que se mostró aquí, o bien el de adaptar esquemas locales, como LHI en otras áreas donde no se le ha considerado.

Nociones básicas de análisis funcional

En éste apéndice prepararemos el marco teórico adecuado para la exposición del enfoque Hilbertiano abstracto en el problema de interpolación de una función dado un conjunto de datos conocido del capítulo 1.

Para ello haremos un breve repaso de las nociones de espacio lineal, métrico y normado, haciendo énfasis en las propiedades de éstos que son relevantes para la teoría de espacios de Hilbert.

A.1. Espacios lineales

Comencemos con un rápido repaso de la teoría de espacios lineales.

Definición A.1.1. Sea \mathbb{K} un campo y X un conjunto. Diremos que dos operaciones $+: X \times X \rightarrow X$, $\cdot: \mathbb{K} \times X \rightarrow X$ –a las que llamaremos suma y producto por escalares, respectivamente– definen una estructura de \mathbb{K} -espacio vectorial sobre X si se verifica que:

1. $(X, +)$ es un grupo conmutativo. Denotamos a su elemento neutro con 0.

2. Para cualesquiera $f, g \in X$ y $\lambda, \mu \in \mathbb{K}$:

I $\lambda(f + g) = \lambda f + \lambda g$.

II $(\lambda + \mu)f = \lambda f + \mu f$.

III $(\lambda\mu)f = \lambda(\mu f)$.

IV $1f = f$.

También se suele decir que X es un espacio vectorial sobre \mathbb{K} , un espacio lineal o espacio \mathbb{K} -lineal. En lo que sigue, \mathbb{K} será o bien el campo de los números reales \mathbb{R} o bien el de los números complejos \mathbb{C} y X un espacio lineal.

Definición A.1.2. Diremos que un subconjunto $V \subset X$ es un subespacio lineal si es en sí mismo un espacio lineal sobre \mathbb{K} con las mismas operaciones de X . Es sencillo mostrar que las condiciones:

1. $0 \in V$
2. $V + V = \{f_1 + f_2 : f_1, f_2 \in V\} \subset V$
3. $\lambda V = \{\lambda f : f \in V\} \subset V$ para todo $\lambda \in \mathbb{K}$

son necesarias y suficientes para tal efecto.

Definición A.1.3. Se dice que un subconjunto $S \subset X$ es linealmente independiente si para cualquier colección $f_1, \dots, f_n \in S$ la ecuación

$$\lambda_1 f_1 + \dots + \lambda_n f_n = 0$$

se satisface únicamente cuando $\lambda_1 = \dots = \lambda_n = 0$. El miembro izquierdo de la misma se conoce como combinación lineal de f_1, \dots, f_n . En particular, $0 \notin S$.

Definición A.1.4. Dado un subconjunto $S \subset X$ definimos el espacio generado por S como el subespacio lineal formado por combinaciones lineales de S :

$$\text{span } S = \left\{ \sum_{\ell=1}^k \lambda_\ell f_\ell : f_\ell \in S, \lambda_\ell \in \mathbb{K}, \text{ con } 1 \leq \ell \leq k \in \mathbb{N} \right\}$$

Es inmediato comprobar que $\text{span } S$ es el subespacio lineal más chico –en el sentido de la contención– que contiene a S . Utilizaremos la notación $\text{span}_{\mathbb{F}} S$ para indicar que los escalares pertenecen a otro campo \mathbb{F} cuando sea relevante. Diremos además que S es un subconjunto generador cuando $\text{span } S = X$.

Definición A.1.5. Una base algebraica B de X es un subconjunto generador y linealmente independiente. Esta noción también se conoce como base de Hamel.

Lema A.1.6. Si $S = \{s_1, \dots, s_n\}$ es un subconjunto linealmente independiente y $R = \{r_1, \dots, r_m\}$ un subconjunto generador, entonces $n \leq m$.

Demostración. Como R es generador, s_1 se escribe como combinación lineal de elementos de R y al menos uno de los escalares involucrados es distinto de cero, pues $s_1 \neq 0$. Sin perder generalidad podemos suponer que es el correspondiente a r_1 , de manera que $R_1 = \{s_1, r_2, \dots, r_m\}$ es también un subconjunto generador de tamaño m . Repetimos este proceso, como consecuencia de la independencia lineal de S en cada paso habremos quitado un elemento de R y podremos continuarlo hasta agotar S . Así, R_n es un conjunto de tamaño m y $S \subset R_n$. \square

Definición A.1.7. La dimensión de X se define como sigue:

- I Si existe una base finita B , definimos la dimensión de X como el número de elementos de B , esto es, $\dim(X) = \#B$.
- II Si existen subconjuntos linealmente independientes de tamaño arbitrariamente grande, diremos que $\dim(X) = \infty$.

Nota A.1.8. El lema A.1.6 nos permite mostrar que la dimensión está bien definida:

1. Si B y \tilde{B} son dos bases finitas de X , por el lema se tiene que $\#B \leq \#\tilde{B}$ pues B es linealmente independiente y \tilde{B} generador. Del mismo modo se obtiene que $\#\tilde{B} \leq \#B$ y por lo tanto la dimensión no depende de la base.
2. Además, I y II no pueden ocurrir simultáneamente. Si B es una base finita de X , por II existe un subconjunto linealmente independiente S con $\#S = \#B + 1$. Por otro lado, puesto que B es -en particular- generador, el lema asegura que $\#S \leq \#B$, una contradicción.

