



TESIS DOCTORAL

ACERCA DE SOLUCIONES EXACTAS Y
APROXIMADAS DE UN PROBLEMA DE
DIFUSIÓN NO LINEAL RELACIONADO CON
LA PRODUCCIÓN DE SEMICONDUCTORES

Santiago Vila Doncel

Departamento de Matemáticas

2015



TESIS DOCTORAL

**ACERCA DE SOLUCIONES EXACTAS Y
APROXIMADAS DE UN PROBLEMA DE
DIFUSIÓN NO LINEAL RELACIONADO CON
LA PRODUCCIÓN DE SEMICONDUCTORES**

Santiago Vila Doncel

Departamento de Matemáticas

Conformidad de los directores:

Fdo: Wojciech Okrański.

Fdo: Mariano Rodríguez-Arias Fernández.

2015

A mi madre.

Agradecimientos

Dada la extrema dificultad que conllevaría elaborar una lista exhaustiva de todas las personas a las que quiero agradecer el apoyo prestado, y para no olvidarme de nadie, voy a utilizar algo que parecerá un subterfugio de teoría de conjuntos. Como es sabido, los conjuntos pueden definirse tanto *por extensión*, nombrando a todos y cada unos de sus elementos, como *por comprensión*, mediante una propiedad que los caracterice.

Así pues, gracias a todos los que me han ayudado a escribir esta tesis, y gracias en general a todos los que han sido generosos conmigo, incluyendo a mi familia, mis amigos, y todo el Departamento de Matemáticas de la Universidad de Extremadura.

Mis dos directores merecen una mención especial, por más razones de las que cabrían en esta página. Gracias a Mariano, entre otras muchas cosas, por creer en mí más que yo mismo, y sobre todo, gracias a Wojtek por todo: por su tiempo, por su interés, por su paciencia... Sin él esta tesis sencillamente no existiría.

Gracias a todos.

Índice general

1. Introducción	1
1.1. La ecuación del medio poroso	1
1.2. Base física de la EMP	1
1.3. Otra forma para la EMP	2
1.4. Soluciones de la EMP	3
1.5. Existencia y unicidad de soluciones débiles	5
2. Una solución en serie de potencias	9
2.1. Introducción	9
2.2. Una prueba de convergencia	11
Algunos resultados numéricos	11
Una serie de potencias auxiliar	11
Demostración del resultado principal	15
3. Comportamiento de la serie de potencias	17
3.1. Motivación	17
3.2. Modelos no lineales	18
3.3. Cálculo del gradiente y del hessiano	19
3.4. El método de Levenberg-Marquardt	20
3.5. En la práctica	23
3.6. Algunas estimaciones	23
3.7. Un método de Marquardt generalizado	27
3.8. El primer modelo	28
3.9. El segundo modelo	31
4. Determinación de la posición del interfase	35
5. Otras soluciones aproximadas	41
5.1. Introducción	41
5.2. Preliminares.	42
5.3. La idea de la aproximación.	48
5.4. El problema particular.	50
5.5. Estimaciones de soluciones no triviales	52
5.6. La propiedad contractiva del operador T	60
5.7. Soluciones aproximadas	67
5.8. Nota final y comentarios.	73

Capítulo 1

Introducción

1.1. La ecuación del medio poroso

Se denomina ecuación del medio poroso (EMP) a la siguiente ecuación en derivadas parciales:

$$u_t = \Delta(u^m) \quad m > 1$$

donde u es una función del espacio y el tiempo, y Δ representa el operador laplaciano solamente en las variables espaciales. Esta ecuación aparece de una forma natural en algunas aplicaciones físicas para describir procesos en los que intervienen difusión o transmisión de calor.

La EMP es un ejemplo de ecuación de evolución no lineal, formalmente de tipo parabólico [14]. En cierto sentido, es la versión no lineal más sencilla posible de la ecuación del calor. Escrita en forma de divergencia

$$u_t = \operatorname{div}(D(u)\nabla u)$$

se ve que el coeficiente de difusión $D(u)$ es igual a mu^{m-1} . Queda claro entonces que la ecuación es parabólica solamente en aquellos puntos en los que $u > 0$, y es degenerada cuando $u = 0$, porque D se anula. Por este motivo se dice que la EMP es una ecuación parabólica degenerada. La aplicación física más conocida donde aparece este modelo tan sencillo sea tal vez la descripción del flujo de un gas isentrópico a través de un medio poroso.

A pesar de la simplicidad de la ecuación y de sus aplicaciones, y debido a su carácter no lineal y degenerado, la teoría matemática para la EMP ha sido desarrollada muy recientemente.

1.2. Base física de la EMP

Tomaremos como ejemplo el flujo de un gas ideal a través de un medio poroso. Según explican Muskat [25] y Vázquez [42], este fenómeno se puede describir en

términos de las variables *densidad*, que representamos por u , *presión* representada por v y *velocidad* representado por V , que son funciones del espacio x y del tiempo t .

Estas magnitudes están relacionadas mediante las siguientes leyes:

(1) *Balance de masas.*

$$\rho u_t + \operatorname{div}(uV) = 0 \quad (1.1)$$

(2) *Ley de Darcy* [12], una ley empírica que describe la dinámica de un flujo a través de un medio poroso.

$$\mu V = -k \nabla v \quad (1.2)$$

(3) *Ecuación de estado*, que para un gas ideal establece que

$$v = v_0 u^\gamma \quad (1.3)$$

donde el exponente γ es 1 para procesos isotérmicos y mayor que 1 para procesos adiabáticos.

Estas consideraciones físicas conducen a la restricción $u \geq 0$.

Los parámetros ρ (la porosidad), μ (la viscosidad), k (la permeabilidad) y v_0 (la presión de referencia) se supone que son positivos y constantes.

Después de algunos cálculos, se reduce (1.1)–(1.3) a

$$u_t = c \Delta(u^m)$$

donde $m = 1 + \gamma$ y c es una constante que puede ser omitida mediante un cambio de variable, de donde finalmente se obtiene la EMP.

En lo que sigue, supondremos que $u = u(x, t)$, es decir, u depende de una única coordenada espacial x y del tiempo t .

1.3. Otra forma para la EMP

Las dos ecuaciones siguientes son equivalentes:

$$u_t = \Delta(u^p) = p(u^{p-1}u_x)_x \quad p > 1 \quad (1.4)$$

$$z_t = (z^q z_x)_x \quad q > 0 \quad (1.5)$$

Se ve fácilmente: si $u(x, t)$ es solución de (1.4) entonces $z(x, t) = ku(x, t)$ es solución de (1.5) para $q = p - 1$ y $k = p^{\frac{1}{p-1}} = (q + 1)^{\frac{1}{q}}$. Esto nos permite utilizar la forma (1.5), también habitual en la literatura, en lugar de la forma (1.4). Así pues, retomando las variables u y m , escribiremos la EMP así:

$$u_t = (u^m u_x)_x \quad m > 0$$

Esta será la forma que utilizaremos a lo largo de esta memoria.

1.4. Soluciones de la EMP

Un ejemplo primero y fundamental de solución para la EMP fue obtenido alrededor de 1950 por Zel'dovich y Kompaneets [45] y Barenblatt [4], quienes analizaron la llamada solución *source-type*. Tiene ese nombre porque toma como datos iniciales una masa de Dirac: Cuando t tiende a cero, $u(x, t)$ tiende a $M\delta(x)$, donde M depende de m y de la constante C .

$$u(x, t) = (m + 1)^{\frac{1}{m}} t^{-\frac{1}{m+2}} \left[C - kx^2 t^{-\frac{2}{m+2}} \right]_+^{\frac{1}{m}}$$

donde $[s]_+ = \max\{s, 0\}$, $k = \frac{m}{2(m+2)(m+1)}$ y $C > 0$ es una constante.

Esta solución fue encontrada posteriormente por Pattle [37], pero en la literatura se le llama la solución de Barenblatt.

Posiblemente la observación más importante de la solución *source-type* es que tiene soporte compacto con respecto a x para cada tiempo t , o dicho en sentido físico, la perturbación se propaga con velocidad finita. Esto contrasta fuertemente con una de las propiedades más contestadas de la ecuación clásica del calor, la velocidad infinita de propagación (una solución no negativa de la ecuación del calor es automáticamente positiva en todo su dominio de definición). En cierto sentido, la propiedad de propagación finita hace que la ecuación sea adecuada, desde un punto de vista físico, para modelar procesos de difusión o de propagación del calor. La aparición de este fenómeno es una consecuencia de la degeneración de la ecuación, esto es, del hecho de que el coeficiente D se anula cuando $u = 0$.

El fenómeno de propagación finita implica la aparición de una frontera libre que separa las regiones donde la solución es positiva (esto es, donde hay gas o fluido) de la “región vacía”, donde $u = 0$.

En el caso de la solución *source-type*, la frontera libre η_0 tiene la forma

$$\eta_0(t) = \sqrt{\frac{C}{k}} t^{\frac{1}{m+2}} \quad (1.6)$$

Esta frontera libre o frente de propagación es una materia importante de la investigación matemática.

Una segunda observación es que la solución *source-type* es continua en su dominio pero no es diferenciable en la frontera libre, de nuevo como consecuencia de la pérdida del carácter parabólico de la ecuación cuando u se anula.

La teoría sistemática de la EMP fue iniciada por Oleñik [35] y sus colaboradores alrededor de 1958, quienes introdujeron un concepto apropiado de soluciones generalizadas o débiles.

Desde entonces, muchos matemáticos de distintos países han considerado la EMP, cuyas aplicaciones son muy variadas: el problema de la radiación del calor en un plasma, biología matemática, infiltración de agua, lubricación, teoría de capa límite y otros campos. Para una monografía, véase [20, 8].

Un caso particular

Algunos fenómenos como el envenenamiento (o dopaje) de un semiconductor [22, 24, 43], y otros que citaremos a continuación, se pueden describir bastante bien mediante la siguiente ecuación y condiciones iniciales y de frontera:

$$u_t = (u^m u_x)_x \quad (x, t) \in S = (0, \infty) \times (0, \infty) \quad (m > 0) \quad (1.7)$$

$$u(0, t) = 1 \quad \text{para } t > 0 \quad (1.8)$$

$$u(x, 0) = 0 \quad \text{para } x > 0 \quad (1.9)$$

Para diferentes problemas físicos el valor del parámetro m puede variar. En las aplicaciones a la producción de semiconductores el problema (1.7)–(1.9) se utiliza con $m = 1$ para modelar la difusión de arsénico y boro en silicio; $m = 2$ para fósforo en silicio; $m = 2, 2,5$ o 3 para cinc en arseniuro de galio [24, 43].

Si $m = 1$ entonces el problema (1.7)–(1.9) describe también la infiltración unidimensional de una impureza en un pozo cuyo nivel de agua es constante [39]. Para $m = 3$, las ecuaciones (1.7)–(1.9) pueden modelar muy bien un fluido viscoso vertiéndose lentamente de un depósito a través de una cañería [44].

También hay algunos modelos en los que la ecuación (1.7) se considera con $m \in (0, 1)$, por ejemplo, la propagación de explosiones violentas [3, 5, 7, 21].

En todos los problemas anteriores, la cuestión de interés es el alcance de los fluidos que se propagan, por lo que con respecto a las aplicaciones, es importante encontrar algún método analítico o numérico que nos permita calcular con suficiente exactitud el valor de u .

Este problema (1.7)–(1.9) tiene una única solución [18] llamada solución débil o generalizada $u = u(x, t)$. La función u es no negativa, acotada y continua en S , y satisface una identidad integral apropiada en lugar de (1.7). Sin embargo u es la solución clásica para aquellos puntos $(x, t) \in S$ en los que $u(x, t) > 0$. Se puede demostrar [16, 18] que en el caso (1.7)–(1.9) la solución débil tiene la forma llamada solución de similitud:

$$u(x, t) = \begin{cases} u(\eta) & \text{si } \eta \leq \eta_0 \\ 0 & \text{si } \eta > \eta_0 \end{cases} \quad (1.10)$$

donde $\eta = x/t^{1/2}$ y η_0 , que depende de m , denota la frontera libre del problema, es decir, si se define $\eta(t) = \sup\{x \in (0, \infty) : u(x, t) > 0\}$ entonces $\eta(t) = \eta_0 t^{1/2}$. La curva $x = \eta_0 t^{1/2}$ se llama el interface o la frontera libre de (1.7)–(1.9).

El caso $m = 0$

Si en la ecuación anterior tomamos $m = 0$, entonces la ecuación se convierte en la ecuación del calor clásica, que es lineal, y la solución del problema está dada

[10] por la fórmula

$$u(x, t) = \operatorname{erfc}\left(\frac{x}{2\sqrt{t}}\right)$$

para $x > 0$ y $t > 0$, donde

$$\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$$

y erf es la función error

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-s^2} ds$$

de donde se sigue además que $u(x, t) > 0$ para todo $x > 0$, $t > 0$.

1.5. Existencia y unicidad de soluciones débiles

Consideraremos un problema un poco más general, a saber:

$$\begin{aligned} u_t &= (u^m u_x)_x & \text{en } S_T = (0, +\infty) \times (0, T] & \quad T > 0 \quad m > 0 \\ u(x, 0) &= u_0(x) & \text{para } x \in (0, +\infty) & \\ u(0, t) &= u_1(t) & \text{para } t \in (0, T] & \end{aligned} \quad (1.11)$$

con las siguientes hipótesis:

$$u_0 \in L^\infty(0, \infty), \quad \operatorname{ess\,inf} u_0 \geq 0$$

$$u_1 \in L^\infty(0, T), \quad \operatorname{ess\,inf} u_1 \geq 0$$

y o bien $u_0 \equiv 0$ c.s. en $(0, \infty)$ o bien $\zeta_0 = \operatorname{ess\,sup} \Omega_0 < \infty$, siendo

$$\Omega_0 = \{x \in (0, \infty) : u_0(x) > 0\} \quad \text{y} \quad m(\Omega_0) > 0.$$

donde $m(A)$ representa la medida del conjunto A .

Denotaremos este problema por $P[u_0, u_1]$. Este problema es, a su vez, un caso particular del planteado por Goncerzewicz en [18], a saber:

$$u_t = \varphi(u)_{xx} + b(x)\varphi(u)_x$$

En nuestro caso estamos utilizando

$$\varphi(u) = \frac{u^{m+1}}{m+1}$$

y

$$b(x) = 0$$

Teniendo esto en cuenta, en nuestro caso particular, definimos:

Definición. Una función u definida en S_T se dice que es una solución débil del problema $P[u_0, u_1]$ si:

- (i) u es no negativa, acotada y continua en S_T
- (ii) u satisface la identidad

$$\iint_{S_T} \left[u\xi_t + \frac{u^{m+1}}{m+1} \xi_{xx} \right] dxdt + \int_0^\infty u_0(x)\xi(x,0)dx + \frac{1}{m+1} \int_0^T u_1(t)^{m+1} \xi_x(0,t)dt = 0$$

para toda $\xi \in C^{2,1}(\bar{S}_T)$ no negativa que se anule para $t = T$, $x = 0$ y x suficientemente grande; donde $C^{2,1}(\bar{S}_T)$ denota las funciones dos veces diferenciables con continuidad respecto de x y una vez respecto de t en dicho conjunto.

Teorema. El problema $P[u_0, u_1]$ tiene una única solución débil.

Para la solución débil u del problema, se denota:

$$\Omega(t) = \{x \in (0, \infty) : u(x, t) > 0\} \quad \text{para } 0 < t \leq T$$

y se define

$$\zeta(t) = \begin{cases} \sup \Omega(t) & \text{si } \Omega(t) \neq \emptyset \\ 0 & \text{si } \Omega(t) = \emptyset \end{cases} \quad (1.12)$$

para $0 < t \leq T$, y

$$\zeta(0) = \begin{cases} \zeta_0 & \text{si } m(\Omega_0) > 0 \\ 0 & \text{si } m(\Omega_0) = 0 \end{cases} \quad (1.13)$$

Se puede demostrar lo siguiente:

Teorema. Si $\zeta(0) = 0$ y $\text{ess inf}\{t \in (0, T] : u_1(t) > 0\} = 0$ entonces

$$\zeta \in C([0, T]) \cap C^1((0, T])$$

Pero en nuestro caso, nos interesa además lo siguiente:

$$\begin{aligned} u_0(x, 0) &= 0 \\ u_1(0, t) &= C \quad (\text{constante}) \end{aligned}$$

Basándonos de nuevo en el trabajo de Goncerzewic [18], se demuestra que si u es la solución débil de $P[0, C]$ entonces

$$u(x, t) = f\left(\frac{x}{\sqrt{t}}\right) \quad \text{para } (x, t) \in S_T \quad (1.14)$$

es decir, u es la llamada solución de similaridad de (1.11), donde $f = f(\eta)$ tiene las siguientes propiedades:

i) $f(0) = C$, $f(\eta) > 0$ para $0 < \eta < \eta_0$ y $f(\eta) \equiv 0$ para $\eta_0 \leq \eta < \infty$, donde η_0 es una constante positiva que depende de C .

ii) f es una solución clásica de la ecuación diferencial ordinaria

$$(f^m f')' + \frac{1}{2} \eta f' = 0 \quad \text{en } (0, \eta_0)$$

iii) $f^m(\eta) f' \in C^1[0, \infty)$

iv) $f'(\eta) < 0$ para $0 \leq \eta < \eta_0$

Definición. Si u es una solución débil del problema considerado, denotaremos su soporte de esta forma:

$$\Omega[u] = \{(x, t) \in S_T : u(x, t) > 0\}$$

Corolario. Sea u la solución débil de $P[0, C]$ mencionada en el Teorema de existencia y unicidad (véase página anterior) y sea ζ la función definida por (1.12) y (1.13). Entonces su soporte es precisamente el siguiente:

$$\Omega[u] = \{(x, t) : 0 < x < \zeta(t), t \in (0, T)\} \quad (1.15)$$

donde $\zeta(t) = \zeta_0 t^{1/2}$ y ζ_0 depende de m y C .

En lo que sigue supondremos $C = 1$. Utilizaremos los resultados arriba presentados para continuar con nuestro objetivo de determinar soluciones suficientemente exactas de $P[0, 1]$ y el valor de η_0 .

Capítulo 2

Una solución en serie de potencias

2.1. Introducción

En este capítulo consideraremos una serie de potencias relacionada con nuestro problema no lineal:

$$c_t = (c^m c_x)_x \quad (x, t) \in S = (0, \infty) \times (0, \infty) \quad (m > 0) \quad (2.1)$$

$$c(0, t) = 1 \quad \text{para } t > 0 \quad (2.2)$$

$$c(x, 0) = 0 \quad \text{para } x > 0 \quad (2.3)$$

Este problema ha sido considerado numéricamente para $m = 1, 2, 3$ en [43]. Se ha presentado algunos métodos de aproximación en [1], [13] y [26].

Un enfoque más general para obtener soluciones aproximadas ha sido estudiado por King en [22, 23]. Él ha considerado la solución de similitud:

$$u(\eta) = \begin{cases} c^m(\eta) & \text{si } \eta \leq \eta_0 \\ 0 & \text{si } \eta > \eta_0 \end{cases}$$

donde $\eta = \frac{x}{\sqrt{t}}$ y $\eta = \eta_0$ denota la frontera libre del problema.

En este caso, (2.1)–(2.2) se reduce a:

$$-\frac{1}{2}\eta u' = uu'' + \frac{1}{m}u'^2 \quad (\eta \in [0, \eta_0])$$
$$u(0) = 1$$

Además deben satisfacerse las siguientes condiciones de frontera libre:

$$u(\eta_0) = 0 \quad \lim_{\eta \rightarrow \eta_0^-} u'(\eta) = -\frac{m}{2}\eta_0$$

Poniendo

$$\eta = \eta_0(1 - s) \quad (s \in [0, 1])$$

y

$$u(\eta) = \eta_0^2 m v(s)$$

el problema anterior (y sus condiciones) se convierte en:

$$\begin{aligned} m v v'' + (v')^2 &= \frac{1}{2}(1 - s)v' & (s \in [0, 1]) \\ v(0) &= 0, \quad \lim_{s \rightarrow 0^+} v'(s) = \frac{1}{2} \end{aligned}$$

La frontera libre viene dada por:

$$\eta_0 = \frac{1}{\sqrt{m v(1)}} \quad (2.4)$$

King ha considerado la expansión en serie de potencias de la solución v al problema anterior.