Definición A.1.9. Dado un subespacio lineal $V \subset X$ consideramos la relación de equivalencia definida por $f \sim g$ si y solo si $f - g \in V$. Para cada $f \in X$ escribiremos a su clase de equivalencia como $f + V = \{f + g : g \in V\}$. Entonces el conjunto de clases de equivalencia (que denotaremos con X/V en lugar de X/\sim) se denomina

el conjunto cociente de X con V y constituye un espacio lineal con las operaciones $(f + V) + (g + V) = (f + g) + V$ y $\lambda(f + V) = \lambda f + V$.

Definición A.1.10. Una función $T : V \rightarrow W$ entre dos espacios \mathbb{K} -lineales se dice transformación lineal si para cualesquiera $f, g \in V$ y $\lambda \in \mathbb{K}$ se tiene que:

1. $T(f + g) = Tf + Tg$

2. $T(\lambda f) = \lambda Tf$

Nótese que escribimos $Tf = T(f)$. El espacio nulo de T se define como

$$\mathcal{N}(T) = \{f \in V : Tf = 0\}$$

y se ve inmediatamente que $\mathcal{N}(T)$ es un subespacio lineal de V y que T es inyectiva si y solo si $\mathcal{N}(T) = \{0\}$. Si además T es biyectiva diremos que es un isomorfismo lineal.

Concluimos ésta sección con la definición de una noción de carácter particularmente geométrico que como veremos más adelante, es parte importante de la estructura de ciertas clases de espacios lineales.

Definición A.1.11. Diremos que un subconjunto $S \subset X$ es convexo si dados $f, g \in X$ también se tiene que $\alpha f + (1 - \alpha)g \in S$ para $\alpha \in [0, 1]$.

A.2. Espacios métricos, normados y con producto interior.

En ésta sección presentaremos un breve resumen de las nociones básicas de distintos tipos de espacios. Mostraremos que se puede establecer una jerarquía de manera natural entre ellos y de esta forma refinar algunos conceptos a situaciones más cercanas a nuestra experiencia con, por ejemplo, espacios euclidianos.

Resumen A.2.1. (Sobre espacios métricos)

1. Sea Ω un conjunto. Decimos que $d : \Omega \times \Omega \rightarrow \mathbb{R}$ es una métrica (o distancia) si para cualesquiera $\omega_1, \omega_2, \omega_3 \in \Omega$ se tiene que:

- I $d(\omega_1, \omega_2) \geq 0$.

- II $d(\omega_1, \omega_2) = d(\omega_2, \omega_1)$.

- III $d(\omega_1, \omega_2) \leq d(\omega_1, \omega_3) + d(\omega_3, \omega_2)$.

- IV $d(\omega_1, \omega_2) = 0$ si y solo si $\omega_1 = \omega_2$.

Si en lugar de la condición IV solo sucede que $d(\omega, \omega) = 0$ para todo $\omega \in \Omega$, diremos que d es una pseudo-métrica.

2. Una función Ψ entre dos espacios pseudo-métricos (Ω, d) y (\mathcal{U}, q) recibe el nombre de isometría (parcial) si para cualesquiera $\omega_1, \omega_2 \in \Omega$ sucede que $d(\omega_1, \omega_2) = q(\Psi(\omega_1), \Psi(\omega_2))$. Adicionalmente diremos que Ψ es

- un encaje (isométrico) si es inyectiva. Se sigue inmediatamente que éste es el caso cuando Ω es un espacio métrico.
- una proyección (isométrica) si es sobreyectiva.
- una isometría global o isomorfismo isométrico si es una biyección. En tal caso diremos que Ω y \mathcal{U} son isométricos, o isométricamente isomorfos.

3. Para cada $\omega \in \Omega$ y $\rho > 0$, el conjunto $B_\rho(\omega) = \{\tilde{\omega} \in \Omega : d(\tilde{\omega}, \omega) < \rho\}$ se conoce como la bola (abierta) de radio ρ alrededor de ω . Si para $\mathcal{S} \subset \Omega$ existen $\omega \in \Omega$ y $\rho > 0$ tales que $\mathcal{S} \subset B_\rho(\omega)$, diremos que \mathcal{S} es acotado.

4. Una pseudo-métrica induce una estructura topológica tomando como base a la colección $\{B_\rho(\omega) : \omega \in \Omega, \rho > 0\}$, de forma que un conjunto es abierto si se puede escribir como unión de bolas abiertas.

Si además es métrica, el espacio topológico que se obtiene resulta ser de Hausdorff. A menos de que se indique lo contrario, las nociones topológicas que mencionemos —como compacidad y conexidad de sus subconjuntos o continuidad de funciones— se referirán a la topología inducida.

5. Se dice que una sucesión $\{\omega_k\}_{k=1}^\infty$ de un espacio pseudo-métrico Ω es
- fundamental (o de Cauchy) si para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para cualesquiera $j, k \geq N$ se tiene que $d(\omega_j, \omega_k) < \varepsilon$.
 - convergente, a $\omega \in \Omega$, si para todo $\varepsilon > 0$ existe $N \in \mathbb{N}$ tal que para cualquier $j \geq N$ se tiene que $d(\omega_j, \omega) < \varepsilon$.

Y claramente toda sucesión convergente es fundamental. Cabe destacar que si Ω sólo es pseudo-métrico el límite puede no ser único.

6. Se puede mostrar que subconjunto \mathcal{S} de un espacio métrico Ω es compacto si y solo si es secuencialmente compacto, es decir, si de toda sucesión $\{\omega_k\}_{k=1}^\infty \subset \mathcal{S}$ se puede extraer una subsucesión convergente. En particular, un compacto es cerrado y acotado pero el recíproco no siempre es cierto.

7. Un subconjunto $\mathcal{S} \subset \Omega$ resulta cerrado —i.e. $\Omega \setminus \mathcal{S}$ abierto— si y solo si toda sucesión fundamental de \mathcal{S} converge y su límite también pertenece a \mathcal{S} .