Suponiendo que

$$v(s) = \sum_{n=1}^{\infty} a_n(m) s^n \quad (2.5)$$

para $|s| < \rho_a(m)$, donde $\rho_a(m)$ es el radio de convergencia, se obtiene a partir de lo anterior la siguiente relación de recurrencia:

$$\begin{aligned} \frac{1}{2}(n+1)(1+nm)a_{n+1}(m) &= -\frac{1}{2}n a_n(m) - \sum_{p=1}^{n-1} (p+1) \\ &\times (n+1+p(m-1))a_{n-p+1}(m)a_{p+1}(m) \quad \text{para } n = 2, 3, 4, \dots, \end{aligned} \quad (2.6)$$

con

$$a_1(m) = \frac{1}{2}, \quad a_2(m) = -\frac{1}{4(m+1)} \quad (2.7)$$

La fórmula (2.4) indica la importancia del valor $v(1)$. Para calcular este valor, debemos saber que $\rho_a(m) > 1$, pero King no da ninguna prueba de convergencia para esta expansión. En [32] se demuestra que la diferencia entre los tres primeros términos de la serie de potencias y $v(s)$ es menor que 0.33 % para $m = 1$ y tiende a cero muy rápidamente cuando m tiende a infinito.

En este capítulo veremos que $\rho_a(m) > 1$ si $m \geq m_0 = \frac{\sqrt{17}-1}{8}$ (esta demostración puede encontrarse también en [33]). Esto nos permitirá (en tales casos) calcular η_0 con cualquier precisión.

2.2. Una prueba de convergencia

Resultado principal. Si $m \geq 1$ entonces la serie de potencias (2.5) con coeficientes dados por (2.6)–(2.7) tiene un radio de convergencia mayor que 2.46.

Comentario. De acuerdo con el resultado principal sabemos que la serie $v(s)$ con coeficientes dados por la relación de recurrencia anterior es la única solución a nuestra ecuación diferencial.

A partir de (2.6) y (2.7) se obtiene

$$\begin{aligned} a_3(m) &= \frac{m}{12(2m+1)(m+1)^2}, \\ a_4(m) &= \frac{m(m+3)}{48(3m+1)(2m+1)(m+1)^3}. \end{aligned} \quad (2.8)$$

La relación de recurrencia (2.6) se puede escribir así:

$$\begin{aligned} \frac{1}{2}(n+1)(mn+1)a_{n+1}(m) &= \frac{2n+m(n^2-3n+2)}{4(m+1)}a_n(m) - \sum_{p=2}^{n-2} (p+1) \\ &\quad \times (n+1+p(m-1))a_{n-p+1}(m)a_{p+1}(m) \quad \text{para } n = 4, 5, 6, \dots \end{aligned} \quad (2.9)$$

Utilizando (2.7), (2.8), y (2.9), probaremos el resultado principal de este capítulo.

Algunos resultados numéricos

De acuerdo con el resultado principal se puede utilizar las sumas $\sum_{k=1}^n a_k(m)$ para calcular un valor aproximado de $v(1)$. Con respecto a (2.4) obtenemos:

$$\begin{aligned} m = 1 & \quad \eta_0 \approx 1,6161254468 \\ m = 2 & \quad \eta_0 \approx 1,0903200288 \\ m = 2,5 & \quad \eta_0 \approx 0,9627664316 \\ m = 3 & \quad \eta_0 \approx 0,8705707561. \end{aligned}$$

En cada uno de los casos anteriores, el valor de η_0 ha sido redondeado, es decir, el error es menor que $0,5 \cdot 10^{-10}$.

Una serie de potencias auxiliar

Consideraremos una nueva serie de potencias con coeficientes positivos $d_n(m)$ ($n = 1, 2, 3, \dots; m > 0$) definida como sigue:

$$d_n(m) = n^2|a_n(m)| \quad n = 1, 2, 3, 4 \quad (2.10)$$

$$\begin{aligned}
d_{n+1}(m) &= \frac{(n+1)(2n+m(n^2-3n+2))}{2(mn+1)(m+1)n^2} d_n(m) \\
&+ \frac{2(n+1)}{mn+1} \sum_{p=2}^{n-2} \frac{n+1+p(m-1)}{(n-p+1)^2(p+1)} d_{n-p+1}(m) d_{p+1}(m) \quad \text{para } n = 4, 5, 6, \dots
\end{aligned} \tag{2.11}$$

Para $m > 0$, se definen los siguientes números:

$$K_1(m, n) = \frac{(n+1)(2n+m(n^2-3n+2))}{2(mn+1)(m+1)n^2} \quad n = 3, 4, 5, \dots \tag{2.12}$$

$$K_2(m, n) = \frac{2(n+1)}{mn+1} \sum_{p=2}^{n-2} \frac{n+1+p(m-1)}{(n-p+1)^2(p+1)} \quad n = 4, 5, 6, \dots \tag{2.13}$$

Lema 1.

$$\frac{ap+b}{(n-p+1)^2(p+1)} = \frac{A}{n-p+1} + \frac{B}{(n-p+1)^2} + \frac{C}{p+1}$$

donde

$$A = \frac{-a+b}{(n+2)^2}, \quad B = \frac{a(n+1)+b}{n+2}, \quad C = \frac{-a+b}{(n+2)^2}$$

Demostración. Se comprueba fácilmente, es la descomposición en fracciones simples. \square

Lema 2. Para todo $m > 0$,

$$\sup_{n \geq 3} K_1(m, n) < \infty \tag{2.14}$$

y

$$\sup_{n \geq 4} K_2(m, n) < \infty \tag{2.15}$$

Demostración. Para la primera parte, calculamos el límite directamente:

$$\lim_{n \rightarrow \infty} K_1(m, n) = \frac{m}{2(m+1)m} = \frac{1}{2(m+1)}$$

de donde se deduce (2.14).

Para la segunda parte, usamos el Lema 1 y escribimos $K_2(m, n)$ como sigue:

$$K_2(m, n) = \frac{2(n+1)}{mn+1} \left[2 \frac{n-m+2}{(n+2)^2} \left(\sum_{p=1}^{n-1} \frac{1}{p} - \frac{3}{2} \right) + \frac{m(n+1)}{(n+2)} \left(\sum_{p=1}^{n-1} \frac{1}{p^2} - \frac{5}{4} \right) \right]$$

Como

$$\sum_{p=1}^q \frac{1}{p} \leq \log q + 1; \quad \lim_{q \rightarrow \infty} \frac{\log q}{q} = 0 \quad \text{y} \quad \sum_{p=1}^{\infty} \frac{1}{p^2} = \frac{\pi^2}{6}$$

entonces

$$\lim_{n \rightarrow \infty} K_2(m, n) = \frac{\pi^2}{3} - \frac{5}{3}$$

Por lo tanto (2.15) es cierto. \square

Sea

$$K_1(m) = \sup_{n \geq 3} K_1(m, n) \quad (2.16)$$

y

$$K_2(m) = \sup_{n \geq 4} K_2(m, n) \quad (2.17)$$

Teorema 3. Sea $R(m)$ la raíz positiva de

$$K_2(m)d_3(m)R^2 + K_1(m)R - 1 = 0 \quad (2.18)$$

Si

$$c(m) = d_3(m)R^3(m). \quad (2.19)$$

entonces

$$d_k(m) \leq \frac{c(m)}{R^k(m)} \quad (2.20)$$

para todo $k \geq 3$.

Demostración. Por (2.19) se obtiene (2.20) para $k = 3$. Como

$$d_4(m) = K_1(m, 3)d_3(m)$$

y (2.20) es cierto para $k = 3$, entonces

$$d_4(m) \leq K_1(m) \frac{c(m)}{R^3(m)} + K_2(m) \frac{c^2(m)}{R^5(m)}.$$

Por (2.18) y (2.19) se ve fácilmente que el miembro derecho de la desigualdad anterior es igual a $c(m)/R^4(m)$. De aquí, (2.20) es verdadero para $k = 4$.

Ahora usamos el mecanismo de inducción. Supongamos que (2.20) es cierto para $k = 3, \dots, n$. A partir de (2.11), (2.16) y (2.17) obtenemos

$$d_{n+1}(m) \leq K_1(m) \frac{c(m)}{R^n(m)} + K_2(m) \frac{c^2(m)}{R^{n+2}(m)} \quad (2.21)$$

Por (2.18) y (2.19) se puede calcular fácilmente que el miembro derecho de (2.21) es igual a $c(m)/R^{n+1}(m)$. Por lo tanto (2.20) es cierto para todo $k \geq 3$. \square

Corolario 4. Sea $\rho_d(m)$ el radio de convergencia de la serie de potencias con coeficientes dados por (2.10)–(2.11). De acuerdo con el Teorema 2 obtenemos que $\rho_d(m) \geq R(m)$.

Lema 5. Si

$$m_0 = \frac{\sqrt{17} - 1}{8}$$

entonces

$$K_1(m_0) = K_1(m_0, 3) = 0,499158\dots \quad (2.22)$$

y

$$K_2(m_0) = K_2(m_0, 16) = 1,499021\dots \quad (2.23)$$

Demostración. Para demostrar (2.22) tenemos que

$$K_1(m_0, n) \leq K_1(m_0, 3) \quad \text{para } 3 \leq n \leq 16.$$

Entonces, tenemos que demostrar que

$$\frac{(n+1)(2n+m_0(n^2-3n+2))}{2(m_0n+1)(m_0+1)n^2} \leq \frac{4(6+2m_0)}{18(3m_0+1)(m_0+1)} \quad \text{para } n \geq 17 \quad (2.24)$$

Pero (2.24) se puede demostrar con la siguiente cadena de desigualdades:

$$\begin{aligned} 9 \left(1 + \frac{1}{n}\right) \left(\frac{2}{n} + m_0 \frac{n^2 - 3n + 2}{n^2}\right) (3m_0 + 1) \\ \leq 9 \left(1 + \frac{1}{17}\right) \left(\frac{2}{17} + m_0\right) (3m_0 + 1) < 8m_0(m_0 + 3) \\ \leq 8 \left(m_0 + \frac{1}{n}\right) (m_0 + 3) \quad \text{para } n \geq 17. \end{aligned}$$

Con lo que finaliza la demostración de (2.22).

Para demostrar (2.23) tenemos que

$$K_2(m_0, n) \leq K_2(m_0, 16) \quad \text{para } 4 \leq n \leq 45$$

Si $m \leq 1$ podemos escribir

$$K_2(m, n) \leq \frac{4}{m(n+2)} \left(\log(n-1) - \frac{1}{2}\right) + \frac{\pi^2}{3} - \frac{5}{2}.$$

Para ver que se satisface (2.23) sólo nos hace falta que

$$\frac{\log(n-1) - \frac{1}{2}}{n+2} < \frac{m_0}{4} \left(K_2(m_0, 16) + \frac{5}{2} - \frac{\pi^2}{3}\right) \quad \text{para } n \geq 46. \quad (2.25)$$

El miembro derecho de (2.25) es igual a 0,0692... Como el miembro izquierdo de (2.25) es decreciente para $n \geq 46$ y en $n = 46$ es igual a 0,0688... entonces se obtiene (2.25). Esto acaba la demostración de (2.23). \square

Teorema 6. Si

$$m_0 = \frac{\sqrt{17} - 1}{8} \text{ entonces } \rho_d(m_0) \geq 1,45.$$

Demostración. Utilizando el Teorema 3 y el Lema 5 obtenemos que

$$R(m_0) \geq 1,45.$$

Por el Corolario 4, el Teorema queda probado. \square

Comentario 7. Se puede demostrar que

$$K_1(m) = \lim_{n \rightarrow \infty} K_1(m, n) = \frac{1}{2(m+1)} \text{ para } m \geq \frac{15}{19}.$$

Como en el Lema 5, se puede calcular exactamente que

$$\begin{aligned} K_2(1) &= K_2(1, 20) \\ K_2(2) &= K_2(2, 33) \\ K_2(2,5) &= K_2(2,5, 43) \\ K_2(3) &= K_2(3, 55). \end{aligned}$$

De acuerdo con el Teorema 3, Comentario 7 y los resultados arriba indicados, obtenemos:

$$\begin{aligned} R(1) &\geq 2,46 \\ R(2) &\geq 3,67 \\ R(2,5) &\geq 4,27 \\ R(3) &\geq 4,87. \end{aligned}$$

Demostración del resultado principal

Consideramos la serie de potencias con coeficientes

$$b_n(m) = |a_n(m)| \quad n = 1, 2, 3, 4 \quad (2.26)$$

y

$$\begin{aligned} \frac{1}{2}(n+1)(nm+1)b_{n+1}(m) &= \frac{2n+m(n^2-3n+2)}{4(m+1)}b_n(m) \\ &+ \sum_{p=2}^{n-2} (p+1)(n+1+p(m-1))b_{n-p+1}(m)b_{p+1}(m) \quad \text{para } n = 4, 5, \dots \end{aligned} \quad (2.27)$$

Comentario 8. Sea $\rho_b(m)$ el radio de convergencia de la serie de potencias con coeficientes $b_n(m)$. Como $d_n(m) = n^2b_n(m)$ entonces $\rho_b(m) = \rho_d(m)$.

Lema 9. Si $m_0 \leq m \leq m'$ entonces

$$b_n(m') \leq b_n(m) \quad (2.28)$$

para todo $n \geq 3$.

Demostración. Sea $m > 0$. La fórmula (2.27) se puede reescribir como sigue:

$$b_{n+1}(m) = f_1(m, n)b_n(m) + \sum_{p=2}^{n-2} f_2(m, n, p)b_{n-p+1}(m)b_{p+1}(m) \quad \text{para } n = 4, 5, \dots; \quad (2.29)$$

donde

$$f_1(m, n) = \frac{2n + m(n^2 - 3n + 2)}{2(m+1)(nm+1)(n+1)}$$

y

$$f_2(m, n, p) = \frac{2(p+1)(n+1+p(m-1))}{(mn+1)(n+1)}$$

Si $m' > m > 0$ entonces $f_1(m, n) \geq f_1(m', n)$ ($n \in \mathbf{N}$) y $f_2(m, n, p) \geq f_2(m', n, p)$ ($n \in \mathbf{N}, 2 \leq p \leq n-2$). Como la función $b_3(m)$ es creciente para $m \in [m_0, \infty)$ entonces $b_4(m) = f_1(m, 3)b_3(m)$ es también creciente en el mismo dominio. Por (2.29) y el argumento de inducción se obtiene (2.28). \square

Corolario 10. Si $m' > m \geq m_0$ entonces $\rho_b(m') \geq \rho_b(m)$.

Demostración del resultado principal. Comparando (2.7)–(2.9) y (2.26)–(2.27) se puede ver fácilmente que

$$|a_n(m)| \leq b_n(m).$$

Esto quiere decir que $\rho_a(m) \geq \rho_b(m)$. Si $m \geq m_0$ entonces por el Corolario 10 tenemos que $\rho_b(m) \geq \rho_b(m_0)$. Como por el Comentario 8

$$\rho_b(m) = \rho_d(m)$$

entonces por el Teorema 6 se obtiene $\rho_b(m_0) \geq 1,45$. Esto implica que $\rho_a(m) \geq 1,45$ para $m \geq m_0$.

Comentario 11. Utilizando razonamientos análogos se puede demostrar que $\rho_a(m) \geq 2,46$ para $m \geq 1$.

Capítulo 3

Comportamiento de la serie de potencias

3.1. Motivación

En el capítulo anterior habíamos llegado a una serie de potencias con coeficientes definidos por esta sucesión recurrente:

$$\begin{aligned} a_1(m) &= \frac{1}{2} & a_3(m) &= \frac{m}{12(2m+1)(m+1)^2} \\ a_2(m) &= -\frac{1}{4(m+1)} & a_4(m) &= \frac{m(m+3)}{48(3m+1)(2m+1)(m+1)^3} \end{aligned}$$

y para $n \geq 4$:

$$\begin{aligned} a_{n+1}(m) &= \frac{2n + m(n^2 - 3n + 2)}{2(m+1)(n+1)(mn+1)} a_n(m) \\ &\quad - \frac{2}{(n+1)(mn+1)} \sum_{p=2}^{n-2} (p+1)(n+1+p(m-1)) a_{p+1}(m) a_{n-p+1}(m) \end{aligned}$$

También se demostró la convergencia de esta serie de potencias para algunos valores de m , lo que nos da una cota inferior de su radio de convergencia, pero no sabemos si la cota es buena o mala. En este capítulo vamos a intentar estimar el radio de convergencia de la serie de potencias por métodos experimentales.

Hemos calculado muchos términos de $a_n(m)$ para distintos valores de m , y hemos buscado posibles modelos de comportamiento realizando ajustes de mínimos cuadrados. Uno de los métodos clásicos para realizar ajustes de este tipo es el llamado método de Levenberg-Marquardt. La explicación que sigue está esencialmente tomada de las páginas 682 y siguientes del libro “*Numerical Recipes in C*” [38].

3.2. Modelos no lineales

Supongamos que disponemos de una serie de datos experimentales (x_i, y_i) para $i = 1, \dots, N$, que creemos que obedecen al modelo $y = y(x; a)$ donde a es un vector de M coordenadas a_1, \dots, a_M , es decir, nuestra función modelo y depende de varios parámetros.

Pretendemos determinar qué parámetros a son los que hacen que mejor se ajuste el modelo a los datos experimentales que tenemos. Para ello, lo usual es utilizar una función de mérito χ^2 , que depende de a , de tal forma que elegiremos aquellos parámetros a que hagan mínimo χ^2 .

Cuando el modelo $y(x; a)$ depende linealmente de a , encontrar el mínimo es bastante fácil, pero en general no es el caso.

Sin embargo, este problema no se considera muy distinto del caso lineal, al menos superficialmente. Se espera que, suficientemente cerca del mínimo, la función χ^2 se aproxime bien mediante una forma cuadrática, que se puede escribir como

$$\chi^2(a) \approx \gamma - d \cdot a + \frac{1}{2} a \cdot D \cdot a$$

donde d es un vector de M coordenadas y D es una matriz de $M \times M$. Si la aproximación es buena, sabemos cómo saltar desde los parámetros de prueba actuales a_{act} a los parámetros con los que se consigue el mínimo a_{min} en un sólo paso, concretamente

$$a_{\text{min}} = a_{\text{act}} + D^{-1} \cdot [-\nabla \chi^2(a_{\text{act}})]$$

Por otro lado, la aproximación mediante una forma cuadrática de χ^2 podría ser una aproximación local bastante pobre de la forma de la función que estamos intentando minimizar en a_{act} . En tal caso, todo lo que podemos hacer es bajar en la dirección del gradiente, como en el método del descenso más pronunciado. En otras palabras:

$$a_{\text{siguiente}} = a_{\text{act}} - \text{constante} \times \nabla \chi^2(a_{\text{act}})$$

donde la constante es suficientemente pequeña como para no agotar la dirección descendente.

Para utilizar cualquiera de las dos ecuaciones anteriores, debemos ser capaces de calcular el gradiente de la función χ^2 en cualquier conjunto de parámetros a . Para usar la primera de estas dos, necesitamos también la matriz D , que es la matriz de segundas derivadas (matriz hessiana) de la función de mérito χ^2 , en cualquier a .

Si no tuviéramos la facultad de calcular directamente la matriz hessiana, sino solamente la función que se va a minimizar y posiblemente el gradiente, entonces recurrir a métodos iterativos es casi obligado, no sólo porque la función sea no lineal, sino también para conseguir así información acerca de la matriz hessiana.

Sin embargo aquí es más sencillo, porque conocemos exactamente la forma de la función χ^2 , dado que está basada en una función modelo que hemos especificado nosotros mismos. Por lo tanto conocemos la matriz hessiana. Así que somos libres de utilizar la primera de estas ecuaciones cuando nos parezca. La única razón para utilizar la otra ecuación será la incapacidad de la primera para mejorar el ajuste, como síntoma de que la aproximación alrededor del mínimo dada por la primera fórmula no es una buena aproximación local.