Denotamos con $\bar{\mathcal{S}}$ a la cerradura de \mathcal{S} i.e. el conjunto cerrado más chico que contiene a \mathcal{S} . Diremos que un subconjunto $\mathcal{S} \subset \Omega$ es denso en Ω si $\bar{\mathcal{S}} = \Omega$ y que Ω es separable si existe un subconjunto denso y numerable.

8. De forma similar, un espacio se dice completo si toda sucesión fundamental converge. Siempre es posible construir un espacio completo $(\bar{\Omega}, \bar{d})$ en el que (Ω, d) se encaje de manera isométrica. El espacio más chico —en el sentido de encajes

isométricos— que posee ésta propiedad recibe el nombre de completación de Ω y evidentemente es único salvo isomorfismos.

Es importante hacer notar que los espacios métricos no necesariamente están equipados con varias de las estructuras que estamos acostumbrados a ver en otro tipo de situaciones. Por ello nos ocuparemos ahora de hablar de espacios lineales que poseen una estructura adicional, compatible con su estructura de espacio vectorial, a través de la cual es posible entenderlos como espacios métricos.

Resumen A.2.2. (Sobre espacios lineales normados)

1. Decimos que $\|\cdot\| : X \rightarrow \mathbb{R}$ define una norma en el espacio lineal X si para cualesquiera $f, g \in X$ y $\lambda \in \mathbb{K}$ se verifica que:

- I $\|f\| \geq 0$.

- II $\|\lambda f\| = |\lambda| \|f\|$.

- III $\|f + g\| \leq \|f\| + \|g\|$.

- IV $\|f\| = 0$ si y solo si $f = 0$.

Si la condición IV no se satisface, llamaremos a $\|\cdot\|$ semi-norma.

Comunmente se escribe $\|f\|_X$ cuando hay más de una norma en discusión.

2. Una (semi-)norma induce una (pseudo-)métrica en X si se define $d(f, g) = \|f - g\|$. Ésta es homogénea e invariante por traslaciones, es decir, para todos $f, g, h \in X$:

- $d(\lambda f, \lambda g) = |\lambda| d(f, g)$.

- $d(f - h, g - h) = d(f, g)$.

respectivamente. Con ella las operaciones de X resultan continuas.

Ésto pone a disposición todas las nociones topológicas y métricas que mencionamos anteriormente. En particular vale la pena resaltar que:

- La completación a un espacio completo \overline{X} se puede hacer de forma que el encaje isométrico $X \rightarrow \overline{X}$ sea lineal.

- El espacio es separable si y solo si existe un subconjunto numerable $S \subset X$ tal que $\overline{\text{span } S} = X$.

3. Diremos que un espacio normado X es de Banach si es completo con ésta métrica. Un subespacio lineal $V \subset X$ se denomina subespacio (de Banach) si es en sí mismo un espacio de Banach con la misma norma que X , es decir, si V es cerrado en X .

4. Si V es un subespacio lineal de X definimos en X/V una semi-norma dada por

$$\|f\|_{X/V} = \inf_{g \in V} \|f + g\|_X$$

Ésta solo será una norma cuando $\|\cdot\|_X$ lo sea y V sea cerrado.

En particular, si $\|\cdot\|$ es una semi-norma, entonces $\mathcal{N}(\|\cdot\|)$ es un subespacio cerrado y a través del cociente se obtiene un espacio normado.

5. Diremos que una aplicación $T : V \rightarrow W$ entre dos espacios semi-normados es acotada si existe $\alpha > 0$ tal que $\|Tf\|_W \leq \alpha\|f\|_V$ para todo $f \in V$. Es muy importante mencionar que ésta condición es equivalente a la continuidad de T . Denotaremos por $\mathcal{B}(V, W)$ al conjunto de transformaciones lineales y acotadas de V en W . En él se define una semi-norma a través de

$$\|T\|_{\mathcal{B}(V, W)} = \inf\{\alpha : \|Tf\| \leq \alpha\|f\|, f \in V\} = \sup_{\|f\| > 0} \frac{\|Tf\|}{\|f\|} = \sup_{\|f\|=1} \|Tf\|$$

y será norma si $\|\cdot\|_W$ lo es. Asimismo si W es de Banach, también $\mathcal{B}(V, W)$.

6. Un espacio lineal normado es el espacio natural para considerar series, dada una colección $\{f_k\}_{k=1}^{\infty} \subset X$ diremos que

- $\sum_{\ell=1}^{\infty} f_{\ell}$ converge si la sucesión de sumas parciales $s_k = \sum_{\ell=1}^k f_{\ell}$ lo hace.
- $\sum_{\ell=1}^{\infty} f_{\ell}$ converge absolutamente si $\sum_{\ell=1}^{\infty} \|f_{\ell}\|$ converge.

Por otra parte evitaremos hablar de series en espacios que solamente son semi-normados porque en ellos el límite no necesariamente es único.

7. Si existe una colección $\{e_k\}_{k=1}^{\infty} \subset X$ tal que para todo $f \in X$ existe una única sucesión de escalares $\{\lambda_k\}_{k=1}^{\infty} \subset \mathbb{K}$ que verifica

$$\lim_{k \rightarrow \infty} \left\| f - \sum_{\ell=1}^k \lambda_{\ell} e_{\ell} \right\| = 0$$

Entonces a $\{e_k\}_{k=1}^{\infty}$ se le llama una base de Schauder de X y es claro que el espacio es separable en esta situación.

8. Dos normas $\|\cdot\|_1$ y $\|\cdot\|_2$ sobre X se dicen equivalentes si existen constantes $\alpha, \beta > 0$ tales que $\alpha\|f\|_1 \leq \|f\|_2 \leq \beta\|f\|_1$ para toda $f \in X$.

Cabe resaltar que si la dimensión de X es finita todas las normas son equivalentes.

9. El subespacio lineal de X^* que consiste de todos los funcionales lineales y continuos se conoce como dual topológico o dual continuo y se escribe X' , éste es siempre un espacio de Banach, según 5. Existe un encaje natural $X \rightarrow X''$ y aquellos espacios que son isométricamente isomorfos a su bidual se llaman reflexivos.

Para finalizar esta sección pasaremos a considerar otra estructura definida sobre un espacio lineal que pretende abstraer las propiedades y características geométricas de espacios euclidianos y llevar estas nociones a un contexto más general.

Como descubriremos más adelante, ésta estructura se presenta de manera natural en varios espacios lineales que ya nos son conocidos, y a través de ésta obtendremos una visión geométrica de los mismos que abrirá las puertas a muchos resultados fuertes y de consecuencias profundas.

Definición A.2.3. Un producto interior, producto interno o producto escalar sobre un espacio lineal X es una aplicación $\langle\langle \cdot | \cdot \rangle\rangle : X \times X \rightarrow \mathbb{K}$ tal que para cualesquiera $f, g, h \in X$ y $\lambda, \mu \in \mathbb{K}$ se verifica que:

1. $\langle\langle \lambda f + \mu g | h \rangle\rangle = \lambda \langle\langle f | h \rangle\rangle + \mu \langle\langle g | h \rangle\rangle$.
2. $\langle\langle g | f \rangle\rangle = \overline{\langle\langle f | g \rangle\rangle}$.
3. $\langle\langle f | f \rangle\rangle \geq 0$.
4. $\langle\langle f | f \rangle\rangle = 0$ si y solo si $f = 0$.

Si la condición 4 no se cumple, diremos que $\langle\langle \cdot | \cdot \rangle\rangle$ es un semi-producto interno, y en ese caso escribiremos $\langle \cdot | \cdot \rangle$ cuando necesitemos enfatizar que no es producto interno.

Definición A.2.4. Diremos que dos elementos $f, g \in X$ son ortogonales si $\langle\langle f | g \rangle\rangle = 0$ y lo denotaremos con $f \perp g$. También diremos que f es ortogonal a un subconjunto

$S \subset X$ si es ortogonal a todos sus elementos y lo escribiremos como $f \perp S$.

Asimismo, $S_1 \perp S_2$ significa que todos los elementos de S_1 son ortogonales a todos los de S_2 . Y finalmente, el complemento ortogonal de un subconjunto S se define como $S^\perp = \{g \in X : \langle f|g \rangle = 0, f \in S\}$ y claramente es un subespacio lineal y cerrado.

La siguiente proposición recoge las propiedades fundamentales que exhibe un semi-producto escalar y que será conveniente tener en mente en lo sucesivo.

Proposición A.2.5. *Sea X un espacio con semi-producto interno y $\|f\| = \langle f|f \rangle^{\frac{1}{2}}$.*

Entonces para cualesquiera $f, g, h \in X$ y $\lambda, \mu \in \mathbb{K}$:

1. $\langle f|\lambda g + \mu h \rangle = \bar{\lambda} \langle f|g \rangle + \bar{\mu} \langle f|h \rangle$.
2. $\|\lambda f\| = |\lambda| \|f\|$.
3. $\|f + g\|^2 = \|f\|^2 + 2\Re \langle f|g \rangle + \|g\|^2$, *la identidad polar.*
4. $|\langle f|g \rangle| \leq \|f\| \|g\|$, *la desigualdad de Cauchy-Bunyakovsky-Schwarz (CBS).*
5. $\|f + g\| \leq \|f\| + \|g\|$, *la desigualdad del triángulo.*
6. $\|f + g\|^2 = \|f\|^2 + \|g\|^2$ *si $f \perp g$, el teorema de Pitágoras.*
7. $\|f + g\|^2 + \|f - g\|^2 = 2\|f\|^2 + 2\|g\|^2$, *la identidad del paralelogramo.*

Si además $\langle \cdot | \cdot \rangle$ es un producto interno,

8. $\|f\| = 0$ *si y solo si $f = 0$.*
9. $\langle f|g \rangle = 0$ *para todo g si y solo si $f = 0$.*

Demostración. .

1. De las propiedades 1 y 2 de la definición de producto interno se sigue que:

$$\langle f|\lambda g + \mu h \rangle = \overline{\langle \lambda g + \mu h|f \rangle} = \overline{\lambda \langle g|f \rangle + \mu \langle h|f \rangle} = \bar{\lambda} \langle f|g \rangle + \bar{\mu} \langle f|h \rangle.$$

2. Del inciso anterior,

$$\|\lambda f\|^2 = \langle \lambda f|\lambda f \rangle = \lambda \bar{\lambda} \langle f|f \rangle = |\lambda|^2 \|f\|^2$$

3. Vemos directamente que

$$\begin{aligned}\|f + g\|^2 &= \langle f + g | f + g \rangle = \langle f | f \rangle + \langle f | g \rangle + \langle g | f \rangle + \langle g | g \rangle \\ &= \|f\|^2 + \langle f | g \rangle + \overline{\langle f | g \rangle} + \|g\|^2 \\ &= \|f\|^2 + 2\Re \langle f | g \rangle + \|g\|^2\end{aligned}$$

4. Sea $\theta \in \mathbb{R}$ tal que $\langle f | g \rangle = |\langle f | g \rangle| e^{-i\theta}$. Usando la identidad polar con $f + te^{-i\theta}g$ obtenemos que para cualquier $t \in \mathbb{R}$:

$$\begin{aligned}0 \leq \|f + te^{-i\theta}g\|^2 &= \|f\|^2 + 2\Re \left(te^{i\theta} \langle f | g \rangle \right) + t^2 \|g\|^2 \\ &= \|f\|^2 + 2t |\langle f | g \rangle| + t^2 \|g\|^2 =: q(t)\end{aligned}$$

Entonces $q(t)$ es un polinomio cuadrático con a lo más una raíz real. Esto significa que su discriminante es no positivo, es decir, $0 \geq (2 |\langle f | g \rangle|)^2 - 4(\|f\|^2)(\|g\|^2)$, de donde se sigue la afirmación.