3.3. Cálculo del gradiente y del hessiano

El modelo que se va a ajustar es

$$y = y(x; a)$$

y la función de mérito χ^2 , es:

$$\chi^2(a) = \sum_{i=1}^N \left[\frac{y_i - y(x_i; a)}{\sigma_i} \right]^2$$

Los σ_i son pesos que dependen del error cometido al realizar las mediciones, y se suponen conocidos. El gradiente de χ^2 con respecto a los parámetros a , que valdrá cero en el mínimo de χ^2 , tiene las siguientes componentes:

$$\frac{\partial \chi^2}{\partial a_k} = -2 \sum_{i=1}^N \frac{[y_i - y(x_i; a)]}{\sigma_i^2} \frac{\partial y(x_i; a)}{\partial a_k} \quad k = 1, 2, \dots, M$$

Tomando una derivada parcial adicional tenemos:

$$\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y(x_i; a)}{\partial a_k} \frac{\partial y(x_i; a)}{\partial a_l} - [y_i - y(x_i; a)] \frac{\partial^2 y(x_i; a)}{\partial a_k \partial a_l} \right]$$

Es habitual hacer desaparecer el número 2 definiendo

$$\beta_k \equiv -\frac{1}{2} \frac{\partial \chi^2}{\partial a_k}$$

$$\alpha_{kl} \equiv \frac{1}{2} \frac{\partial^2 \chi^2}{\partial a_k \partial a_l}$$

haciendo que el vector $[\alpha]$ sea $\frac{1}{2}D$ en la primera ecuación (en la que se halla a_{\min} a partir de a_{act}), en cuyos términos se puede reescribir esa ecuación como el conjunto de ecuaciones lineales:

$$\sum_{l=1}^M \alpha_{kl} \delta a_l = \beta_k$$

Este conjunto de ecuaciones se resuelve para los incrementos δa_l que, añadidos a la aproximación actual, dan la siguiente aproximación. En el contexto de los mínimos cuadrados, la matriz $[\alpha]$, igual a la mitad de la matriz hessiana, se llama habitualmente la *matriz de curvatura*.

La ecuación siguiente, la fórmula del descenso más pronunciado, se convierte en

$$\delta a_l = \text{constante} \times \beta_l$$

Nótese que las componentes α_{kl} de la matriz hessiana dependen tanto de las primeras derivadas como de las segundas derivadas de las funciones base con respecto a sus parámetros. Algunos tratamientos simplemente descartan la segunda derivada sin más comentario. Nosotros tampoco la tendremos en cuenta, pero solamente *después* de algunos comentarios.

La segunda derivada aparece porque el gradiente ya depende de $\partial y / \partial a_k$, así que la siguiente derivada debe contener términos con $\partial^2 y / \partial a_l \partial a_k$. El término con la segunda derivada puede descartarse cuando es cero (como sucede en el caso lineal), o cuando es suficientemente pequeño comparado con el término de la primera derivada. Tiene también una posibilidad adicional de ser realmente pequeño en la práctica: El término que multiplica la segunda derivada en esa ecuación es $[y_i - y(x_i; a)]$. Si el modelo es bueno, este término debe ser justamente la medida de error aleatoria para cada punto. Este error puede tener cualquier signo, y en general no debe tener correlación con el modelo. Por lo tanto, los términos con la segunda derivada tienden a cancelarse cuando se suman todos con respecto a i .

La inclusión del término de la segunda derivada puede de hecho producir inestabilidad si el modelo no se ajusta bien o está contaminado por puntos extraños poco propensos a ser compensados por otros puntos de signo opuesto. De ahora en adelante, usaremos siempre como definición de α_{kl} la fórmula

$$\alpha_{kl} = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial y(x_i; a)}{\partial a_k} \frac{\partial y(x_i; a)}{\partial a_l} \right]$$

Esta expresión se asemeja más a su contrapartida lineal. Debe entenderse que enredar un poco (o incluso mucho) con $[\alpha]$ no tiene absolutamente ningún efecto sobre qué conjunto final de parámetros a se alcanzará, sino que afecta únicamente al camino que nos lleva a ese punto. La condición en el mínimo de χ^2 , que $\beta_k = 0$ para todo k , es independiente de cómo se defina $[\alpha]$.

3.4. El método de Levenberg-Marquardt

Marquardt ha puesto en marcha un método elegante, relacionado con una sugerencia anterior de Levenberg, para cambiar paulatinamente entre los extremos del método de la inversa del hessiano y el método del descenso más pronunciado. El segundo método se utiliza lejos del mínimo, cambiando de forma continua

al primer método conforme nos vamos aproximando al mínimo. Este *método de Levenberg-Marquardt* (también llamado *método de Marquardt*) funciona bastante bien en la práctica y durante mucho tiempo se ha considerado el estándar de las rutinas de mínimos cuadrados no lineales.

Este método está basado en dos ideas, elementales pero importantes. Considérese la constante de la ecuación $\delta a_l = \text{constante} \times \beta_l$. ¿Cuál debería ser, incluso en orden de magnitud? ¿Qué establece su escala? No hay información sobre la respuesta en el gradiente. Eso indica solamente la pendiente, pero no hasta dónde se extiende esa pendiente. La primera idea de Marquardt es que las componentes de la matriz hessiana, incluso si no se pueden utilizar de forma precisa, dan *alguna* información sobre la escala (en orden de magnitud) del problema.

La cantidad χ^2 es adimensional, es decir, es un número puro, esto es evidente a partir de su definición. Por otro lado, β_k tiene las dimensiones de $1/a_k$, que bien puede tener dimensiones, por ejemplo, puede tener unidades como m^{-1} o $\text{kw} \cdot \text{h}$, o lo que sea (de hecho, cada componente de β_k puede tener distintas dimensiones). La constante de proporcionalidad entre β_k y δa_k debe por lo tanto tener las dimensiones de a_k^2 . Examinando las componentes de $[\alpha]$ veremos que solamente hay una cantidad obvia con esas dimensiones, que es $1/a_{kk}$, el inverso del elemento de la diagonal. Así que eso debe establecer la escala de la constante. Pero la escala propiamente dicha podría ser demasiado grande. Así que dividamos la constante por algún factor ficticio (adimensional) λ , con la posibilidad de establecer λ mucho mayor que uno para reducir el salto. En otras palabras, sustituimos la ecuación $\delta a_l = \text{constante} \times \beta_l$ por

$$\delta a_l = \frac{1}{\lambda \alpha_{ll}} \beta_l \quad \text{o bien} \quad \lambda \alpha_{ll} \delta a_l = \beta_l$$

Es necesario que todos los α_{ll} sean positivos, pero esto está garantizado por definición (otra razón más para adoptar la última definición considerada de α_{kl}).

La segunda idea de Marquardt es que la ecuación anterior y el sistema de ecuaciones lineales se pueden combinar si se define una nueva matriz α' mediante la siguiente receta:

$$\begin{aligned} \alpha'_{jj} &\equiv \alpha_{jj}(1 + \lambda) \\ \alpha'_{jk} &\equiv \alpha_{jk} \quad (j \neq k) \end{aligned}$$

y entonces las dos ecuaciones se pueden sustituir por una sola:

$$\sum_{l=1}^M \alpha'_{kl} \delta a_l = \beta_k$$

Cuando λ es muy grande, la matriz α' está obligada a ser *diagonal-dominante*, de forma que esta nueva ecuación es casi idéntica que la ecuación de la constante.

Por otro lado, conforme λ se aproxima a cero, la ecuación nos lleva al sistema de ecuaciones lineales.

Dada una estimación inicial del conjunto de parámetros que se han de ajustar, a , la receta recomendada de Marquardt es como sigue:

- Calcular $\chi^2(a)$.
- Escoger un valor modesto para λ , digamos $\lambda = 0,001$.
- (*) Resolver el sistema de ecuaciones lineales para δa y evaluar $\chi^2(a + \delta a)$.
- Si $\chi^2(a + \delta a) \geq \chi^2(a)$, *aumentar* λ en un factor de 10 (o cualquier otro factor sustancial) y volver a (*).
- Si $\chi^2(a + \delta a) < \chi^2(a)$, *disminuir* λ en un factor de 10, actualizar la solución de prueba $a \leftarrow a + \delta a$, y volver a (*).

También es necesaria una condición de paro. Iterar hasta alcanzar la convergencia (hasta la exactitud de la máquina o el límite de redondeo) es generalmente un derroche innecesario dado que el mínimo es, en el mejor de los casos, solamente una estimación estadística de los parámetros a . Un cambio en los parámetros que haga cambiar χ^2 en una cantidad mucho menor que uno nunca es estadísticamente significativa.

Además, no es infrecuente encontrar que los parámetros bailen alrededor del mínimo en un valle plano de topología complicada. La razón es que el método de Marquardt generaliza el método de las ecuaciones normales, de aquí que tenga el mismo problema que ese método con respecto a la casi-degeneración del mínimo. Un fallo completo a causa de un pivote nulo es posible, pero poco probable. Más a menudo, un pivote pequeño provoca una gran corrección que es entonces rechazada, incrementándose entonces del valor de λ . Para λ suficientemente grande, la matriz $[a']$ es definido positiva y puede no tener ningún pivote pequeño. De esta forma el método tiende a mantenerse alejado de los pivotes nulos, pero a costa de una tendencia a bailar mediante el descenso más pronunciado en valles degenerados con muy poca pendiente.

Estas consideraciones sugieren que, en la práctica, uno debería también dejar de iterar a la primera o segunda ocasión en que χ^2 desciende en una cantidad despreciable, digamos menos que 0.01 en términos absolutos o (en caso de que el redondeo impida que eso se alcance) alguna cantidad fraccionaria como 10^{-3} . No se debe parar después de un paso en el que χ^2 *aumenta*. Eso solamente demuestra que λ no se ha ajustado óptimamente.

3.5. En la práctica

Es raro tener que implementar este método *a mano*, pues hay muchos programas que ya se encargan de ello. Por ejemplo, la mayoría de los ajustes individuales que presentaremos en las siguientes secciones han sido realizados con la función `fit` del programa `gnuplot` (para Unix).

La cosa cambia cuando queremos hacer cien, doscientos, o mil ajustes de forma automática, o peor todavía, como veremos en la sección 3.7, cuando el método de ajuste que queremos aplicar no es *exactamente* el método de Marquardt sino una variante del mismo.

3.6. Algunas estimaciones

A simple vista, la sucesión $a_n(m)$ contiene términos positivos y negativos, pero cada vez más pequeños. Lo primero que llama nuestra atención es que $\log |a_n(m)|$ depende de n casi de forma lineal. En las figuras 3.1 y 3.2 tenemos representaciones de $\log |a_n(m)|$ contra n para $m = 1$ y $m = 0,1$ respectivamente.

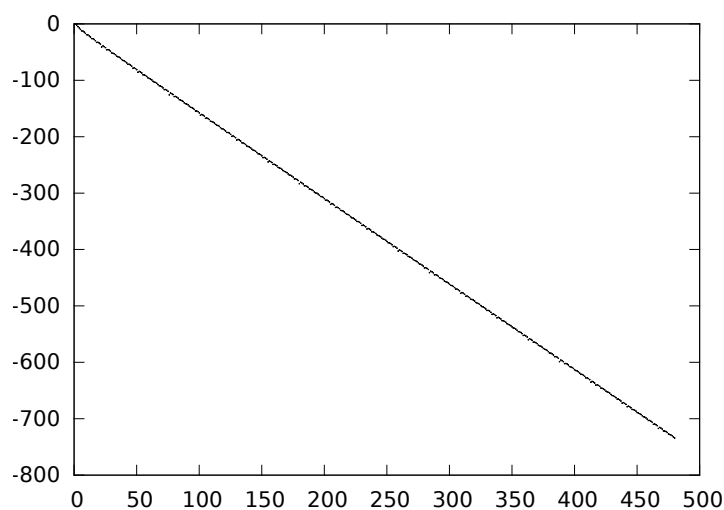


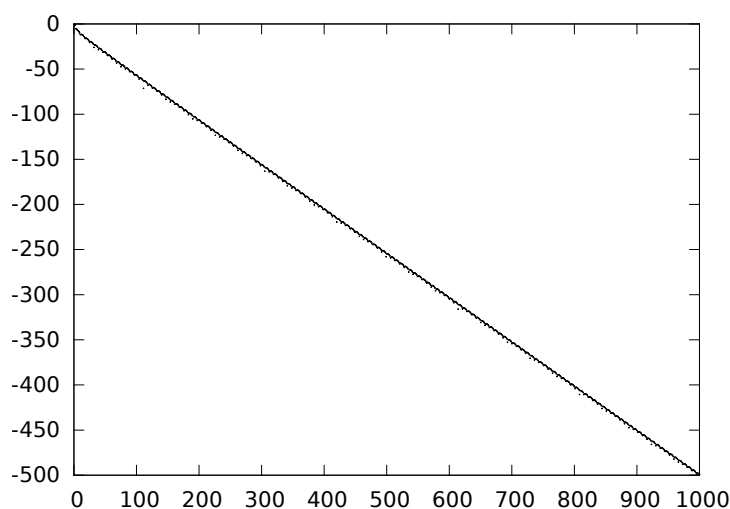
Figura 3.1: $\log |a_n(m)|$ para $m = 1$

Tenemos entonces que la aproximación:

$$\log |a_n(m)| \approx k_1(m)n + k_2(m)$$

es bastante buena, incluso cuando n no es muy grande. De aquí se obtiene, naturalmente, que

$$|a_n(m)| \approx e^{k_1(m)n + k_2(m)}$$

Figura 3.2: $\log |a_n(m)|$ para $m = 0,1$

con lo que el radio de convergencia aproximado de $a_n(m)$ sería $e^{-k_1(m)}$. Con esto tenemos una primera forma de estimar el radio de convergencia.

Veamos cómo se comportan $k_1(m)$ y $k_2(m)$ con respecto a m .

A tal fin, se han generado valores de m desde 0.01 hasta 10 igualmente espaciados en sentido logarítmico (empleando 64 subdivisiones por cada factor de 10). Para cada uno de esos valores, se han calculado 200 términos de la serie, a los que se les ha efectuado el logaritmo del valor absoluto, y a los datos que de aquí resultan, la recta de regresión de mínimos cuadrados, obteniendo así la pendiente y la ordenada en el origen.

Tenemos así un conjunto finito de valores de $k_1(m)$ y $k_2(m)$. Veamos ahora cómo puede obtenerse una estimación de estas funciones.

Para empezar notemos que cuando se representa $k_2(m)$ contra $\log m$ se obtiene la nube de puntos que se muestra en la figura 3.3, que ajustada a una recta, nos da como resultado que:

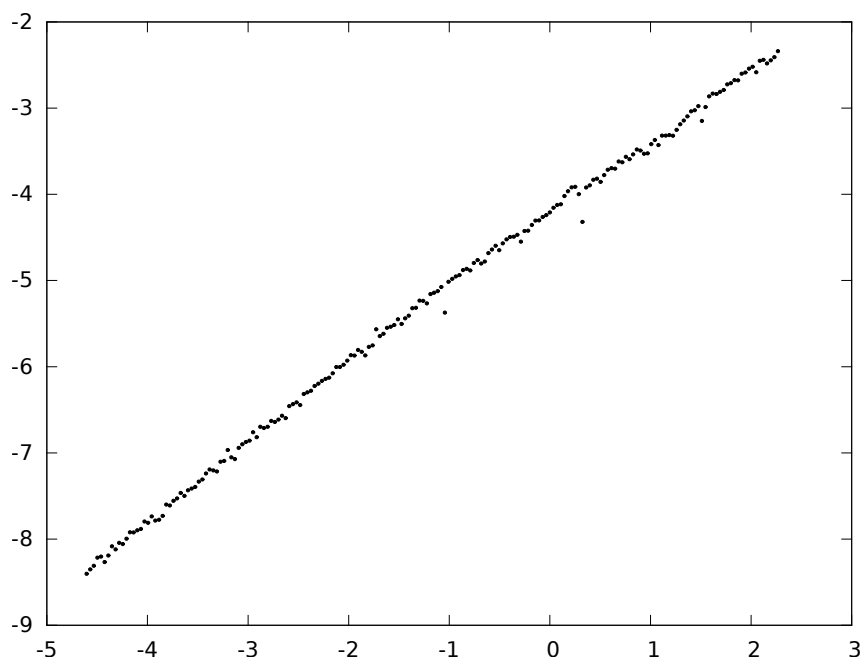
$$k_2(m) \approx q_1 \log m + q_0$$

donde $q_1 = 0,873536$ y $q_0 = -4,24425$. Por las irregularidades de la nube, la suma de cuadrados no es muy buena: 1,2 para un total de 192 puntos.

La figura 3.4 representa $k_1(m)$ contra $\log m$, el otro término que nos queda. Esto puede aproximarse muy bien con una función de la forma $f(x) = -1/(r_1 + r_2 e^{-r_3 x})$ donde $r_1 = 0,151661$, $r_2 = 0,502076$ y $r_3 = 0,556134$, con lo que podemos escribir:

$$k_1(m) \approx \frac{-1}{r_1 + r_2 e^{-r_3 \log m}}$$

En este caso la suma de cuadrados es bastante buena, $4,2 \cdot 10^{-3}$ para 192 puntos.

Figura 3.3: $k_2(m)$ contra $\log m$

Ahora lo lógico es preguntarse qué ocurrirá si de la sucesión $a_n(m)$ eliminamos la componente exponencial que acabamos de aproximar, es decir: ¿Cómo se comporta $e^{-k_1(m)n-k_2(m)}a_n(m)$?

Bien, los datos obtenidos ponen de manifiesto que cuando el parámetro m es pequeño, los dos primeros términos $a_1(m)$ y $a_2(m)$ valen (en el límite) $\frac{1}{2}$ y $-\frac{1}{4}$ respectivamente. Su orden de magnitud, pues, difiere enormemente del resto de los términos. Por ejemplo $a_3(m)$ y los términos que le siguen tienen el mismo orden de magnitud que m . Por ello, es mejor no tener en cuenta estos dos primeros términos en lo que sigue.

La figura 3.5 muestra, para $m = 0,1$, el aspecto de los primeros términos de $e^{-k_1(m)n-k_2(m)}a_n(m)$ a partir de $n = 3$:

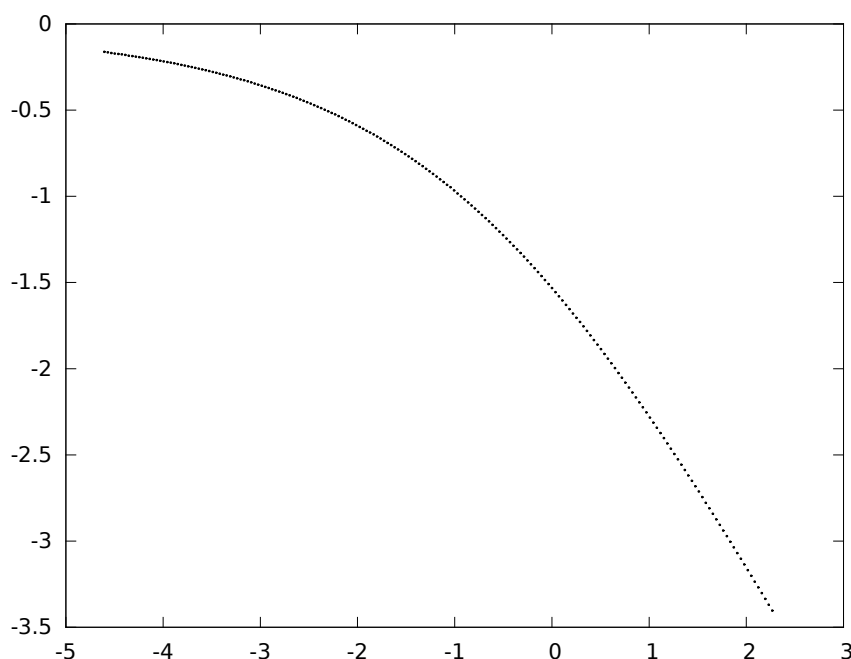
Cualquiera que sea el valor de m , se obtiene una onda senoidal (de mayor o menor periodo). Esto sugiere inmediatamente la siguiente aproximación:

$$a_n \approx e^{k_1 n + k_2} \text{sen}(k_3 n + k_4)$$

donde los parámetros k_1, k_2, k_3, k_4 dependen de m .