5. La desigualdad CBS nos dice que $\Re \langle f | g \rangle \leq |\langle f | g \rangle| \leq \|f\| \|g\|$. Haciendo uso de la identidad polar:

$$\|f + g\|^2 \leq \|f\|^2 + 2\|f\| \|g\| + \|g\|^2 = (\|f\| + \|g\|)^2$$

6. Se sigue inmediatamente de la identidad polar, pues $\langle f | g \rangle = 0$.

7. Usando la identidad polar, $\|f - g\|^2 = \|f\|^2 - 2\Re \langle f | g \rangle + \|g\|^2$, luego sumando $\|f + g\|^2$ obtenemos la identidad del paralelogramo.

8. Es inmediato de la propiedad 4 en la definición de producto escalar.

9. Evidentemente si $f = 0$, $\langle f | g \rangle = 0$ para cualquier g . Conversamente si $\langle f | g \rangle = 0$ para cualquier g , en particular sucede que $0 = \langle f | f \rangle = \|f\|^2$ y por el inciso anterior, que $f = 0$.

□

Nota A.2.6. La proposición [A.2.5](#) muestra que un (semi-)producto interno induce una (semi-)norma definiendo $\|f\| = \langle f | f \rangle^{\frac{1}{2}}$.

A.3. Espacios de Hilbert

En esta sección presentamos una breve introducción a la teoría general de espacios de Hilbert, las propiedades que los caracterizan y los resultados más importantes que permiten utilizar ésta teoría para la resolución de ecuaciones diferenciales parciales.

Definición A.3.1. Un espacio de Hilbert es un espacio lineal \mathcal{H} provisto de un producto interno $\langle\langle \cdot | \cdot \rangle\rangle$ cuya norma inducida convierte a \mathcal{H} en un espacio métrico completo.

Ejemplo A.3.2. Veamos algunos ejemplos importantes de espacios con producto interior que constituyen un espacio de Hilbert.

1. El conjunto de sucesiones (reales o complejas) $\varphi = (\varphi_k)_{k=1}^{\infty}$ tales que la suma $\sum_{k=1}^{\infty} |\varphi_k|^2$ converge se suele denotar como l_2 o $l_2(\mathbb{N})$.

Es un espacio de Hilbert con el producto escalar dado por

$$\langle\langle \varphi | \psi \rangle\rangle_{l_2} = \sum_{k=1}^{\infty} \varphi_k \overline{\psi_k}$$

2. El espacio de funciones complejas cuadrado integrables con respecto a una medida boreliana μ se denota con $L^2(\mathbb{R}, \mu)$ o $L^2(\mu)$ y es un espacio de Hilbert con el producto interno definido como

$$\langle\langle f | g \rangle\rangle_{L^2(\mu)} = \int_{\mathbb{R}} f(t) \overline{g(t)} dt$$

La misma definición funciona para cualquier espacio de medida (X, \mathcal{M}, μ) y denotaremos al espacio correspondiente con $L^2(X, \mathcal{M}, \mu)$, o simplemente $L^2(X, \mu)$ cuando la σ -álgebra \mathcal{M} esté clara del contexto.

3. El espacio $H^k(\Omega) = \{f \in L^2(\Omega) : D^\alpha f \in L^2(\Omega), |\alpha| \leq k\}$, donde $\Omega \subset \mathbb{R}^d$, se conoce como el espacio de Sobolev y es un espacio de Hilbert con el producto

$$\langle\langle f | g \rangle\rangle_{H^k(\Omega)} = \sum_{|\alpha| \leq k} \int_{\Omega} D^\alpha f \overline{D^\alpha g}$$

Aquí $\alpha \in \mathbb{N}^d$ es un multi-índice y seguiremos las convenciones usuales para éstos:

- La norma del multi-índice $\alpha = (\alpha_1, \dots, \alpha_d) \in \mathbb{N}^d$ se define como $|\alpha| = \sum_{\ell=1}^d \alpha_\ell$.
- El operador de derivada D^α se entiende como la derivada $\frac{\partial^{|\alpha|}}{\partial x_1^{\alpha_1} \dots \partial x_d^{\alpha_d}}$.
- Las desigualdades entre multi-índices se refieren al orden lexicográfico, es decir, $\alpha \leq \tilde{\alpha}$ si y solo si $\alpha_k \leq \tilde{\alpha}_k$ para todo $k \in \{1, \dots, d\}$.

La estructura de producto interno nos permite entender a los subespacios de un espacio de Hilbert desde un punto de vista geométrico muy similar al que empleamos al estudiar espacios euclidianos. En particular se tiene el siguiente resultado, del que se desprenderán fuertes consecuencias.

Teorema A.3.3. (*Teorema de proyección ortogonal de Hilbert*)

Sea $S \subset \mathcal{H}$ un subconjunto convexo, cerrado y no vacío y sea $h \in \mathcal{H}$. Entonces existe un único elemento $s_0 \in S$ tal que

$$\|h - s_0\| = \inf_{s \in S} \|h - s\| = \text{dist}(h, S)$$

Si además S es un subespacio lineal y cerrado de \mathcal{H} se tiene que $(h - s_0) \perp S$. Conversamente si $s_0 \in S$ es tal que $(h - s_0) \perp S$, entonces se tiene que $\|h - s_0\| = \text{dist}(h, S)$ y s_0 es el único elemento de S con dicha propiedad.

Definición A.3.4. Dado un subespacio lineal y cerrado $V \subset \mathcal{H}$ definimos $P_V : \mathcal{H} \rightarrow \mathcal{H}$, la proyección ortogonal sobre V , como la función lineal que a cada $f \in \mathcal{H}$ le asocia el -único- elemento que provee el teorema de proyección ortogonal.

Nota A.3.5. Se puede comprobar fácilmente que un operador de proyección P_V así definido satisface

- $\|P_V h\| \leq \|h\|$ para todo $h \in \mathcal{H}$.
- $P_V^2 = P_V$, es decir P_V es idempotente.
- $\mathcal{N}(P_V) = V^\perp$ y $P_V \mathcal{H} = V$.

Estamos en condiciones de presentar otro de los resultados que juegan un papel crucial en el desarrollo de la teoría, que nos permitirá caracterizar la clase de funcionales lineales y continuos a través de una identificación con el espacio original.

Teorema A.3.6. (Teorema de representación de Riesz)

Sea \mathcal{H} un espacio de Hilbert y $\varphi \in \mathcal{B}(\mathcal{H}, \mathbb{K})$ un funcional lineal y acotado.

Entonces existe un único $h_\varphi \in \mathcal{H}$ tal que

$$\varphi(h) = \langle\langle h | h_\varphi \rangle\rangle_{\mathcal{H}}$$

y además se tiene que $\|\varphi\|_{\mathcal{B}(\mathcal{H})} = \|h_\varphi\|_{\mathcal{H}}$.

Demostración. .

Sea $\mathcal{N} = \mathcal{N}(\varphi)$ el espacio nulo de φ . Como φ es lineal y continuo, se tiene que \mathcal{N} es un subespacio lineal y cerrado de \mathcal{H} , que además es propio excepto cuando $\varphi = 0$, y en éste caso se sigue inmediatamente de la condición $\|\varphi\| = \|h_\varphi\|$ que $h_\varphi = 0$.

Supongamos entonces que $\varphi \neq 0$ -i.e. $\mathcal{N} \subsetneq \mathcal{H}$ - y tomemos $h_0 \in \mathcal{N}^\perp \setminus \{0\}$, de manera que para cualquier $h \in \mathcal{H}$ se cumple que

$$\varphi\left(h - \frac{\varphi(h)}{\varphi(h_0)}h_0\right) = \varphi(h) - \frac{\varphi(h)}{\varphi(h_0)}\varphi(h_0) = 0$$

es decir, $\left(h - \frac{\varphi(h)}{\varphi(h_0)}h_0\right) \in \mathcal{N}$ para todo $h \in \mathcal{H}$. Por lo tanto, como $h_0 \perp \mathcal{N}$,

$$0 = \left\langle\left\langle h - \frac{\varphi(h)}{\varphi(h_0)}h_0 \middle| h_0 \right\rangle\right\rangle = \langle\langle h | h_0 \rangle\rangle - \frac{\varphi(h)}{\varphi(h_0)}\|h_0\|^2$$

Y así

$$\varphi(h) = \frac{\varphi(h_0)}{\|h_0\|^2} \langle\langle h | h_0 \rangle\rangle = \left\langle\left\langle h \middle| h_0 \frac{\varphi(h_0)}{\|h_0\|^2} \right\rangle\right\rangle$$

Pongamos $h_\varphi = \frac{\varphi(h_0)}{\|h_0\|^2}h_0$. Como $\varphi(h_\varphi) = \|h_\varphi\|^2$, entonces $\|\varphi\| \geq \|h_\varphi\|$. Por otra parte, de la desigualdad CBS se sigue que $\|\varphi\| \leq \|h_\varphi\|$ y así $\|\varphi\| = \|h_\varphi\|$, como se pedía.

Para establecer la unicidad, supongamos que existe \widetilde{h}_φ con las mismas propiedades, entonces para todo $h \in \mathcal{H}$ se tiene que

$$\left\langle\left\langle h \middle| h_\varphi - \widetilde{h}_\varphi \right\rangle\right\rangle = \varphi(h) - \varphi(h) = 0$$

En particular poniendo $h = h_\varphi - \widetilde{h}_\varphi$ vemos que $\|h_\varphi - \widetilde{h}_\varphi\| = 0$ y por lo tanto $\widetilde{h}_\varphi = h_\varphi$. \square

Pasemos ahora a encontrar una noción apropiada de base para un espacio de Hilbert. Si bien existe una base algebraica siendo éste un espacio lineal, ésta en general resulta demasiado grande y poco manejable en la práctica. Además, como estamos a punto de ver, el producto interior proporciona estructura suficiente para encontrar lo más cercano a una base sobre la cual se permitan combinaciones lineales infinitas.

Definición A.3.7. Decimos que un subconjunto $\mathcal{E} \subset \mathcal{H}$ es ortonormal -o que es un sistema ortonormal, abreviado s.o.n.- si sus elementos son ortogonales dos a dos y tienen norma igual a 1. Si además sucede que el único vector ortogonal a \mathcal{E} es el 0, diremos que \mathcal{E} es un conjunto completo y nos referiremos a él como una base ortonormal (b.o.n.). Es posible demostrar que cualesquiera dos bases ortonormales de \mathcal{H} son de la misma cardinalidad y a través de ésta cantidad común definimos la dimensión de \mathcal{H} .