Aunque en la gráfica anterior ya puede apreciarse, más adelante veremos que esta aproximación es algo imperfecta, porque habiendo descontado supuestamente ya el “efecto exponencial”, la amplitud de las ondas es cada vez más grande.

Lo que nos interesa ahora mismo es: ¿qué valores deben tener los parámetros k_i

Figura 3.4: $k_1(m)$ contra $\log m$

para que esta aproximación concuerde lo más posible con los datos experimentales? Para ello, deberemos hacer un ajuste de mínimos cuadrados. Pero antes, ¿qué debemos entender por “concordar”, en la frase anterior? ¿Hallamos simplemente la diferencia entre a_n y $e^{k_1 n + k_2} \text{sen}(k_3 n + k_4)$, elevamos al cuadrado, y sumamos? No, porque entonces los primeros términos tendrían mucho más “peso” en el ajuste que los últimos. Para este problema en particular, pensamos que es mucho más razonable considerar la siguiente función de mérito:

$$\chi^2(k) = \sum_{n=n_i}^{n_f} \left(a_n e^{-k_1 n - k_2} - \text{sen}(k_3 n + k_4) \right)^2$$

De esta forma, conseguimos que el peso de cada punto en el ajuste, *después* de la corrección exponencial, sea siempre el mismo.

Esta función de mérito *especial* nos impide aplicar el método de Marquardt tal cual, pues tendríamos un modelo $y = y(x; a)$ y una función de mérito χ^2 en la que no se contempla que y_i esté multiplicado por nada que dependa del vector de parámetros a .

Bien es cierto que una posible solución a este problema hubiera sido utilizar el método de Marquardt dando valores distintos de 1 a los pesos σ_i . Pero buscar unos pesos adecuados, habría que tomarlos de la primera estimación de k_1 y k_2 , y entonces esos pesos podrían diferir de los resultados finales para k_1 y k_2 dados por

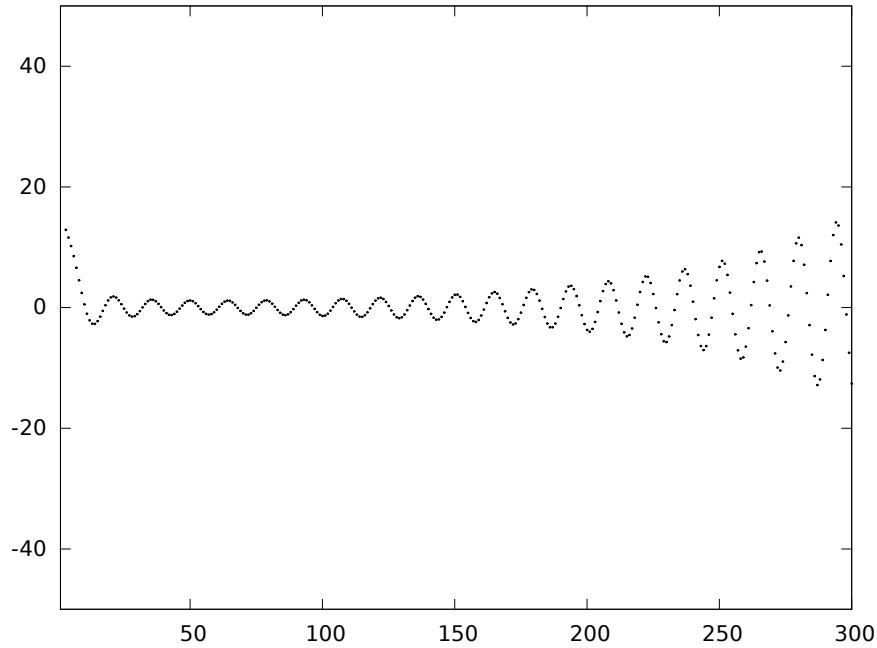


Figura 3.5: $e^{-k_1(m)n - k_2(m)} a_n(m)$

el ajuste.

Así que hay que inventarse otro método de ajuste ligeramente distinto para la función de mérito que hemos escogido, que afortunadamente consiste en una mera generalización del método de Marquardt. Esto se verá con detalle en la siguiente sección.

3.7. Un método de Marquardt generalizado

En nuestro caso particular, el modelo y la función de mérito pueden escribirse de la siguiente forma:

$$y = y(x_i; a) = \frac{r(x_i; a)}{s(a)}$$

$$\chi^2(a) = \sum_{i=1}^N \left[\frac{s(a)y_i - r(x_i; a)}{\sigma_i} \right]^2$$

Por lo tanto el método de Marquardt explicado en la sección 3.4 no se puede aplicar directamente, pues allí la función de mérito era distinta. Pero sí podemos aplicar a este nuevo caso los razonamientos exactos del caso anterior para llegar a un resultado completamente análogo.

Para simplificar la notación, definimos:

$$f(x_i, y_i; a) = r(x_i; a) - s(a)y_i$$

con lo cual podemos decir que las coordenadas del gradiente de χ^2 con respecto a los parámetros a son:

$$\frac{\partial \chi^2}{\partial a_k} = -2 \sum_{i=1}^N \frac{[s(a)y_i - r(x_i; a)]}{\sigma_i^2} \frac{\partial f(x_i, y_i; a)}{\partial a_k} \quad k = 1, 2, \dots, M$$

Y tomando una derivada parcial adicional tenemos que:

$$\frac{\partial^2 \chi^2}{\partial a_k \partial a_l} = 2 \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial f(x_i, y_i; a)}{\partial a_k} \frac{\partial f(x_i, y_i; a)}{\partial a_l} - [s(a)y_i - y(x_i; a)] \frac{\partial^2 f(x_i, y_i; a)}{\partial a_k \partial a_l} \right]$$

y por el mismo motivo por el que entonces despreciábamos las segundas derivadas, al aparecer entonces multiplicadas por $y_i - y(x_i, a)$ y sumadas para todo valor de i , podemos ahora despreciarlas también, al estar multiplicadas por $s(a)y_i - y(x_i; a)$.

En consecuencia, podemos definir:

$$\beta_k = \sum_{i=1}^N \frac{[s(a)y_i - r(x_i; a)]}{\sigma_i^2} \frac{\partial f(x_i, y_i; a)}{\partial a_k}$$

$$\alpha_{kl} = \sum_{i=1}^N \frac{1}{\sigma_i^2} \left[\frac{\partial f(x_i, y_i; a)}{\partial a_k} \frac{\partial f(x_i, y_i; a)}{\partial a_l} \right]$$

y el resto del método de Marquardt es exactamente igual.

3.8. El primer modelo

Llamando k al vector de parámetros (en lugar de a), los primeros ajustes se han realizado con funciones de la forma:

$$y(x_i; k) = e^{k_1 x_i + k_2} \cdot \text{sen}(k_3 x_i + k_4)$$

es decir, si $r(x_i; k) = \text{sen}(k_3 x_i + k_4)$ y si $s(k) = e^{-k_1 x_i - k_2}$, tenemos lo siguiente:

$$f(x_i, y_i; k) = \text{sen}(k_3 x_i + k_4) - y_i e^{-k_1 x_i - k_2}$$

Las derivadas parciales de esta función son las siguientes:

$$\frac{\partial f}{\partial k_1}(x_i, y_i; k) = y_i x_i e^{-k_1 x_i - k_2}$$

$$\frac{\partial f}{\partial k_2}(x_i, y_i; k) = y_i e^{-k_1 x_i - k_2}$$

$$\frac{\partial f}{\partial k_3}(x_i, y_i; k) = \cos(k_3 x_i + k_4) x_i$$

$$\frac{\partial f}{\partial k_4}(x_i, y_i; k) = \cos(k_3 x_i + k_4)$$

Como las ondas no tienen todas la misma amplitud, vamos a hacer varios ajustes, uno por cada *onda*, entendiendo por tal a un número de términos consecutivos de la sucesión a_n ; digamos a_p, a_{p+1}, \dots, a_q ; tales que las siguientes condiciones se cumplen:

- $a_p \geq 0$.
- $a_{p-1} < 0$ si $p > 3$.
- $q = \inf\{n : n > p, a_n < 0, a_{n+1} \geq 0\}$.

Para realizar los ajustes necesitamos estimaciones iniciales de los cuatro parámetros en cuestión. Supongamos que tenemos una onda a_{n_i}, \dots, a_{n_f} . Entonces:

- Para k_1 y k_2 , utilizaremos las funciones obtenidas en la sección 3.6.
- Para estimar k_3 , que es la “velocidad angular”, basta considerar que una estimación del periodo es $n_f - n_i + 1$. Por lo tanto una estimación de k_3 es

$$\frac{2\pi}{n_f - n_i + 1}$$

- Para estimar k_4 , que es la “fase”, tendremos en cuenta que la onda se encuentra en la mitad de su recorrido cuando n es aproximadamente $\frac{n_i + n_f}{2}$. Por lo tanto, una estimación de k_4 es

$$\pi - k_3 \frac{n_i + n_f}{2}$$

En realidad estas estimaciones sólo van a ser necesarias para la primera onda, pues al no diferir demasiado cada onda de la que le sigue, los parámetros finales de ajuste para una determinada onda nos sirven perfectamente como estimaciones iniciales para los parámetros de la onda siguiente.

Efectuados los ajustes, los resultados son buenos, pero hay “algo” que falla. La siguiente tabla es una muestra de los resultados obtenidos para $m = 1$:

n_i	n_f	k_1	k_2	k_3	k_4	χ^2
3	9	-1.9183	0.78	0.699824	5.4830	1.4e-02
10	17	-1.6225	-1.70	0.762163	5.2456	3.9e-04
18	26	-1.5784	-2.47	0.757687	5.3427	6.4e-05
27	34	-1.5593	-2.96	0.755587	5.4063	3.8e-06
35	42	-1.5485	-3.33	0.755350	5.4215	1.4e-06
43	51	-1.5416	-3.62	0.755511	5.4186	2.4e-06
52	59	-1.5368	-3.87	0.755243	5.4347	4.6e-07
60	67	-1.5332	-4.08	0.755198	5.4397	1.6e-07
68	75	-1.5304	-4.26	0.755217	5.4402	2.2e-07
76	84	-1.5283	-4.43	0.755272	5.4371	3.3e-07
85	92	-1.5265	-4.58	0.755174	5.4463	4.1e-08

La siguiente tabla es una muestra de los resultados obtenidos para $m = 0,1$:

n_i	n_f	k_1	k_2	k_3	k_4	χ^2
3	17	-0.6763	-2.91	0.356975	5.8679	1.7e-01
18	31	-0.5354	-4.98	0.434409	4.9495	1.6e-04
32	46	-0.5171	-5.54	0.436629	4.9003	4.4e-05
47	60	-0.5093	-5.90	0.436775	4.9006	5.1e-06
61	75	-0.5049	-6.17	0.437029	4.8896	4.3e-06
76	89	-0.5021	-6.38	0.437031	4.8919	1.0e-06
90	103	-0.5001	-6.55	0.437077	4.8898	4.9e-07
104	118	-0.4986	-6.71	0.437139	4.8847	5.9e-07
119	132	-0.4975	-6.84	0.437129	4.8870	1.7e-07
133	146	-0.4966	-6.96	0.437144	4.8858	1.2e-07
147	161	-0.4959	-7.07	0.437171	4.8827	1.6e-07

Estos ajustes confirman por un lado nuestra hipótesis de utilizar una función senoidal, pues tanto la frecuencia como la fase se mantienen estables de una onda para otra.

Pero por otro lado los resultados correspondientes al segundo parámetro k_2 dejan mucho que desear. Este parámetro no es estable de una onda para otra, sino que es cada vez más negativo.

También hay otro problema: Al calcular los términos de la sucesión $a_n(m)$, los números se hacen demasiado pequeños, con lo que, a los ojos del ordenador, se convierten en cero muy fácilmente (y eso que estamos usando números de doble precisión, que pueden ser tan pequeños como 10^{-324}). Como resultado, el número máximo de términos de la sucesión que se pueden calcular efectivamente es bastante limitado, como se puede apreciar en la siguiente tabla:

m	calculables
10	216
3	315
2	366
1	487
0.3	856
0.1	1497
0.03	2833
0.01	5047

Hay varias soluciones para esto:

- Definir una nueva sucesión:

$$\bar{a}_n(m) = r^n a_n(m) \quad \forall n \in \mathbf{N}$$

donde r es una estimación del radio de convergencia de a_n . No sería necesario que la estimación fuera perfecta, siempre que almacenemos el valor de r junto con los datos \bar{a}_n . Este método nos conduciría a otra relación de recurrencia para \bar{a}_n , casi idéntica que la correspondiente a la sucesión anterior, a_n . El problema es que estaríamos trabajando con dos sucesiones al mismo tiempo, una la real, y otra la calculada, y tendríamos que estar convirtiendo los resultados obtenidos con la sucesión calculada a lo que corresponda para la sucesión real. En un cierto momento hicimos esto que contamos aquí, pero se trata de una complicación innecesaria teniendo en cuenta la solución que finalmente adoptamos, que es la siguiente.

- Usar doble precisión extendida (en C, el tipo `long double`), con lo que nos evitamos los problemas antes mencionados manteniendo la simplicidad del modelo. El tipo `long double` no solamente tiene más precisión que el tipo `double`, sino que también permite que los números puedan ser más grandes, y más pequeños, que es lo que interesa en este caso. No todos los compiladores de C tienen el tipo `long double`, pero dado que nosotros usamos el compilador de C de GNU, `gcc`, que sí lo tiene, y no es previsible que deje de tenerlo, nos parece la mejor solución.

3.9. El segundo modelo

En la sección anterior hemos considerado la aproximación:

$$a_n \approx e^{k_1 n + k_2} \operatorname{sen}(k_3 n + k_4)$$

donde los parámetros k_1, k_2, k_3, k_4 dependen de m . Sin embargo, como ya hemos visto, esta aproximación no es perfecta, porque el valor de k_2 dado por los ajustes de mínimos cuadrados decrece con n .

Ahora proponemos una aproximación mejor:

$$a_n \approx \frac{e^{k_1 n + k_2}}{n^{k_5}} \operatorname{sen}(k_3 n + k_4) \quad (3.1)$$

donde todos los k_i dependen de m .

De acuerdo con la notación introducida en la sección anterior, pero teniendo en cuenta que x_i se refiere realmente a n y que y_i es realmente a_n , escribiremos:

$$f(n, a_n; k) = \operatorname{sen}(k_3 n + k_4) - a_n n^{k_5} e^{-k_1 n - k_2}$$

Ahora, las derivadas parciales de esta función son las siguientes:

$$\begin{aligned}\frac{\partial f}{\partial k_1}(n, a_n; k) &= na_n n^{k_5} e^{-k_1 n - k_2} \\ \frac{\partial f}{\partial k_2}(n, a_n; k) &= a_n n^{k_5} e^{-k_1 n - k_2} \\ \frac{\partial f}{\partial k_3}(n, a_n; k) &= \cos(k_3 n + k_4) n \\ \frac{\partial f}{\partial k_4}(n, a_n; k) &= \cos(k_3 n + k_4) \\ \frac{\partial f}{\partial k_5}(n, a_n; k) &= -a_n n^{k_5} e^{-k_1 n - k_2} \log n\end{aligned}$$

Como estimaciones iniciales de los parámetros, tomaremos las mismas que antes para k_1, \dots, k_4 , y 0 para k_5 . De esta forma comenzamos con un vector de parámetros inicial cuya suma de cuadrados es, al menos, tan buena como si sólo tuviéramos los cuatro primeros parámetros.

Veamos qué resultados se consiguen con este modelo.

La tabla siguiente es una muestra de los resultados obtenidos para $m = 1$:

n_i	n_f	k_1	k_2	k_3	k_4	k_5	χ^2
3	9	-1.309258	2.641	0.80143928	4.861404	3.1894	2.1e-04
10	17	-1.441257	2.239	0.77074262	5.128937	2.4630	7.1e-05
18	26	-1.507681	0.756	0.75878528	5.318221	1.5475	9.9e-06
27	34	-1.516667	0.140	0.75654632	5.376818	1.2878	2.0e-07
35	42	-1.507735	0.839	0.75611616	5.391944	1.5719	5.9e-08
43	51	-1.510376	0.543	0.75574313	5.407592	1.4630	4.0e-08
52	59	-1.511259	0.366	0.75554580	5.417832	1.4067	2.7e-09
60	67	-1.509397	0.678	0.75547033	5.422393	1.5101	2.2e-09
68	75	-1.508431	0.908	0.75541263	5.426223	1.5802	5.0e-10
76	84	-1.509067	0.774	0.75535282	5.430674	1.5381	1.2e-09
85	92	-1.509112	0.776	0.75531746	5.433630	1.5376	3.1e-10

Y esta otra tabla corresponde a $m = 0,1$:

n_i	n_f	k_1	k_2	k_3	k_4	k_5	χ^2
3	17	-0.368039	0.184	0.41541373	5.197787	2.7885	5.1e-02
18	31	-0.485879	-2.320	0.43659688	4.895379	1.2145	3.6e-06
32	46	-0.488023	-2.529	0.43716333	4.879454	1.1341	2.2e-08
47	60	-0.488534	-2.596	0.43719808	4.877926	1.1103	2.3e-10
61	75	-0.488674	-2.622	0.43720227	4.877679	1.1020	1.1e-11
76	89	-0.488723	-2.634	0.43720318	4.877612	1.0984	7.3e-13
90	103	-0.488748	-2.642	0.43720340	4.877591	1.0961	1.2e-13
104	118	-0.488759	-2.646	0.43720353	4.877578	1.0950	3.1e-14
119	132	-0.488766	-2.649	0.43720359	4.877571	1.0942	7.1e-15
133	146	-0.488771	-2.651	0.43720361	4.877568	1.0936	2.2e-15
147	161	-0.488773	-2.653	0.43720363	4.877565	1.0932	9.7e-16

Comparando estas tablas con las obtenidas con el primer modelo, salta a la vista que el parámetro k_1 es más estable que antes (converge más rápidamente), el parámetro k_2 se convierte en estable (antes no lo era), y los parámetros k_3 y k_4 son también más estables, especialmente cuando m es pequeño. La suma de cuadrados ha mejorado mucho, especialmente también cuando m es pequeño. En cuanto al quinto parámetro, no solamente parece estable sino que los indicios apuntan a la siguiente aproximación, que merecería un estudio más detallado:

$$k_5 \approx 1 + \frac{m}{m+1}$$

Para reunir una cantidad suficiente de datos, se han tomado 256 valores distintos de m , desde 0,001 hasta 10, y para cada uno de ellos, se han calculado los tres mil primeros términos de la sucesión $a_n(m)$, después se han hecho ajustes onda por onda, y nos hemos quedado con el valor de k_1 correspondiente a la última onda. Representando de nuevo $k_1(m)$ contra $\log m$ (como ya hicimos en la sección 3.6), obtenemos la figura 3.6.