Proposición A.3.8. .

1. Si $\mathcal{E} = \{e_1, \dots, e_k\} \subset \mathcal{H}$ es un conjunto ortonormal, entonces $P = P_{\text{span } \mathcal{E}}$ está dada por $Ph = \sum_{\ell=1}^k \langle h | e_\ell \rangle e_\ell$ para todo $h \in \mathcal{H}$.
2. Dado un subconjunto $S = \{s_k\}_{k=1}^\infty \subset \mathcal{H}$ linealmente independiente, existe un subconjunto ortonormal $\mathcal{E} = \{e_k\}_{k=1}^\infty$ tal que $\text{span } \mathcal{E} = \text{span } S$. El procedimiento para generar \mathcal{E} se llama proceso de ortogonalización de Gram-Schmidt.
3. Si $\mathcal{E} = \{e_\ell\}_{\ell=1}^\infty \subset \mathcal{H}$ es un conjunto ortonormal y $h \in \mathcal{H}$, se cumple que $\sum_{\ell=1}^\infty |\langle h | e_\ell \rangle|^2 \leq \|h\|^2$, la desigualdad de Bessel.
4. Si además, en el inciso anterior, \mathcal{E} es completo, la igualdad tiene lugar. Ésta se conoce como la identidad de Parseval.

Demostración. .

1. En efecto, puesto que para todo $h \in \mathcal{H}$ se tiene que $(h - \langle h | e_\ell \rangle e_\ell) \perp e_\ell$, de la ortogonalidad de \mathcal{E} se sigue que $\left(h - \sum_{\ell=1}^k \langle h | e_\ell \rangle e_\ell\right)$ es ortogonal a todos los elementos del mismo y por lo tanto a $\text{span } \mathcal{E}$. Pero, por definición, Ph es el único elemento de $\text{span } \mathcal{E}$ tal que $(h - Ph) \perp \text{span } \mathcal{E}$. Luego $Ph = \sum_{\ell=1}^k \langle h | e_\ell \rangle e_\ell$.

2. La construcción de Gram-Schmidt es como sigue:

$$\begin{aligned}
 e_1 &= \frac{s_1}{\|s_1\|}, \\
 e_2 &= \frac{r_2}{\|r_2\|}, & \text{con } r_2 &= s_2 - P_1 s_2, & \text{donde } P_1 &= P_{\text{span}\{e_1\}} \\
 &\vdots & &\vdots & & \\
 e_k &= \frac{r_k}{\|r_k\|}, & \text{con } r_k &= s_k - P_k s_k, & \text{donde } P_k &= P_{\text{span}\{e_1, \dots, e_k\}} \\
 &\vdots & &\vdots & &
 \end{aligned}$$

Note que $\|r_k\| \neq 0$ para toda k debido a la independencia lineal de S .

Entonces es claro que $e_k \perp e_\ell$ para $1 \leq \ell \leq k$ y que $\|e_k\| = 1$, es decir, que \mathcal{E} es un sistema ortonormal. Ahora, por construcción, cada e_k es combinación lineal de s_1, \dots, s_k , y con ello, $\mathcal{E} \subset \text{span } S$. No es difícil ver de la misma forma que $S \subset \text{span } \mathcal{E}$ y por lo tanto $\text{span } \mathcal{E} = \text{span } S$.

3. Recordemos que $\|Ph\| \leq \|h\|$ para una proyección P y tomemos $P_k = P_{\text{span}\{e_1, \dots, e_k\}}$. Según el primer inciso y –en virtud de la ortogonalidad de \mathcal{E} – por el teorema de Pitágoras, para cualquier k tenemos que:

$$\sum_{\ell=1}^k |\langle h|e_\ell\rangle|^2 = \left\| \sum_{\ell=1}^k \langle h|e_\ell\rangle e_\ell \right\|^2 = \|P_k h\|^2 \leq \|h\|^2$$

Dejando $k \rightarrow \infty$ se obtiene la desigualdad de Bessel.

4. La serie $\sum_{\ell=1}^{\infty} \langle h|e_\ell\rangle e_\ell$ converge en \mathcal{H} como consecuencia de la convergencia de $\sum_{\ell=1}^{\infty} |\langle h|e_\ell\rangle|^2$, y además ésta es su norma. Como $\left(h - \sum_{\ell=1}^{\infty} \langle h|e_\ell\rangle e_\ell \right)$ es ortogonal a \mathcal{E} , que es un conjunto completo, se sigue que $h = \sum_{\ell=1}^{\infty} \langle h|e_\ell\rangle e_\ell$.

□

Sobre el operador \mathcal{M}

En el tercer capítulo se introdujo la expresión diferencial

$$\mathcal{M} = I + \beta \mathcal{L}^* \mathcal{L}$$

en cuyos términos se expresa el problema a resolver para la variable de estado. Es de particular interés en éste trabajo el caso en el que la expresión diferencial \mathcal{L} es de la forma $\mathcal{L} = \rho \Delta + \omega \cdot \nabla$, con lo que \mathcal{L}^* , su adjunto formal, queda dado por $\mathcal{L}^* = \rho \Delta - \omega \cdot \nabla - \nabla \cdot \omega$.

En tal situación, procedemos a calcular la expresión explícita del operador \mathcal{M} :

$$\begin{aligned} \mathcal{L}^* \mathcal{L} y &= (\rho \Delta - \omega \cdot \nabla - \nabla \cdot \omega) (\rho \Delta + \omega \cdot \nabla) y \\ &= \rho \Delta (\rho \Delta + \omega \cdot \nabla) y \\ &\quad - \omega \cdot \nabla (\rho \Delta + \omega \cdot \nabla) y \\ &\quad - (\nabla \cdot \omega) (\rho \Delta + \omega \cdot \nabla) y \\ &= \rho^2 \Delta^2 y + \rho \Delta (\omega \cdot \nabla y) \\ &\quad - \rho \omega \cdot \nabla \Delta y - \omega \cdot \nabla (\omega \cdot \nabla y) \\ &\quad - (\nabla \cdot \omega) (\rho \Delta + \omega \cdot \nabla) y \\ &= \rho^2 \Delta^2 y + \rho ((\Delta \omega) \cdot (\nabla y) + 2\mathcal{J}^\omega : \mathcal{H}^y + \omega \cdot (\nabla \Delta y)) \\ &\quad - \rho \omega \cdot (\nabla \Delta y) - \omega \cdot (\mathcal{H}^y \omega) - (\mathcal{J}^\omega \omega) \cdot (\nabla y) \\ &\quad - (\nabla \cdot \omega) (\rho \Delta + \omega \cdot \nabla) y \\ &= \rho^2 \Delta^2 y + (2\rho \mathcal{J}^\omega - \omega \omega^t) : \mathcal{H}^y + (\rho \Delta \omega - \mathcal{J}^\omega \omega) \cdot (\nabla y) \\ &\quad - (\nabla \cdot \omega) (\rho \Delta + \omega \cdot \nabla) y \\ &= \rho^2 \Delta^2 y + (2\rho \mathcal{J}^\omega - \omega \omega^t - \rho (\nabla \cdot \omega) I) : \mathcal{H}^y + (\rho \Delta \omega - \mathcal{J}^\omega \omega - (\nabla \cdot \omega) \omega) \cdot (\nabla y) \\ &= \rho^2 \Delta^2 y + (2\rho \mathcal{J}^\omega - \omega \omega^t - \rho (\nabla \cdot \omega) I) : \mathcal{H}^y + (\rho \Delta \omega - \nabla \cdot (\omega \omega^t)) \cdot (\nabla y) \end{aligned}$$

en donde \mathcal{J}^ω es la matriz jacobiana del campo vectorial ω , \mathcal{H}^y la matriz hessiana de y y la notación $A : B$ para dos matrices (reales) A, B se refiere al producto de Frobenius, dado por $A : B = \text{tr}(A^t B) = \sum_{i,j} a_{ij} b_{ij}$.

Hay que aclarar que en el cálculo anterior hemos considerado al coeficiente de difusión ρ como constante, mientras que sobre ω no hemos impuesto ninguna condición, salvo por los requisitos de regularidad necesarios para que la aplicación de los operadores diferenciales involucrados sea posible en el sentido clásico.

Ahora bien, un caso particular que aparece frecuentemente es aquél en el que $\nabla \cdot \omega = 0$, es decir, cuando ω es incompresible.

En este caso $\mathcal{L}^* \mathcal{L}y$ se reduce a

$$\mathcal{L}^* \mathcal{L}y = \rho^2 \Delta^2 y + (2\rho \mathcal{J}^\omega - \omega \omega^t) : \mathcal{H}^y + (\rho \Delta \omega - \mathcal{J}^\omega \omega) \cdot (\nabla y)$$

Finalmente, si es ω es constante, $\mathcal{L}^* \mathcal{L}y$ queda dado por

$$\mathcal{L}^* \mathcal{L}y = \rho^2 \Delta^2 y - \omega^t \mathcal{H}^y \omega$$

Bibliografía

- [1] Becker, R. and Vexler, B. (2006). Optimal-Control of the Convection-Diffusion equation using Stabilized-Finite-Element-Methods. *Comput. Methods Appl. Mech. Engrg.*, 10:1–15. [iii](#), [41](#)
- [2] Chen, W., Fu, Z.-J., and Chen, C. (2014). *Recent Advances in Radial-Basis-Function Collocation Methods*. University of Oxford. [6](#)
- [3] Chinchapatnam, P., Djidjeli, K., and Nair, P. (2005). Unsymmetric and symmetric meshless schemes for the unsteady convection–diffusion equation. *Comput. Methods Appl. Mech. Engrg.*, 27:1201–1230. [24](#)
- [4] González-Casanova, P., Muñoz-Gómez, J., and Rodríguez-Gómez, G. (2009). Node-Adaptive Domain-Decomposition method by Radial-Basis-Functions. *Numer. Methods Partial Differential Eq.*, 25:1482–1501. [34](#)
- [5] Pearson, J. W. (2013). A radial-basis-function method for solving PDE-constrained optimization-problems. *Num. Alg.*, 64:481–506. [iii](#), [iii](#), [25](#), [41](#)
- [6] Pearson, J. W. and Wathen, A. J. (2013). Fast iterative solvers for convection-diffusion control-problems. *Electronic Transactions on Numerical Analysis.*, 40:294–310. [iii](#), [35](#)
- [7] Rees, T. (2010). *Preconditioning Iterative-Methods for PDE-Constrained-Optimization*. University of Oxford. [iii](#), [17](#), [25](#), [35](#), [37](#), [41](#)
- [8] Shu, C., Ding, H., and Yeo, K. (2003). Local radial-basis-function-based differential-quadrature method and its application to solve two-dimensional incompressible navier–stokes equations. *Comput. Methods Appl. Mech. Engrg.*, 192:941–954. [24](#)
- [9] Stevens, D., Power, H., Lees, M., and Morvan, H. (2010). A Local-Hermitian RBF Meshless numerical method for the solution of Multi-Zone-Problems. *Numer. Methods Partial Differential Eq.*, 27:1201–1230. [iii](#), [6](#), [41](#)