Como entonces, esta gráfica se aproxima bastante bien mediante una función de la forma $f(x) = -1/(r_1 + r_2 e^{-r_3 x})$, de tal modo que se puede escribir:

$$k_1(m) \approx \frac{-1}{r_1 + r_2 e^{-r_3 \log m}}$$

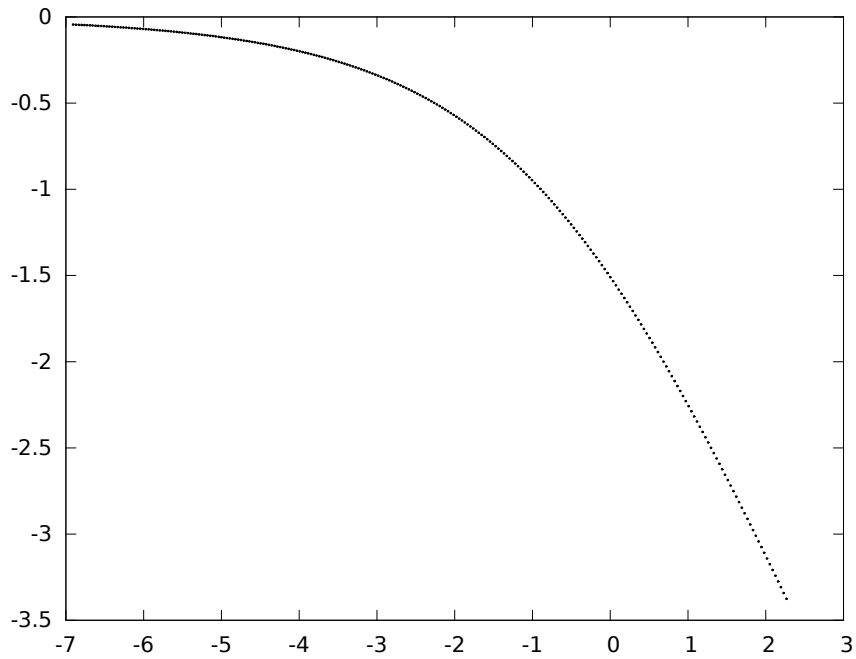
En este caso, los ajustes dan $r_1 = 0,156414$, $r_2 = 0,507456$ y $r_3 = 0,567237$, siendo la suma de cuadrados igual a $2,7 \cdot 10^{-3}$, para 256 puntos (mejor que entonces).

De acuerdo con (3.1), el radio de convergencia de $a_n(m)$ sería:

$$\rho_a(m) = e^{-k_1(m)}$$

y de acuerdo con la aproximación hallada para $k_1(m)$, tendríamos que

$$k_1(m) \leq 0 \quad \forall m > 0 \quad \text{y} \quad \lim_{m \rightarrow 0^+} k_1(m) \rightarrow 0$$

Figura 3.6: $k_1(m)$ contra $\log m$ (segundo modelo)

lo cual querría decir que

$$\rho_a(m) \geq 1 \quad \forall m > 0 \quad \text{y} \quad \lim_{m \rightarrow 0^+} \rho_a(m) = 1$$

Esto es lo que los datos sugieren, aunque obviamente, nos gustaría poder demostrarlo con todo rigor, como se hizo para $m \geq 1$ en el capítulo 2.

Capítulo 4

Determinación de la posición del interfase

Volvemos al problema planteado capítulos atrás:

$$c_t = (c^m c_x)_x \quad (x, t) \in S = (0, \infty) \times (0, \infty) \quad (m > 0) \quad (4.1)$$

$$c(0, t) = 1 \quad \text{para } t > 0 \quad (4.2)$$

$$c(x, 0) = 0 \quad \text{para } x > 0 \quad (4.3)$$

Sabemos que este problema tiene una única solución débil, que además tiene la forma de una solución de similaridad:

$$c(x, t) = \begin{cases} c(\eta) & \text{si } \eta \leq \eta_0, \\ 0 & \text{si } \eta > \eta_0, \end{cases} \quad (4.4)$$

donde $\eta = x/\sqrt{t}$ y $\eta = \eta_0$ denota el interfase (o frontera libre) del problema.

En este capítulo nos gustaría presentar algunos métodos analíticos y numéricos, que nos permiten calcular de forma suficientemente exacta el valor de η_0 . Estos resultados se pueden encontrar en [34].

Si consideramos $u = c^m$ en lugar de c , entonces (4.1)-(4.2) se reduce a:

$$-\frac{1}{2}\eta u' = uu'' + \frac{1}{m}u'^2 \quad (\eta \in [0, \eta_0]) \quad (4.5)$$

$$u(0) = 1 \quad (4.6)$$

Además deben satisfacerse la siguientes condiciones de frontera:

$$u(\eta_0) = 0, \quad \lim_{\eta \rightarrow \eta_0^-} u'(\eta) = -\frac{m}{2}\eta_0 \quad (4.7)$$

Poniendo

$$\eta = \eta_0(1 - s) \quad (s \in [0, 1])$$

y

$$u(\eta) = \eta_0^2 m v(s)$$

el problema anterior se convierte en:

$$m v v'' + (v')^2 = \frac{1}{2}(1-s)v' \quad (4.8)$$

$$v(0) = 0, \quad \lim_{s \rightarrow 0^+} v'(s) = \frac{1}{2} \quad (4.9)$$

La interfase está dada por:

$$\eta_0 = \frac{1}{\sqrt{m v(1)}} \quad (4.10)$$

Como en [32], se puede demostrar que el problema (4.8)-(4.9) tiene la solución única $v \in C([0, 1]) \cap C^2((0, 1])$ que satisface la siguiente desigualdad:

$$v_3(s) \leq v(s) \leq v_3(s) + M(s) \quad (4.11)$$

para $s \in [0, 1]$, donde

$$v_3(s) = \frac{s}{2} + a_2(m)s^2 + a_3(m)s^3,$$

$$M(s) = \frac{2}{3} \left[\frac{1}{2} \left(1 + a_2(m) + \frac{m(2^{1/m} - 1)}{3(m+1)} \right) \right]^{\frac{m-1}{m}} \frac{a_3(m)(m+3)(v_3(1))^{\frac{1}{m}-1} s^{2+\frac{1}{m}}}{(2m+1)(4m+1)}$$

y

$$a_2(m) = -\frac{1}{4(m+1)} \quad (4.12)$$

$$a_3(m) = \frac{m}{12(2m+1)(m+1)^2} \quad (4.13)$$

Si denotamos por

$$\eta_3 \doteq \frac{1}{\sqrt{m v_3(1)}} \quad (4.14)$$

entonces, usando el lado izquierdo de (4.11), se obtiene:

$$\left| 1 - \frac{\eta_3}{\eta_0} \right| \leq \frac{v(1) - v_3(1)}{2v_3(1)} \quad (4.15)$$

Por (4.11) y (4.15) se obtiene:

$$\left| 1 - \frac{\eta_3}{\eta_0} \right| \leq \frac{1}{3} \left[\frac{1}{2} \left(1 + a_2(m) + \frac{m(2^{\frac{1}{m}} - 1)}{3(m+1)} \right) \right]^{\frac{m-1}{m}} \frac{a_3(m)(m+3)[v_3(1)]^{\frac{1}{m}-2}}{(2m+1)(4m+1)} \quad (4.16)$$

Hemos determinado el valor de η_3 para algunos m y hemos utilizado la inecuación anterior para calcular una estimación E_r del error entre η_0 y η_3 . La siguiente tabla muestra los resultados de estos cálculos.

m	η_3	$E_r(\%)$
1	1,61808	0,16
2	1,090609	$3,57 \cdot 10^{-2}$
2,5	0,9629128	$2,00 \cdot 10^{-2}$
3	0,8706533	$1,22 \cdot 10^{-2}$
10	0,4575805	$3,59 \cdot 10^{-4}$

A partir de (4.16) es fácil de ver que el error tiende muy rápidamente a cero cuando m tiende a infinito. De acuerdo con la tabla presentada y (4.16) podemos decir que para $m \geq 1$ el valor:

$$\eta_3 = 2(m+1) \left[\frac{3(2m+1)}{[3(2m+1)^2(m+1)+m]m} \right]^{\frac{1}{2}} \quad (4.17)$$

aproxima a η_0 muy exactamente. Sin embargo para $m < 1$ tenemos que buscar otros métodos.

En [23] King ha considerado la expansión formal en serie de potencias para la solución v de (4.8)–(4.9). Suponiendo que:

$$v(s) = \frac{s}{2} + \sum_{n=2}^{\infty} a_n(m)s^n \quad (4.18)$$

para $|s| < \rho_a(m)$, donde $\rho_a(m)$ es el radio de convergencia, se obtienen $a_2(m)$ y $a_3(m)$ tal y como aparecen en (4.12) y (4.13), respectivamente, y:

$$a_4(m) = \frac{m(m+3)}{48(3m+1)(2m+1)(m+1)^3} \quad (4.19)$$

$$(n+1)(mn+1)a_{n+1}(m) = \frac{2n+m(n^2-3n+2)}{2(m+1)}a_n(m) - 2 \sum_{p=2}^{n-2} (p+1)(n+1+p(m-1))a_{p+1}(m)a_{n-p+1}(m) \quad (4.20)$$

Para calcular η_0 debemos saber que $\rho_a(m) > 1$, pero necesitábamos una prueba de convergencia para la expansión (4.18). Finalmente en [33] se demostró que $\rho_a(m) > 1$ para $m \geq m_0 = \frac{\sqrt{17}-1}{8}$, pero sin ninguna estimación de la diferencia $v(s) - v_n(s)$, donde

$$v_n(s) = \frac{s}{2} + \sum_{k=2}^n a_k(m)$$

Se ha encontrado una aproximación muy buena de $a_n(m)$, a saber:

$$a_n(m) \approx \frac{A}{\rho_a^n(m)n^E} \operatorname{sen}(\omega n + \alpha) \quad (4.21)$$

para

$$n \geq \operatorname{Int} \left[5,45387 + \frac{3,86911}{\sqrt{m}} \right] + 1 \quad (4.22)$$

donde A , ω , α y $E > 0$ dependen de m . En el caso $m \leq 1$ tenemos:

$$A \leq 1 \quad (4.23)$$

Además el radio de convergencia se puede aproximar mediante la fórmula

$$\rho_a(m) \approx \begin{cases} (m+1) \left[1 + \frac{1,5649\sqrt{m}}{1+0,1874m} \right] & m < \frac{1}{2} \\ (m+1) \left[2,8375 - \frac{0,9382}{m+0,6292} \right] & m \geq \frac{1}{2} \end{cases} \quad (4.24)$$

Obviamente $\rho_a(m) > 1$. Con el $\rho_a(m)$ determinado numéricamente podemos establecer algunas estimaciones. Sea $m \leq 1$.

Por (4.21) y (4.23) se obtiene

$$|a_n(m)\rho_a^n(m)| \leq 1$$

para n que satisface (4.22), y de aquí

$$|v(1) - v_n(1)| \leq \sum_{k=n+1}^{\infty} \rho_a^{-k}(m)$$

lo que da la siguiente estimación:

$$|v(1) - v_n(1)| \leq \frac{1}{\rho_a^n(m)} (\rho_a(m) - 1) \quad (4.25)$$

Para cualquier $m \leq 1$ consideraremos únicamente n tales que

$$\frac{1}{\rho_a^n(m)} (\rho_a(m) - 1) \leq 2 \cdot 10^{-7} \quad (4.26)$$

Es fácil comprobar numéricamente que tales n satisfacen (4.22) y entonces $a_n(m)$ se puede sustituir por su aproximación (4.21). Por el lado izquierdo de (4.11) se infiere que:

$$v(1) \geq 0,25 \quad (4.27)$$

Por (4.25)–(4.27) se obtiene:

$$v_n(1) \geq 0,2499998 \quad (4.28)$$

Si denotamos por

$$\eta_n = \frac{1}{\sqrt{mv_n(1)}} \quad (4.29)$$

entonces, como en (4.15), se obtiene:

$$\left| 1 - \frac{\eta_n}{\eta_0} \right| \leq \frac{|v(1) - v_n(1)|}{2v_n(1)} \quad (4.30)$$

Por (4.28) y (4.30) podemos escribir:

$$\left| 1 - \frac{\eta_n}{\eta_0} \right| \leq \frac{1}{0,499996\rho_a^n(m)(\rho_a(m) - 1)} \quad (4.31)$$

Consideraremos errores $E_r \leq 4 \cdot 10^{-5}$ (%)

De acuerdo con (4.31) podemos decir que para

$$n = \text{Int} \left[-\frac{\log(E_r \cdot 0,004999996(\rho_a(m) - 1))}{\log \rho_a(m)} \right] + 1 \quad (4.32)$$

el valor de η_n aproxima η_0 con un error menor que E_r .

Como ejemplo presentamos, para algunos m , el valor de η_n que aproxima η_0 con un error menor que 0,00004%. La siguiente tabla muestra valores de m y n determinados por (4.32) y el valor de η_n .

m	n	η_n
1	10	1,6161254468
0,1	33	5,9273943901
0,01	111	19,7626800537

Algunas consideraciones numéricas muestran que el número n no es el más pequeño, pero da la información acerca de cuántos términos debemos tener en cuenta para obtener η_0 con un error menor que E_r . Esta estimación es útil para distintas aplicaciones. El conocimiento de una buena aproximación de η_0 es muy importante para determinar numéricamente la solución a (4.1)-(4.3).

Capítulo 5

Otras soluciones aproximadas

5.1. Introducción

En el estudio de las soluciones de similitud de algunos problemas de difusión no lineal tales como la infiltración de fluidos en pozos, fabricación de semiconductores, etc ([23], [30], [39], [44]) aparece la siguiente ecuación diferencial

$$(k(u)u')' = f(x)u' \quad (5.1)$$

junto con las condiciones

$$\begin{aligned} u(0) &= 0 \\ \lim_{x \rightarrow 0^+} k(u(x))u'(x) &= 0 \end{aligned} \quad (5.2)$$

Obviamente $u \equiv 0$ es la solución trivial del problema, pero la existencia, unicidad y aproximaciones de soluciones no triviales tienen un interés físico.

Durante los últimos años se han publicado artículos [31], [36], [27], en los que se presentan algunos teoremas acerca de la existencia y unicidad de soluciones no triviales a problemas de este tipo. Pero ninguno de estos trabajos mencionados proporciona buenas soluciones aproximadas a (5.1)–(5.2) tan importantes con respecto a las aplicaciones.

En este capítulo nos gustaría presentar un método que nos permite estimar muy exactamente la solución no trivial de (5.1)–(5.2) cuando k está muy próximo a u^m ($m > 0$). Este último caso es muy frecuente en las aplicaciones. Estos resultados se pueden encontrar en [29].

Veamos de qué forma la ecuación $(k(u)u')' = f(x)u'$ es un caso particular de $c_t = (c^m c_x)_x$. La solución de similitud de $c_t = (c^m c_x)_x$ es de la forma

$$c(x, t) = u\left(\frac{x}{\sqrt{t}}\right) = u(xt^{-1/2})$$

Tenemos entonces que:

$$c_t = \frac{\partial c(x, t)}{\partial t} = u'(xt^{-1/2}) \cdot x \cdot \frac{-1}{2} \cdot t^{-3/2}$$

$$c_x(x, t) = \frac{\partial c(x, t)}{\partial x} = u'(xt^{-1/2})t^{-1/2}$$

$$c^m c_x = u(xt^{-1/2})^m u'(xt^{-1/2})t^{-1/2}$$

$$(c^m c_x)_x = t^{-1/2}(u(xt^{-1/2})^m u'(xt^{-1/2}))_x =$$

$$= t^{-1/2}[mu(xt^{-1/2})^{m-1}u'(xt^{-1/2})t^{-1/2}u'(xt^{-1/2}) + u(xt^{-1/2})^m u''(xt^{-1/2})t^{-1/2}]$$

Con la notación $\eta = xt^{-1/2}$, la ecuación $c_t = (c^m c_x)_x$ se convierte en:

$$-\frac{1}{2}u'(\eta)\eta t^{-1} = t^{-1}[mu(\eta)^{m-1}u'(\eta)^2 + u(\eta)^m u''(\eta)]$$

es decir:

$$-\frac{1}{2}\eta u'(\eta) = (u^m(\eta)u'(\eta))'$$

Podemos considerar, entonces, que si $k(u) = u^m$ y $f(x) = -\frac{1}{2}x$ lo anterior equivale a $(k(u)u')' = f(x)u'$

5.2. Preliminares.

Ahora haremos suposiciones más precisas acerca de f y k en el caso general. Sea $I = [0, a]$ ($a > 0$) un intervalo. Suponemos que $f : I \rightarrow [0, +\infty)$ es una función diferenciable decreciente y continua tal que $f' < 0$ en I . Además, $k \in C([0, +\infty)) \cap C^1((0, +\infty))$ es tal que $k(0) = 0$, $k'(u) > 0$ para $u > 0$ y

$$\int_0^\delta \frac{k(u)}{u} du < \infty \quad (\delta > 0)$$

Estamos interesados en soluciones u del problema (5.1)–(5.2) tales que $u \in C(I) \cap C^2(0, a]$ y $u(x) > 0$ para $x > 0$.

Sea

$$K = \{u \in C(I) : u \geq 0 \text{ en } I\}$$

y

$$K_0 = \{u \in K : u(0) = 0, u > 0 \text{ en } (0, a]\}$$

Como en [23], [30] podemos demostrar

Lema 12. La función $u \in K_0$ es una solución de (5.1)–(5.2) si y sólo si:

$$u = T(u) \quad \text{en } I \quad (5.3)$$

donde

$$T(u) = W^{-1}(Lu), \quad (5.4)$$

$$Lu(x) = \int_0^x [f(s) - f'(s)(x-s)]u(s)ds \quad (5.5)$$

y W^{-1} es la función inversa de

$$W(x) = \int_0^x k(s)ds \quad (5.6)$$

Demostración. Para ver que $u = W^{-1}(Lu)$ veremos que $W(u(x)) = Lu(x)$, Si $x = 0$ se da la igualdad, veamos que coinciden las derivadas.

Podemos escribir $Lu(x)$ como sigue:

$$Lu(x) = \int_0^x f(s)u(s)ds - x \int_0^x f'(s)u(s) + \int_0^x s f'(s)u(s)ds$$

Su derivada será pues la siguiente:

$$\begin{aligned} (Lu)'(x) &= f(x)u(x) - 1 \cdot \int_0^x f'(s)u(s)ds - x(f'(x)u(x)) + x f'(x)u(x) \\ &= f(x)u(x) - \int_0^x f'(s)u(s)ds \end{aligned}$$

Derivando $W(u(x))$ se obtiene: $W'(u(x))u'(x) = k(u(x))u'(x)$ y queremos ver que eso es igual a lo anterior.

Como antes, veamos que coinciden las derivadas:

$$(k(u(x))u'(x))' = f(x)u'(x) = f'(x)u(x) + f(x)u'(x) - f'(x)u'(x)$$

□

Este último resultado nos permite considerar la ecuación (5.3) en lugar del problema original. Se puede demostrar también la siguiente estimación *a priori* ([23], [30]).

Lema 13. Si $u \in K_0$ es una solución de (5.3) entonces u es una función estrictamente creciente que satisface

$$\varphi(x) \leq u(x) \leq \Phi(x) \quad (5.7)$$

donde

$$\varphi(x) = V^{-1} \left(\int_0^x f(s)ds \right) \quad (5.8)$$

$$\Phi(x) = V^{-1}(f(0)x) \quad (5.9)$$

y V^{-1} denota la función inversa de

$$V(x) = \int_0^x \frac{k(s)}{s} ds \quad (5.10)$$

Demostración. Nos basaremos en los siguientes hechos. La ecuación (5.3) se puede escribir como

$$W(u(x)) = \int_0^x [f(s) - f'(s)(x-s)]u(s)ds \quad x \in [0, a]$$

Diferenciando se obtiene:

$$k(u(x))u'(x) = f(x)u(x) - \int_0^x f'(s)u(s)ds$$

para $x > 0$.

Dividiendo todo entre $u(x)$ se obtiene que:

$$\frac{k(u(x))u'(x)}{u(x)} = f(x) - \int_0^x \frac{f'(s)u(s)}{u(x)} ds \geq f(x)$$

La última desigualdad se debe a que $f'(x) < 0$ en I y $u(x) > 0$ para $x > 0$.

Por lo tanto lo siguiente es cierto:

$$f(x) \leq \frac{k(u(x))u'(x)}{u(x)} \leq f(x) - \int_0^x \frac{f'(s)u(s)}{u(x)} ds \quad (5.11)$$

Donde la última desigualdad es realmente una igualdad.

Como $f(x) > 0$ para $x > 0$, lo anterior implica que $u'(x) > 0$ para $x > 0$.

Además, como u es estrictamente creciente (y $f' < 0$), se tiene que:

$$f(x) - \int_0^x \frac{f'(s)u(s)}{u(x)} ds \leq f(x) - \int_0^x f'(s)ds = f(x) - [f(x) - f(0)] = f(0)$$

Se tiene por tanto la siguiente cadena de desigualdades:

$$f(x) \leq \frac{k(u(x))u'(x)}{u(x)} \leq f(0)$$

Si ahora integramos entre 0 y x sale lo siguiente:

$$\int_0^x f(s)ds \leq \int_0^x \frac{k(u(s))u'(s)}{u(s)} ds \leq \int_0^x f(0)dx = f(0)x$$

y aplicando V^{-1} a todo se obtiene (5.7), esto es:

$$\varphi(x) \leq u(x) \leq \Phi(x)$$

Que sale $\varphi(x)$ a la izquierda y $\Phi(x)$ a la derecha está bastante claro por la definición. Para ver que lo que sale en medio es $u(x)$ bastará ver que

$$\int_0^x \frac{k(u(s))u'(s)}{u(s)} ds = V(u(x))$$

Veamos, derivando sale que:

$$\frac{k(u(x))u'(x)}{u(x)} = V'(u(x))u'(x) = \frac{k(u(x))}{u(x)}u'(x)$$

luego efectivamente es cierto. \square

La estimación (5.7), en general, es muy pobre. Como veremos más adelante con la ayuda de (5.11) podemos obtener mejores estimaciones de u .

Antes de la formulación del siguiente lema, introducimos el siguiente operador:

$$Su \doteq Lu - W(u) \tag{5.12}$$

para $u \in K_0$

Lema 14. *Si φ y Φ están dados por (5.8) y (5.9) respectivamente, entonces:*

$$T(\varphi) \geq \varphi \tag{5.13}$$

y

$$T(\Phi) \leq \Phi \tag{5.14}$$

Demostración. Dado que

$$T(\varphi) = W^{-1}(L\varphi) \geq \varphi \iff L\varphi \geq W(\varphi)$$

para probar este lema es suficiente demostrar que

$$S(\varphi) \geq 0$$

y

$$S(\Phi) \leq 0$$

lo cual haremos por el siguiente razonamiento:

$$S(\varphi)(0) = S(\Phi)(0) = 0, \quad S'(\varphi) \geq 0 \text{ y } S'(\Phi) \leq 0.$$

Veamos:

$$S(\varphi)(0) = (L\varphi - W(\varphi))(0) = L\varphi(0) - W(\varphi(0)) = 0 - 0 = 0$$

y lo mismo pasa con $S(\Phi)(0)$.

Nota: En lo que sigue, se entiende, dado que S es un operador, que $S'(\varphi)$ es la función derivada de $S(\varphi)$.

Veamos ahora que $S'(\varphi) \geq 0$.

$$S\varphi = L\varphi - W(\varphi)$$

$$\begin{aligned} (L\varphi)(x) &= \int_0^x [f(s) - f'(s)(x-s)]\varphi(s)ds \\ &= \int_0^x f(s)\varphi(s)ds - x \int_0^x f'(s)\varphi(s)ds + \int_0^x sf'(s)\varphi(s)ds \end{aligned}$$

por lo tanto:

$$\begin{aligned} (L\varphi)'(x) &= f(x)\varphi(x) - 1 \cdot \int_0^x f'(s)\varphi(s)ds - x \cdot f'(x)\varphi(x) + xf'(x)\varphi(x) \\ &= f(x)\varphi(x) - \int_0^x f'(s)\varphi(s)ds \end{aligned}$$

Para derivar la segunda parte se aplica la regla de la cadena y la definición de W :

$$W(\varphi)'(x) = k(\varphi(x))\varphi'(x)$$

Por lo tanto, en definitiva, lo que hay que ver para ver que $S'(\varphi)(x) \geq 0$ es que:

$$f(x)\varphi(x) - \int_0^x f'(s)\varphi(s)ds - k(\varphi(x))\varphi'(x) \geq 0$$

La derivada de $\varphi(x)$ en realidad se puede calcular: Si

$$V(\varphi(x)) = \int_0^x f(s)ds$$

entonces, derivando:

$$V'(\varphi(x))\varphi'(x) = f(x)$$

y teniendo en cuenta la definición de V , resulta que $V'(\varphi(x)) = \frac{k(\varphi(x))}{\varphi(x)}$ y de aquí resulta que

$$\varphi'(x) = \frac{f(x)\varphi(x)}{k(\varphi(x))}$$

Resumiendo, para ver que $S'(\varphi(x)) \geq 0$ solamente hace falta ver que

$$- \int_0^x f'(s)\varphi(s)ds \geq 0$$

lo cual es cierto porque $f' < 0$ y $\varphi \geq 0$.

Para ver que $S'(\Phi) \leq 0$ hacemos lo mismo que antes y llegamos a que basta comprobar lo siguiente:

$$f(x)\Phi(x) - \int_0^x f'(s)\Phi(s)ds - k(\Phi(x))\Phi'(x) \leq 0$$

Para calcular la derivada de $\Phi(x)$, usamos la regla de la cadena en la relación $V(\Phi(x)) = f(0)x$ y se obtiene que:

$$V'(\Phi(x))\Phi'(x) = f(0) \Rightarrow \Phi'(x) = \frac{f(0)\Phi(x)}{k(\Phi(x))}$$

Sustituyendo se obtiene que lo que hay que comprobar es que:

$$f(x)\Phi(x) - \int_0^x f'(s)\Phi(s)ds - f(0)\Phi(x) \leq 0$$

y eso es cierto porque es una función que vale 0 en el 0, y su derivada vale:

$$\begin{aligned} f'(x)\Phi(x) + f(x)\Phi'(x) - f'(x)\Phi(x) - f(0)\Phi'(x) \\ = (f(x) - f(0))\Phi'(x) = \frac{(f(x) - f(0))f(0)\Phi(x)}{k(\Phi(x))} \leq 0 \end{aligned}$$

ya que f es decreciente, $\Phi \geq 0$ y $k \geq 0$. \square

A partir de este lema se obtiene muy fácilmente el siguiente:

Corolario 15. *Para todo $n \in \mathbf{N}$*

$$\varphi \leq T^n(\varphi) \leq T^{n+1}(\varphi) \leq u \leq T^{n+1}(\Phi) \leq T^n(\Phi) \leq \Phi \quad (5.15)$$

Nota previa a la demostración: Si $u_1, u_2 \in K$ son tales que $u_1 \leq u_2$ entonces $T(u_1) \leq T(u_2)$.

Demostración.

$$\varphi \leq u \Rightarrow T(\varphi) \leq T(u) = u$$

por otro lado:

$$\varphi \leq T(\varphi) \Rightarrow T(\varphi) \leq T^2(\varphi)$$

Repitiendo el proceso se obtiene el Lema:

$$\varphi \leq T^n(\varphi) \leq T^{n+1}(\varphi) \leq u$$

Del mismo modo se obtiene lo correspondiente para Φ a partir de $T(\Phi) \leq \Phi$. \square

Los teoremas presentados en [23], [30] garantizan la existencia y unicidad de soluciones no triviales de (5.3) y por tanto de (5.1)–(5.2). En este momento debemos hacer notar que en el caso $k(u) = u^m$ ($m > 0$) también podemos utilizar el método de la métrica proyectiva de Hilbert [9] para probar la existencia y unicidad de soluciones no triviales de la ecuación de operadores (5.3).

Teniendo todo esto en cuenta podemos presentar brevemente la idea de nuestro método de soluciones aproximadas.

5.3. La idea de la aproximación.

Presentamos los pasos principales de nuestra idea de aproximación de soluciones no triviales de (5.3).

(1) Es fácil de comprobar que también $S''(\varphi) \geq 0$ en I (Esto no es cierto para $S(\Phi)$).

Demostración. La última vez que lo calculamos obtuvimos que

$$S'(\varphi)(x) = - \int_0^x f'(s)\varphi(s)ds$$

Su derivada será por tanto:

$$S''(\varphi)(x) = -f'(x)\varphi(x) \geq 0 \quad \text{en } I$$

ya que $f' < 0$ y $\varphi \geq 0$ en I . □

Por lo tanto en alguna clase de funciones de K_0 que contenga a φ buscaremos una función $\underline{\varphi}$ que satisfaga:

$$S''(\underline{\varphi}) \geq 0 \tag{5.16}$$

tal que $\varphi < \underline{\varphi}$ para $x > 0$. Por el Corolario 15 obtenemos $\underline{\varphi} \leq u$.

(2) Usando $\underline{\varphi}$ intentaremos cambiar la estimación de u anterior para obtener, tal vez, una mejor. Para este fin, usaremos el lado derecho de (5.11). Sea $\underline{\Phi}$ una nueva estimación superior de u .

Denotaremos

$$P = \{u \in K_0 : \underline{\varphi} \leq u \leq \underline{\Phi}\}$$

(3) Suponemos que existe $w \in K_0$ tal que

$$\sup_{(0,a]} \frac{\underline{\Phi} - \underline{\varphi}}{w} < \infty \tag{5.17}$$

Si para algún $u \in C(I)$ tenemos que $\sup_{(0,a]} \frac{|u|}{w} < \infty$ entonces definimos $\|u\|_w \doteq \sup_{(0,a]} \frac{|u|}{w}$. Obviamente, si $u_1, u_2 \in P$ entonces $\|u_1 - u_2\|_w < \infty$.

(4) Suponemos la propiedad contractiva:

Si $u_1, u_2 \in P$, entonces $\|T(u_1) - T(u_2)\|_w < \infty$ y

$$\|T(u_1) - T(u_2)\|_w \leq \gamma \|u_1 - u_2\|_w \tag{5.18}$$

donde $\gamma \in (0, 1)$ es una constante fija independiente de u_1 y u_2 .

(5) Habiendo satisfecho todos los puntos anteriores, podemos probar el siguiente lema:

Lema 16. Si $p \in P$ satisface $S(p) \geq 0$ y

$$\left\| \frac{S(p)}{k(\underline{\varphi})} \right\|_w \leq \alpha < \infty$$

entonces

$$p \leq u \leq p + \frac{\alpha}{1-\gamma} w \quad \text{en } I \quad (5.19)$$

Demostración. Tenemos

$$\begin{aligned} \|u - p\|_w &= \|T(u) - p\|_w \\ &= \|T(u) - T(p) + T(p) - p\|_w \leq \|T(u) - T(p)\|_w + \|T(p) - p\|_w \end{aligned}$$

Por (4) obtenemos

$$\|u - p\|_w \leq \gamma \|u - p\|_w + \|T(p) - p\|_w \quad (5.20)$$

Como

$$|Tp - p| = |W^{-1}(Lp) - W^{-1}(W(p))| \quad (5.21)$$

entonces podemos aplicar el Teorema de los incrementos finitos de Lagrange y se obtiene:

$$W^{-1}(Lp)(x) - W^{-1}(W(p))(x) = (W^{-1})'(c(x))(Lp(x) - W(p(x)))$$

donde $W(p(x)) \leq c(x) \leq Lp(x)$. Teniendo en cuenta que $W' = k$ resulta que

$$(W^{-1})'(z) = \frac{1}{k(W^{-1}(z))}$$

y por lo tanto:

$$|Tp(x) - p(x)| = \left| \frac{Lp(x) - W(p(x))}{k(W^{-1}(c(x)))} \right|$$

Como k y W^{-1} son crecientes y $p \geq \underline{\varphi}$ se tiene que:

$$k(\underline{\varphi}(x)) \leq k(p(x)) = k(W^{-1}(W(p(x)))) \leq k(W^{-1}(c(x)))$$

así que de lo anterior resulta:

$$|Tp(x) - p(x)| \leq \left| \frac{Lp(x) - W(p(x))}{k(\underline{\varphi}(x))} \right|$$

Dividiendo entonces entre w y tomando supremos tenemos que:

$$\|Tp - p\|_w \leq \left\| \frac{Lp - W(p)}{k(\underline{\varphi})} \right\|_w = \left\| \frac{S(p)}{k(\underline{\varphi})} \right\|_w \leq \alpha \quad (5.22)$$

A partir de (5.20) y (5.22) se obtiene lo siguiente:

$$\|u - p\|_w \leq \gamma \|u - p\|_w + \alpha \Rightarrow (1 - \gamma) \|u - p\|_w \leq \alpha$$

y por lo tanto:

$$\|u - p\|_w \leq \frac{\alpha}{1 - \gamma} \quad (5.23)$$

Como $S(p) \geq 0$ entonces $T(p) \geq p$ y por el Corolario 15, $u \geq p$.

Por lo tanto a partir de (5.23) obtenemos:

$$0 \leq \frac{u(x) - p(x)}{w(x)} \leq \frac{\alpha}{1 - \gamma}$$

para $x > 0$, de donde sale lo que queríamos demostrar, esto es, que:

$$p \leq u \leq p + \frac{\alpha}{1 - \gamma} \omega$$

Con esto termina la demostración del lema. □

Corolario 17. *Si ponemos $p \equiv \underline{\varphi}$ entonces*

$$\underline{\varphi} \leq u \leq \underline{\varphi} + \frac{\alpha}{1 - \gamma} \omega \text{ en } I \quad (5.24)$$

(6) Sea $k_1 \leq k \leq k_2$, $f_1 \leq f \leq f_2$ y $f'_1 \geq f' \geq f'_2$. Aplicando el Corolario 17 obtenemos

$$\underline{\varphi}_1 \leq u \leq \underline{\varphi}_2 + \frac{\alpha_2}{1 - \gamma_2} \omega_2, \quad (5.25)$$

donde $\underline{\varphi}_i$, w_i , γ_i , α_i ($i = 1, 2$) son funciones apropiadas descritas en los puntos anteriores.

En el siguiente apartado aplicaremos esta idea a un caso más particular del problema considerado.

5.4. El problema particular.

Consideraremos el problema (5.1)–(5.2) suponiendo adicionalmente que

$$c_1 u^m \leq k(u) \leq c_2 u^m$$

donde $0 < c_1 \leq c_2 < \infty$ y $m > 0$. Esta suposición es típica para las aplicaciones [23], [30], [39], [44]. Podemos escribir la estimación

$$f_1(x) \leq f(x) \leq f_2(x), \quad x \in [0, a]$$

donde

$$\begin{aligned} f_i(x) &= A_i(a-x) + f(a) \\ A_1 &= - \max_{x \in [0,a]} f'(x) \\ A_2 &= - \min_{x \in [0,a]} f'(x) \end{aligned}$$

Comentario: De dónde sale que $f_1(x) \leq f(x) \leq f_2(x)$

Aplicando el Teorema de los incrementos finitos de Lagrange se tiene que si $x \in [0, a]$ entonces $f(a) - f(x) = f'(c)(a-x)$. Dado que $-A_2 \leq f'(c) \leq -A_1$ se tiene que:

$$-A_2(a-x) \leq f(a) - f(x) \leq -A_1(a-x)$$

Despejando $f(x)$ se obtiene entonces que:

$$f_1(x) = A_1(a-x) + f(a) \leq f(x) \leq A_2(a-x) + f(a) = f_2(x)$$

Sea $k_i(u) = c_i u^m$ ($i = 1, 2$). Para los f_i y k_i definidos, la ecuación (5.3) toma la forma

$$u_i(x) = \left[\frac{1}{c_i(m+1)} \int_0^x [A_i a + f_i(a) + A_i(x-2s)] u_i(s) ds \right]^{1/(m+1)} \quad x \in [0, a] \quad (5.26)$$

Se definen $k_i(u) = c_i u^m$ para $i = 1, 2$, veamos cómo queda la ecuación $T(u) = u$ para esos k_i , f_i , etc.

$$W_i(x) = \int_0^x k_i(s) ds = \int_0^x c_i s^m ds = c_i \frac{x^{m+1}}{m+1}$$

La inversa será pues:

$$W_i^{-1}(x) = \left(\frac{(m+1)x}{c_i} \right)^{1/(m+1)}$$

Por otro lado:

$$L u_i(x) = \int_0^x [f_i(s) - f'_i(s)(x-s)] u_i(s) ds = \int_0^x [A_i(a-s) + f(a) + A_i(x-s)] u_i(s) ds$$

Juntándolo todo sale que:

$$u_i(x) = \left(\frac{m+1}{c_i} \int_0^x [A_i a + f(a) + A_i(x-2s)] u_i(s) ds \right)^{1/(m+1)}$$

Sustituyendo

$$\left[\frac{(A_i a + f_i(a))^2}{A_i c_i (m+1)} \right]^{1/m} u \left(\frac{A_i x}{A_i a + f(a)} \right)$$

en lugar de u_i ($i = 1, 2$) podemos reducir la ecuación (5.26) a la siguiente:

$$u(x) = \left[\int_0^x (1 + x - 2s)u(s)ds \right]^{1/(m+1)} \quad x \in [0, 1] \quad (5.27)$$

Ahora estudiaremos soluciones aproximadas de (5.27)

5.5. Estimaciones de soluciones no triviales de (5.27)

De acuerdo con el Lema 13 podemos formular:

Corolario 18. Si $u_0 \in K_0$ es la solución no trivial de (5.27) entonces:

$$\left(m \left(x - \frac{x^2}{2} \right) \right)^{1/m} \leq u(x) \leq (mx)^{1/m} \quad (5.28)$$

para $x \in [0, 1]$.

Demostración. Tomando $f(x) = 1 - x$ y $k(x) = x^m$ sale lo siguiente:

$$\begin{aligned} V(x) &= \int_0^x \frac{s^m}{s} ds = \int_0^x s^{m-1} ds = \left[\frac{s^m}{m} \right]_0^x = \frac{x^m}{m} \\ &\Rightarrow V^{-1}(x) = (mx)^{1/m} \\ \Phi(x) &= V^{-1}(f(0)x) = (mx)^{1/m} \\ \int_0^x f(s)ds &= \int_0^x (1-s)ds = \left[s - \frac{s^2}{2} \right]_0^x = x - \frac{x^2}{2} \\ \varphi(x) &= V^{-1} \left(\int_0^x f(s)ds \right) = \left(m \left(x - \frac{x^2}{2} \right) \right)^{1/m} \end{aligned}$$

Luego efectivamente se cumple que:

$$\left(m \left(x - \frac{x^2}{2} \right) \right)^{1/m} \leq u(x) \leq (mx)^{1/m}$$

Faltaría ver, para que cuadre todo, que

$$u(x) = \left[\int_0^x (1 + x - 2s)u(s)ds \right]^{1/(m+1)} = W^{-1}(Lu)$$

Veamos:

$$\begin{aligned} (Lu)(x) &= \int_0^x [f(s) - f'(s)(x-s)]u(s)ds \\ &= \int_0^x (1-s - (-1)(x-s))u(s)ds = \int_0^x (1+x-2s)u(s)ds \end{aligned}$$

$$W(x) = \int_0^x k(s)ds = \int_0^x s^m ds = \frac{x^{m+1}}{m+1}$$

$$W^{-1}(Lu) = \left((m+1) \int_0^x (1+x-2s)u(s)ds \right)^{1/(m+1)}$$

□

Pero esta estimación es bastante pobre para las aplicaciones. Nos gustaría encontrar estimaciones mejores. Usando la condición (5.16) intentaremos encontrar una estimación mejor de u a partir de lo anterior. Formulamos el siguiente lema basado en cálculos sencillos:

Lema 19. Si $\varphi_0(x) = [m(x - d_1x^2 + d_2x^3)]^{1/m}$ ($d_1 \leq \frac{1}{2}, d_2 \geq 0$) entonces

$$\begin{aligned} S''(\varphi_0)(x) &= \varphi_0^{1-m}(x)x[2d_1(m+1) - 1 \\ &\quad - [2d_1(d_1(m+2) - 1) + 3d_2(1+2m)]x \\ &\quad + [4d_1(2m+3) - 3]d_2x^2 - 9d_2^2x^3] \end{aligned} \quad (5.29)$$

Comentario: Se trata de encontrar una función sencilla que verifique la relación $\varphi(x) \leq \varphi_0(x) \leq u(x)$. Dado que $\varphi(x)^m$ es un polinomio de grado 2, es lógico buscar φ_0 como raíz m de un polinomio que extienda al anterior.

Demostración. Calculemos cuánto vale $S''(\varphi_0)$. Como antes, vamos a suponer que estamos en el caso en el que $f(x) = 1 - x$, $k(x) = x^m$, se tiene entonces que $f'(x) = -1$.

Dado que

$$S(\varphi_0) = L\varphi_0 - W(\varphi_0)$$

vamos a calcular primero $L\varphi_0$.

$$\begin{aligned} L\varphi_0(x) &= \int_0^x [f(s) - f'(s)(x-s)]\varphi_0(s)ds \\ &= \int_0^x [1-s + (x-s)]\varphi_0(s)ds = \int_0^x (1+x-2s)\varphi_0(s)ds \end{aligned}$$

es decir:

$$L\varphi_0(x) = (1+x) \int_0^x \varphi_0(s)ds - 2 \int_0^x s\varphi_0(s)ds$$

Por otro lado:

$$\begin{aligned} S'(\varphi_0)(x) &= (L\varphi_0)'(x) - W'(\varphi_0(x))\varphi_0'(x) \\ &= 1 \cdot \int_0^x \varphi_0(s)ds + (1+x)\varphi_0(x) - 2x\varphi_0(x) - k(\varphi_0(x))\varphi_0'(x) \end{aligned}$$

es decir:

$$S'(\varphi_0)(x) = \int_0^x \varphi_0(s) ds + (1-x)\varphi_0(x) - k(\varphi_0(x))\varphi_0'(x)$$

Ya se puede calcular la segunda derivada:

$$S''(\varphi_0)(x) = \varphi_0(x) + (-1)\varphi_0(x) + (1-x)\varphi_0'(x) - [k'(\varphi_0(x))\varphi_0'(x)^2 + k(\varphi_0(x))\varphi_0''(x)]$$

O sea:

$$S''(\varphi_0)(x) = (1-x)\varphi_0'(x) - k'(\varphi_0(x))\varphi_0'(x)^2 - k(\varphi_0(x))\varphi_0''(x)$$

En este caso, $k(x) = x^m$, por lo que $k'(x) = mx^{m-1}$. Calculemos primero la derivada de φ_0 :

$$\varphi_0(x) = [m(x - d_1x^2 + d_2x^3)]^{1/m} = [m \cdot p(x)]^{1/m}$$

donde $p(x) = x - d_1x^2 + d_2x^3$, por lo que

$$\varphi_0'(x) = \frac{1}{m} [mp(x)]^{\frac{1}{m}-1} mp'(x) = [mp(x)]^{\frac{1}{m}-1} p'(x)$$

$$\begin{aligned} \varphi_0''(x) &= \left(\frac{1}{m} - 1 \right) [mp(x)]^{\frac{1}{m}-2} mp'(x)p'(x) + [mp(x)]^{\frac{1}{m}-1} p''(x) \\ &= (1-m)[mp(x)]^{\frac{1}{m}-2} p'(x)^2 + [mp(x)]^{\frac{1}{m}-1} p''(x) \end{aligned}$$

Sustituyendo cada cosa por lo que vale, se obtiene:

$$\begin{aligned} S''(\varphi_0)(x) &= (1-x)[mp(x)]^{\frac{1}{m}-1} p'(x) - m[mp(x)]^{\frac{m-1}{m}} \left[[mp(x)]^{\frac{1}{m}-1} p'(x) \right]^2 \\ &\quad - mp(x) \left[(1-m)[mp(x)]^{\frac{1}{m}-2} p'(x)^2 + [mp(x)]^{\frac{1}{m}-1} p''(x) \right] = \\ &= [mp(x)]^{\frac{1}{m}-1} \left[(1-x)p'(x) - mp'(x)^2 - (1-m)p'(x)^2 - mp(x)p''(x) \right] = \\ &= [mp(x)]^{\frac{1}{m}-1} \left[(1-x)p'(x) - p'(x)^2 - mp(x)p''(x) \right] = \\ &= [mp(x)]^{\frac{1}{m}-1} \left[p'(x)(1-x-p'(x)) - mp(x)p''(x) \right] \end{aligned}$$

Ya es hora de sustituir cada cosa por su valor:

$$p(x) = x - d_1x^2 + d_2x^3$$

$$p'(x) = 1 - 2d_1x + 3d_2x^2$$

$$p''(x) = -2d_1 + 6d_2x$$

Podemos entonces escribir:

$$S''(\varphi_0)(x) = [mp(x)]^{\frac{1}{m}-1} q(x)$$

donde

$$q(x) = (1 - 2d_1x + 3d_2x^2)(1 - x - (1 - 2d_1x + 3d_2x^2)) - m(x - d_1x^2 + d_2x^3)(-2d_1 + 6d_2x)$$

es decir:

$$\begin{aligned} q(x) &= (1 - 2d_1x + 3d_2x^2)(-x + 2d_1x - 3d_2x^2) - m(x - d_1x^2 + d_2x^3)(-2d_1 + 6d_2x) = \\ &= -x + 2d_1x - 3d_2x^2 + 2d_1x^2 - 4d_1^2x^2 + 6d_1d_2x^3 - 3d_2x^3 + 6d_1d_2x^3 - 9d_2^2x^4 \\ &\quad + 2md_1x - 6md_2x^2 - 2md_1^2x^2 + 6md_1d_2x^3 + 2md_1d_2x^3 - 6md_2^2x^4 = \\ &= (-1 + 2d_1 + 2md_1)x + (-3d_2 + 2d_1 - 4d_1^2 - 6md_2 - 2md_1^2)x^2 + \\ &\quad (6d_1d_2 - 3d_2 + 6d_1d_2 + 6md_1d_2 + 2md_1d_2)x^3 + (-9d_2^2 - 6md_2^2)x^4 = \\ &= [2d_1(m + 1) - 1]x + [2d_1(1 - 2d_1 - d_1m) - 3d_2(1 + 2m)]x^2 + \\ &\quad [d_2(12d_1 - 3 + 8md_1)]x^3 + [3d_2^2(-3 - 2m)]x^4 = \\ &= [2d_1(m + 1) - 1]x + [2d_1(-d_1(m + 2) + 1) - 3d_2(1 + 2m)]x^2 + \\ &\quad [d_2(4d_1(3 + 2m) - 3)]x^3 + [3d_2^2(-3 - 2m)]x^4 \end{aligned}$$

En resumen:

$$\begin{aligned} S''(\varphi_0)(x) &= \varphi_0^{1-m}(x)x[2d_1(m + 1) - 1 - [2d_1(d_1(m + 2) - 1) + 3d_2(2m + 1)]x \\ &\quad + [4d_1(2m + 3) - 3]d_2x^2 - 3d_2^2(2m + 3)x^3] \end{aligned}$$

□

Usando el último lema podemos escribir

Lema 20. Sea \underline{P} la clase de funciones de la forma

$$[m(x - d_1x^2)]^{1/m} \quad \text{con } d_1 \leq \frac{1}{2}.$$

Si $\varphi_0 \in \underline{P}$ es tal que $S''(\varphi_0) \geq 0$ en $[0, 1]$ entonces

$$\varphi_0(x) \leq \left[m \left(x - \frac{1}{2(m+1)}x^2 \right) \right]^{1/m} \doteq \underline{\varphi}_1(x) \quad (5.30)$$

Demostración. Haciendo $d_2 = 0$ en (5.29) se obtiene:

$$S''(\varphi_0)(x) = \varphi_0^{1-m}(x)x[2d_1(m + 1) - 1 - (2d_1(d_1(m + 2) - 1))x] \geq 0$$

Dado que $\varphi_0(x)^{1-m}x > 0$ para $x > 0$, lo anterior se traduce en que:

$$2d_1(m + 1) - 1 \geq [2d_1(d_1(m + 2) - 1)]x \quad x \in (0, 1]$$

Cuando $x \rightarrow 0$ se obtiene que

$$2d_1(m+1) - 1 \geq 0 \Rightarrow 2d_1(m+1) \geq 1 \Rightarrow d_1 \geq \frac{1}{2(m+1)} \Rightarrow -d_1 \leq -\frac{1}{2(m+1)}$$

De lo anterior sale:

$$m(x - d_1x^2) \leq m \left(x - \frac{1}{2(m+1)}x^2 \right)$$

lo cual implica

$$\varphi_0(x) = [m(x - d_1x^2)]^{\frac{1}{m}} \leq \left[m \left(x - \frac{1}{2(m+1)}x^2 \right) \right]^{\frac{1}{m}} = \underline{\varphi}_1(x)$$

con lo que el lema queda demostrado. \square

A partir de ahora denotaremos

$$d_1(m) \doteq \frac{1}{2(m+1)} \quad (5.31)$$

Con la ayuda de los resultados del Lema 20 podemos formular el siguiente:

Lema 21. Sea \underline{P} la clase de funciones de la forma

$$[m(x - d_1(m)x^2 + d_2x^3)]^{1/m} \quad \text{con } d_2 \geq 0,$$

donde $d_1(m) = \frac{1}{2(m+1)}$. Si $\varphi_0 \in \underline{P}$ es tal que $S''(\varphi_0) \geq 0$ en $[0, 1]$ entonces

$$\varphi_0(x) \leq \left[m \left(x - d_1(m)x^2 + \frac{m}{6(m+1)^2(2m+1)}x^3 \right) \right]^{1/m} \doteq \underline{\varphi}(x) \quad (5.32)$$

Demostración. En el Lema 19, sustituyamos d_1 por su valor y veamos qué significa en tal caso que $S''(\varphi_0) \geq 0$.

Como antes, podemos descartar el factor $\varphi_0^{1-m}(x)x$ ya que no aporta nada. Se tiene entonces lo siguiente:

$$2 \cdot \frac{1}{2(m+1)} \cdot (m+1) - 1 - \left[2 \cdot \frac{1}{2(m+1)} \left(\frac{m+2}{2(m+1)} - 1 \right) + 3d_2(2m+1) \right] x \\ + \left[4 \frac{2m+3}{2(m+1)} - 3 \right] d_2x^2 - 3d_2^2(2m+3)x^3 \geq 0$$

Dado que el término independiente de lo anterior desaparece, se puede extraer de nuevo un factor x que no aporta nada. Teniendo en cuenta que el resultado es mayor o igual que cero para x en $(0, 1]$, podemos considerar que también es cierto cuando $x \rightarrow 0$, con lo que se obtiene lo siguiente:

$$- \left[\frac{1}{m+1} \left(\frac{m+2}{2(m+1)} - \frac{2m+2}{2(m+1)} \right) + 3d_2(2m+1) \right] \geq 0$$

es decir:

$$\begin{aligned} \frac{m}{2(m+1)^2} - 3d_2(2m+1) &\geq 0 \\ \Rightarrow \frac{m}{2(m+1)^2} &\geq 3d_2(2m+1) \\ \Rightarrow d_2 &\leq \frac{m}{6(m+1)^2(2m+1)} \end{aligned}$$

Lo que queremos ver es que

$$\varphi_0(x) \leq \left[m \left(x - d_1(m)x^2 + \frac{m}{6(m+1)^2(2m+1)}x^3 \right) \right]^{\frac{1}{m}}$$

es decir, elevando a m y quitando lo que sobra, queremos ver que:

$$d_2x^3 \leq \frac{m}{6(m+1)^2(2m+1)}x^3$$

lo cual es inmediatamente cierto a partir de la última desigualdad obtenida antes. \square

A partir de aquí denotaremos

$$d_2(m) \doteq \frac{m}{6(m+1)^2(2m+1)} \quad (5.33)$$

Corolario 22. De los Lemas 20 y 21 se infiere

$$\left[m \left(x - \frac{x^2}{2} \right) \right]^{1/m} \leq \underline{\varphi}_1(x) \leq \underline{\varphi}(x) \leq u(x) \quad (5.34)$$

Demostración. Veamos: La primera desigualdad está clara porque

$$x - \frac{1}{2}x^2 \leq x - \frac{1}{2(m+1)}x^2$$

La segunda está clara porque

$$x - d_1(m)x^2 \leq x - d_1(m)x^2 + d_2(m)x^3$$

\square

Tomaremos la función $\underline{\varphi}$ como una nueva estimación de u por debajo. Ahora damos una nueva estimación de u a partir de lo anterior.

Lema 23. Si $u \in K_0$ es la solución no trivial de (5.27) entonces

$$u(x) \leq \underline{\Phi}(x) \quad \text{para } x \in [0, 1], \quad (5.35)$$

donde

$$\underline{\Phi}(x) \doteq \left[m \left(x - d_1(m)x^2 + \frac{2^{1/m}}{3(m+1)(2m+1)}x^3 \right) \right]^{1/m}$$

Demostración. La ecuación (5.11) dice

$$\frac{k(u(x))u'(x)}{u(x)} \leq f(x) - \int_0^x \frac{f'(s)u(s)}{u(x)} ds$$

En nuestro caso, $k(s) = s^m$ y $f(s) = 1 - s$, por lo que lo anterior se convierte en:

$$u^{m-1}(x)u'(x) \leq 1 - x + \int_0^x \frac{u(s)}{u(x)} ds \quad x \in (0, 1]$$

Teniendo en cuenta que $u(x) \leq (mx)^{1/m}$ sale que:

$$u^{m-1}(x)u'(x) \leq 1 - x + \int_0^x \frac{(ms)^{1/m}}{u(x)} ds = 1 - x + \frac{1}{u(x)} \int_0^x (ms)^{1/m} ds =$$

$$1 - x + \frac{1}{u(x)} \left(m^{1/m} \frac{x^{1+1/m}}{\frac{1}{m} + 1} \right) = 1 - x + \frac{m^{1/m}}{u(x)} \frac{m}{m+1} x^{1+1/m}$$

Aprovechamos ahora que $\varphi_1(x) \leq u(x)$ (lo cual estaba por ver) y obtenemos que lo anterior es menor o igual que:

$$1 - x + \frac{m^{1/m}}{\left[m \left(x - \frac{1}{2(m+1)} x^2 \right) \right]^{1/m}} \frac{m}{m+1} = 1 - x + \frac{m}{m+1} x \left(1 - \frac{x}{2(m+1)} \right)^{-\frac{1}{m}}$$

Es decir:

$$u^{m-1}(x)u'(x) \leq 1 - x + \frac{m}{m+1} x \left(1 - \frac{x}{2(m+1)} \right)^{-1/m} \quad (5.36)$$

Ahora usamos que

$$\left(1 - \frac{x}{2(m+1)} \right)^{-1/m} \leq 1 + \frac{2^{1/m}}{m(2m+1)} x \quad \text{para } x \in [0, 1]$$

cosa que demostraremos después. Se tiene entonces que lo anterior es menor o igual que:

$$1 - x + \frac{m}{m+1} x \left(1 + \frac{2^{1/m}}{m(2m+1)} x \right) = 1 - x + \frac{m}{m+1} x + \frac{2^{1/m}}{(m+1)(2m+1)} x^2$$

Por lo tanto:

$$u^{m-1}(x)u'(x) \leq 1 - \frac{1}{m+1} x + \frac{2^{1/m}}{(m+1)(2m+1)} x^2$$

Ahora integramos esta desigualdad y sale lo siguiente:

$$\int_0^x u^{m-1}(s)u'(s) ds = \frac{u^m(x)}{m} \leq x - \frac{1}{2(m+1)} x^2 + \frac{2^{1/m}}{3(m+1)(2m+1)} x^3$$

Por lo tanto:

$$u(x) \leq \left[m \left(x - \frac{1}{2(m+1)}x^2 + \frac{2^{1/m}}{3(m+1)(2m+1)}x^3 \right) \right]^{\frac{1}{m}} = \underline{\Phi}(x)$$

Como queríamos demostrar.

Demostremos ahora eso que dejamos pendiente:

$$\left(1 - \frac{x}{2(m+1)} \right)^{-\frac{1}{m}} \leq 1 + \frac{2^{\frac{1}{m}}}{m(2m+1)}x$$

Basta ver que la función

$$f(x) = \left(1 - \frac{x}{2(m+1)} \right)^{-\frac{1}{m}} - 1 - \frac{2^{\frac{1}{m}}}{m(2m+1)}x \leq 0$$

Dado que $f(0) = 0$, será suficiente ver que $f'(x) \leq 0$ para $x \in [0, 1]$.

Calculando pues la derivada, queremos ver que:

$$-\frac{1}{m} \left(1 - \frac{x}{2(m+1)} \right)^{-\frac{1}{m}-1} \left(-\frac{1}{2(m+1)} \right) - \frac{2^{\frac{1}{m}}}{m(2m+1)} \leq 0$$

Multiplicando por m , será suficiente comprobar que:

$$\frac{1}{2(m+1)} \left(1 - \frac{x}{2(m+1)} \right)^{-\frac{1}{m}-1} - \frac{2^{\frac{1}{m}}}{(2m+1)} \leq 0$$

Dado que se trata de una función creciente, bastará ver que se verifica para $x = 1$, es decir:

$$\frac{1}{2(m+1)} \left(1 - \frac{1}{2(m+1)} \right)^{-\frac{1}{m}-1} - \frac{2^{\frac{1}{m}}}{(2m+1)} \leq 0$$

O sea:

$$\frac{1}{2m+1} \cdot \frac{2m+1}{2m+2} \left(\frac{2m+1}{2m+2} \right)^{-\frac{1}{m}-1} - \frac{2^{\frac{1}{m}}}{(2m+1)} \leq 0$$

lo cual equivale, multiplicando todo por $2m+1$ a que:

$$\left(\frac{2m+1}{2m+2} \right)^{-\frac{1}{m}} \leq 2^{\frac{1}{m}}$$

es decir, elevando a m :

$$\frac{2m+2}{2m+1} \leq 2$$

lo cual es cierto ya que $2m+2 \leq 4m+2$. □

Comentario 24. La estimación $u(x) \leq \underline{\Phi}(x)$ es mejor solamente para $m \geq 1$ pero tiene una forma más apropiada para nuestras consideraciones posteriores.

A partir de aquí consideraremos $P = \{u \in K_0 : \underline{\varphi} \leq u \leq \underline{\Phi}\}$ donde $\underline{\varphi}$ está definido por (5.32) y $\underline{\Phi}$ por (5.35).

5.6. La propiedad contractiva del operador T .

Tal y como se ha indicado en el punto 3 de la idea de la aproximación debemos tratar de encontrar una función w tal que (5.17) se satisfice. Teniendo en cuenta la forma de $\underline{\Phi}$ y $\underline{\varphi}$ intentaremos considerar por ejemplo las siguientes funciones w :

$$\begin{aligned} w_1(x) &= x^{2+1/m} \quad (m > 0) \\ w_2(x) &= x^{1/m} \quad (m > 0) \\ w_3(x) &= x^{3/m} \quad (m > 1) \end{aligned}$$

Es fácil de ver que para estas funciones, (5.17) se satisfice, es decir:

$$\sup_{(0,a]} \frac{\underline{\Phi} - \underline{\varphi}}{w} < \infty$$

Veámoslo: Dado que se trata de funciones continuas, bastará ver que el límite cuando x tiende a cero es cero.

Veamos primero que es cierto con w_1 . Las otras dos pueden deducirse fácilmente una vez que se ha hecho la primera.

Para simplificar la notación, escribimos:

$$\begin{aligned} \underline{\varphi}(x) &= [mxp(x)]^{1/m} \\ \underline{\Phi}(x) &= [mxq(x)]^{1/m} \end{aligned}$$

donde

$$p(x) = 1 - d_1(m)x + d_2(m)x^2$$

y

$$q(x) = 1 - d_1(m)x + \frac{2^{1/m}}{3(m+1)(2m+1)}x^2$$

Se tiene entonces que

$$\lim_{x \rightarrow 0} \frac{\underline{\Phi}(x) - \underline{\varphi}(x)}{w_1(x)} = \lim_{x \rightarrow 0} \frac{[mxq(x)]^{1/m} - [mxp(x)]^{1/m}}{x^{2+1/m}}$$

Simplificando un poco, eso es igual a:

$$\lim_{x \rightarrow 0} \frac{[mq(x)]^{1/m} - [mp(x)]^{1/m}}{x^2}$$

Al ser cero tanto los límites del numerador y del denominador, podemos aplicar la

regla de L'Hopital:

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{[mq(x)]^{1/m} - [mp(x)]^{1/m}}{x^2} \\ &= \lim_{x \rightarrow 0} \frac{\frac{1}{m}[mq(x)]^{\frac{1}{m}-1}(mq'(x)) - \frac{1}{m}[mp(x)]^{\frac{1}{m}-1}(mp'(x))}{2x} \\ &= \lim_{x \rightarrow 0} \frac{[mq(x)]^{\frac{1}{m}-1}q'(x) - [mp(x)]^{\frac{1}{m}-1}p'(x)}{2x} \end{aligned}$$

pero volvemos a tener cero como límite de numerador y denominador, ya que $p'(0) = q'(0) = 0$, así que volvemos a aplicar la regla de L'Hopital y sale 2 en el denominador, y en el numerador lo siguiente:

$$\begin{aligned} \left(\frac{1}{m} - 1\right) [mq(x)]^{\frac{1}{m}-2}mq'(x)^2 + [mq(x)]^{\frac{1}{m}-1}q''(x) \\ - \left(\frac{1}{m} - 1\right) [mp(x)]^{\frac{1}{m}-2}mp'(x)^2 - [mp(x)]^{\frac{1}{m}-1}p''(x) \end{aligned}$$

Por lo tanto el límite sale igual a

$$\frac{m^{1/m-1}(q''(0) - p''(0))}{2}$$

que es finito, como queríamos demostrar.

Teniendo en cuenta que $w_1(x) = x^{2+1/m} = x^2x^{1/m} = x^2w_2(x)$ se obtiene inmediatamente que

$$\lim_{x \rightarrow 0} \frac{\Phi(x) - \varphi(x)}{w_2(x)} = 0 < \infty$$

Por último, teniendo en cuenta que cuando $x \rightarrow 0$

$$x^{2+1/m} \leq x^{3/m} \iff \frac{3}{m} \leq 2 + \frac{1}{m} \iff 3 \leq 2m + 1 \iff 2 \leq 2m \iff 1 \leq m$$

se obtiene análogo resultado para w_3 cuando $m \geq 1$

Obviamente es posible encontrar otra w pero, en nuestra opinion, la presentadas anteriormente son las más naturales con respecto a las φ y Φ consideradas. Nos gustaría elegir, si es posible, la mejor w con respecto a nuestras aproximaciones.

Ahora intentaremos comprobar si T tiene la propiedad contractiva con respecto a las distancias inducidas por las funciones w consideradas. Para ello formulamos primero el siguiente lema:

Lema 25. *Si L está dado por (5.5) entonces*

$$L(w_i)(x) \leq \lambda_i \varphi_1^m(x) w_i(x) \quad \text{para } x \in [0, 1] \quad (i = 1, 2, 3), \quad (5.37)$$

donde

$$\begin{aligned}\lambda_1 &= \frac{1}{3m+1} \\ \lambda_2 &= \frac{1}{m+1} \\ \lambda_3 &= \frac{1}{m+3} \quad (m > 1)\end{aligned}\tag{5.38}$$

Demostración. Sabemos, porque lo hemos hecho antes, que

$$Lw(x) = (1+x) \int_0^x w(s)ds - 2 \int_0^x sw(s)ds$$

Para $w_1(x) = x^{2+1/m}$ sale lo siguiente:

$$Lw_1(x) = (1+x) \int_0^x s^{2+1/m}ds - 2 \int_0^x s^{3+1/m}ds$$

es decir, queremos ver que:

$$(1+x) \cdot \frac{x^{3+1/m}}{3+\frac{1}{m}} - 2 \cdot \frac{x^{4+1/m}}{4+\frac{1}{m}} \leq \frac{1}{3m+1} \left[m \left(x - \frac{1}{2(m+1)}x^2 \right) \right] x^{2+1/m}$$

Dividiendo todo entre $x^{2+1/m}$ y multiplicando por $\frac{3m+1}{m}$ sale que:

$$(1+x)x - \frac{2(3m+1)}{4m+1}x^2 \leq x - \frac{1}{2(m+1)}x^2$$

Restando x de ambos miembros de la igualdad y dividiendo entre x^2 se obtiene que:

$$1 + \frac{1}{2(m+1)} = \frac{2m+3}{2(m+1)} \leq \frac{2(3m+1)}{4m+1}$$

En resumidas cuentas, queremos ver que:

$$\begin{aligned}(2m+3)(4m+1) &\leq 4(m+1)(3m+1) \\ \iff 8m^2 + 14m + 3 &\leq 12m^2 + 16m + 4 \\ &\iff 4m^2 + 2m + 1 \geq 0\end{aligned}$$

Lo cual es cierto.

Repitiendo lo mismo de antes con w_2 sale que:

$$(1+x) \cdot \frac{x^{1+1/m}}{1+\frac{1}{m}} - 2 \cdot \frac{x^{2+1/m}}{2+\frac{1}{m}} \leq \frac{1}{m+1} \left[m \left(x - \frac{1}{2(m+1)}x^2 \right) \right] x^{1/m}$$

Casi igual que antes, dividiendo todo entre $x^{1/m}$ y multiplicando por $\frac{m+1}{m}$ sale que:

$$x + x^2 - \frac{2(m+1)}{2m+1}x^2 \leq x - \frac{1}{2(m+1)}x^2$$

lo cual, como antes, equivale a:

$$1 + \frac{1}{2(m+1)} = \frac{2m+3}{2(m+1)} \leq \frac{2(m+1)}{2m+1}$$

O sea:

$$(2m+3)(2m+1) \leq 4(m+1)^2 \iff 4m^2 + 8m + 3 \leq 4m^2 + 8m + 4 \iff 0 \leq 1$$

Lo cual es cierto.

Por último, tomando $w_3(x) = x^{3/m}$ resulta:

$$(1+x) \cdot \frac{x^{1+3/m}}{1+\frac{3}{m}} - 2 \cdot \frac{x^{2+3/m}}{2+\frac{3}{m}} \leq \frac{1}{m+3} \left[m \left(x - \frac{1}{2(m+1)}x^2 \right) \right] x^{3/m}$$

Como de costumbre, dividiendo todo entre $x^{3/m}$ y multiplicando por $\frac{m+3}{m}$ sale que:

$$x + x^2 - \frac{2(m+3)}{2m+3}x^2 \leq x - \frac{1}{2(m+1)}x^2$$

es decir:

$$1 + \frac{1}{2(m+1)} = \frac{2m+3}{2(m+1)} \leq \frac{2(m+3)}{2m+3}$$

Es decir:

$$(2m+3)^2 \leq 4(m+1)(m+3) \iff 4m^2 + 12m + 9 \leq 4m^2 + 16m + 12 \iff 4m+3 \geq 0$$

lo cual es cierto, incluso si $m \leq 1$. \square

Ahora podemos formular

Lema 26. Si $u_1, u_2 \in P$ entonces $\|T(u_1) - T(u_2)\|_{w_i} < \infty$ ($i = 1, 2, 3$) y

$$\|T(u_1) - T(u_2)\|_{w_i} \leq \lambda_i \|u_1 - u_2\|_{w_i} \quad (5.39)$$

Demostración. Dado que

$$W(x) = \int_0^x k(s) ds = \frac{x^{m+1}}{m+1}$$

se tiene que $T(u) = W^{-1}(Lu) = [(m+1)Lu]^{1/(m+1)}$.

Teniendo en cuenta que $f(x) = 1 - x$, como antes, sale que:

$$|(Lu_1)(x) - (Lu_2)(x)| = \left| \int_0^x (1 + x - 2s)(u_1(s) - u_2(s)) ds \right|$$

Como $x \in [0, 1]$, dentro de la integral tenemos que $s \leq x \leq 1$, por lo tanto resulta que $1 + x - 2s \geq 0$. Así pues, la integral anterior es menor o igual que:

$$\dots \leq \int_0^x (1 + x - 2s) \cdot |u_1(s) - u_2(s)| ds = \int_0^x (1 + x - 2s) w_i(s) \frac{|u_1(s) - u_2(s)|}{w_i(s)} ds$$

Por la definición de la norma se tiene que

$$\frac{|u_1(s) - u_2(s)|}{w_i(s)} \leq \|u_1 - u_2\|_{w_i}$$

así que la anterior integral será menor o igual que:

$$\int_0^x (1 + x - 2s) w_i(s) \|u_1 - u_2\|_{w_i} ds = \|u_1 - u_2\|_{w_i} (Lw_i)(x)$$

De aquí se obtiene que

$$|(Lu_1)(x) - (Lu_2)(x)| \leq \|u_1 - u_2\|_{w_i} (Lw_i)(x) \quad (5.40)$$

Como veremos más adelante, $T(u_j) \geq \underline{\varphi}_1$ ($j = 1, 2$), así que:

$$|T(u_1)(x) - T(u_2)(x)| \leq \frac{|(Lu_1)(x) - (Lu_2)(x)|}{\underline{\varphi}_1^m(x)} \quad (5.41)$$

Explicación: Aplicando el Teorema de Lagrange, resulta:

$$\begin{aligned} |T(u_1)(x) - T(u_2)(x)| &= |W^{-1}(Lu_1)(x) - W^{-1}(Lu_2)(x)| \\ &= |(W^{-1})'(c(x))(Lu_1(x) - Lu_2(x))| \end{aligned}$$

donde $c(x)$ está entre $Lu_1(x)$ y $Lu_2(x)$.

Por otro lado:

$$(W^{-1})'(c(x)) = \frac{1}{k(W^{-1}(c(x)))} \leq \frac{1}{k(W^{-1}(Lu_1(x)))} = \frac{1}{k(T(u_1)(x))} \leq \frac{1}{\underline{\varphi}_1^m(x)}$$

Juntando esto con lo anterior se obtiene (5.41).

Finalmente,

$$\frac{|T(u_1)(x) - T(u_2)(x)|}{w_i(x)} \leq \frac{|Lu_1(x) - Lu_2(x)|}{w_i(x) \underline{\varphi}_1^m(x)} \leq \frac{\|u_1 - u_2\|_{w_i} (Lw_i)(x)}{w_i(x) \underline{\varphi}_1^m(x)} \leq \lambda_i \|u_1 - u_2\|$$

para $x \in (0, 1]$, donde la primera desigualdad es por (5.41), la segunda por (5.40) y la tercera por (5.37).

Tomando $\sup_{x \in (0,1]}$ se obtiene que:

$$\|T(u_1) - T(u_2)\| \leq \lambda_i \|u_1 - u_2\|$$

como queríamos demostrar.

Veamos ahora lo que dejamos pendiente:

$$T(u_j) \geq \underline{\varphi}_1 \quad (j = 1, 2)$$

Concretamente, veremos que si $u \in P = \{u \in K_0 : \underline{\varphi} \leq u \leq \Phi\}$ entonces $T(u) \geq \underline{\varphi}_1$. Dado que por (5.34) tenemos que $\underline{\varphi}_1 \leq \underline{\varphi}$, bastará ver que $\underline{\varphi} \leq T(\underline{\varphi})$, pues entonces:

$$\underline{\varphi} \leq u \Rightarrow \underline{\varphi}_1 \leq \underline{\varphi} \leq T(\underline{\varphi}) \leq T(u)$$

Para ver que $\underline{\varphi} \leq T(\underline{\varphi})$ será suficiente probar que $S(\underline{\varphi}) \geq 0$, siendo $S(\underline{\varphi}) = L\underline{\varphi} - W(\underline{\varphi})$, es decir, queremos ver que:

$$\int_0^x (1+x-2s)\underline{\varphi}(s)ds - \frac{(\underline{\varphi}(x))^{m+1}}{m+1} \geq 0$$

Cuando $x = 0$ se da la igualdad, por lo tanto bastará ver que la derivada es mayor o igual que cero.

La función es

$$(1+x) \int_0^x \underline{\varphi}(s)ds - 2 \int_0^x s\underline{\varphi}(s)ds - \frac{\underline{\varphi}(x)^{m+1}}{m+1}$$

Su derivada será:

$$\begin{aligned} 1 \cdot \int_0^x \underline{\varphi}(s)ds + (1+x)\underline{\varphi}(x) - 2x\underline{\varphi}(x) - (\underline{\varphi}(x))^m \underline{\varphi}'(x) \\ = \int_0^x \underline{\varphi}(s)ds + (1-x)\underline{\varphi}(x) - (\underline{\varphi}(x))^m \underline{\varphi}'(x) \end{aligned}$$

Dado que esto también vale 0 cuando $x = 0$, bastará ver que su derivada es mayor o igual que cero.

Llegados a este punto haremos uso del Lema 19, en el que tenemos el valor de S'' para funciones con la forma que tiene $\underline{\varphi}$.

$$\begin{aligned} S''(\underline{\varphi})(x) = \underline{\varphi}^{1-m}(x)x[2d_1(m+1) - 1 - [2d_1(d_1(m+2) - 1) + 3d_2(2m+1)]x \\ + [4d_1(2m+3) - 3]d_2x^2 - 3d_2^2(2m+3)x^3] \end{aligned}$$

Bastará con ver qué ocurre si $d_1 = d_1(m)$ y $d_2 = d_2(m)$.

Concentrándonos en lo que hay dentro de los corchetes, el término independiente sale:

$$2d_1(m+1) - 1 = 2 \cdot \frac{1}{2(m+1)}(m+1) - 1 = 0$$

El término en x vale:

$$\begin{aligned} & 2d_1(d_1(m+2) - 1) + 3d_2(2m+1) \\ &= 2 \cdot \frac{1}{2(m+1)} \left(\frac{1}{2(m+1)}(m+2) - 1 \right) + \frac{3m(2m+1)}{6(m+1)^2(2m+1)} \\ &= \frac{1}{m+1} \left(\frac{m+2}{2m+2} - \frac{2m+2}{2m+2} \right) + \frac{m}{2(m+1)^2} \\ &= \frac{1}{m+1} \left(\frac{-m}{2(m+1)} \right) + \frac{m}{2(m+1)^2} = 0 \end{aligned}$$

Para los términos que quedan nos bastará con que se cumpla:

$$(4d_1(2m+3) - 3)d_2x^2 - 3d_2^2(2m+3)x^3 \geq 0$$

Quitando el factor d_2x^2 es suficiente ver que:

$$4d_1(2m+3) - 3 - 3d_2(2m+3)x \geq 0$$

es decir:

$$4 \cdot \frac{1}{2(m+1)}(2m+3) - 3 - \frac{3m(2m+3)x}{6(m+1)^2(2m+1)} \geq 0$$

o bien:

$$\frac{4m+6}{m+1} - \frac{3m+3}{m+1} - \frac{3m(2m+3)x}{6(m+1)^2(2m+1)} \geq 0$$

es decir

$$\frac{m+3}{m+1} - \frac{m(2m+3)}{2(m+1)^2(2m+1)}x \geq 0$$

o sea:

$$m+3 \geq \frac{m(2m+3)}{2(m+1)(2m+1)}$$

es decir:

$$x \leq \frac{2(m+3)(m+1)(2m+1)}{m(2m+3)}$$

Si queremos que esto sea cierto $\forall x \in (0, 1]$ bastará con que sea cierto para $x = 1$:

$$\begin{aligned} m(2m+3) &\leq 2(m+3)(m+1)(2m+1) \\ 2m^2 + 3m &\leq 2(m^2 + 4m + 3)(2m+1) \\ 2m^2 + 3m &\leq 2(2m^2 + m^2 + 8m^2 + 4m + 6m + 3) \end{aligned}$$

$$2m^2 + 3m \leq 2(2m^3 + 9m^2 + 10m + 3)$$

$$2m^2 + 3m \leq 4m^3 + 18m^2 + 20m + 6$$

$$4m^3 + 16m^2 + 17m + 6 \geq 0$$

Lo cual es cierto $\forall m \geq 0$. □

Comentario 27. El último lema muestra que el valor del coeficiente contractivo λ_i depende de la elección de w_i .

5.7. Soluciones aproximadas para la ecuación (5.27).

Primero intentamos encontrar la estimación de $S(\underline{\varphi})$ a partir de lo anterior.

Lema 28. Si $\underline{\varphi}(x) = [m(x - d_1(m)x^2 + d_2(m))x^3]^{1/m}$ como en (5.32) entonces:

$$S(\underline{\varphi})(x) \leq 2m^{1/m}d_2(m) \frac{(m+3)m}{(m+1)(3m+1)(4m+1)} x^{4+1/m} \quad (5.42)$$

para $x \in [0, 1]$, donde $d_2(m)$ está definido por (5.33).

Demostración. Es fácil de ver que $S(\underline{\varphi})(0) = [S(\underline{\varphi})]'(0) = 0$.

Veamos esto primero:

$$S(\underline{\varphi})(x) = (L\underline{\varphi} - W(\underline{\varphi}))(x)$$

$$S(\underline{\varphi})(0) = (L\underline{\varphi})(0) - W(\underline{\varphi}(0)) = 0 - 0 = 0$$

$$S(\underline{\varphi})(x) = (1+x) \int_0^x \underline{\varphi}(s) ds - 2 \int_0^x s \underline{\varphi}(s) ds - W(\underline{\varphi}(x))$$

$$[S(\underline{\varphi})]'(x) = 1 \cdot \int_0^x \underline{\varphi}(s) ds + (1+x)\underline{\varphi}(x) - 2x\underline{\varphi}(x) - k(\underline{\varphi}(x))\underline{\varphi}'(x)$$

Cuando x vale 0, la integral desaparece, $\underline{\varphi}(0) = 0$ y $k(\underline{\varphi}(0)) = [\underline{\varphi}(0)]^m = 0$, con lo que todo lo anterior vale cero.

Con ayuda de (5.29) tomando

$$d_1(m) = \frac{1}{2(m+1)} \quad \text{y} \quad d_2(m) = \frac{m}{6(m+1)^2(2m+1)}$$

vamos a obtener lo siguiente:

$$[S(\underline{\varphi})]''(x) \leq 2m^{1/m}d_2(m) \frac{(m+3)}{(m+1)m} x^{2+1/m}$$

para $x \in (0, 1]$. Integrando dos veces la última desigualdad obtenemos (5.42). Veámoslo: De (5.29) sale lo siguiente:

$$S''(\varphi)(x) = \varphi^{1-m}(x)xq(x)$$

El término independiente de $q(x)$ es:

$$2d_1(m+1) - 1 = \frac{2(m+1)}{2(m+1)} - 1 = 0$$

El término en x de $q(x)$ es:

$$\begin{aligned} & - [2d_1(d_1(m+2) - 1) + 3d_2(2m+1)] \\ &= - \left[\frac{2}{2(m+1)} \left(\frac{m+2}{2(m+1)} - 1 \right) + \frac{3m(2m+1)}{6(m+1)^2(2m+1)} \right] \\ &= - \left[\frac{1}{m+1} \left(\frac{m+2 - (2m+2)}{2(m+1)} \right) + \frac{m}{2(m+1)^2} \right] \\ &= - \left[\frac{-m}{2(m+1)^2} + \frac{m}{2(m+1)^2} \right] = 0 \end{aligned}$$

Lo que hay que demostrar es entonces que:

$$\begin{aligned} \underline{\varphi}^{1-m}(x)x [(4d_1(2m+3) - 3)d_2x^2 - 3d_2^2(2m+3)x^3] \\ \leq 2m^{\frac{1}{m}}d_2(m) \frac{(m+3)}{(m+1)m} x^{2+1/m} \end{aligned}$$

Eliminando un factor d_2x^2 sale lo siguiente:

$$\underline{\varphi}^{1-m}(x)x [(4d_1(2m+3) - 3) - 3d_2(2m+3)x] \leq 2m^{\frac{1}{m}} \frac{(m+3)}{(m+1)m} x^{1/m}$$

O sea:

$$\underline{\varphi}^{1-m}(x)x \left[\left(\frac{4(2m+3)}{2(m+1)} - 3 \right) - \frac{3(2m+3)m}{6(m+1)^2(2m+1)} x \right] \leq 2m^{\frac{1}{m}} \frac{(m+3)}{(m+1)m} x^{1/m}$$

Es decir:

$$\underline{\varphi}^{1-m}(x)x \left[\left(\frac{8m+12-6m-6}{2(m+1)} \right) - \frac{(2m+3)m}{2(m+1)^2(2m+1)} x \right] \leq 2m^{\frac{1}{m}} \frac{(m+3)}{(m+1)m} x^{1/m}$$

O lo que es lo mismo, multiplicando todo por $m+1$

$$\underline{\varphi}^{1-m}(x)x \left[m+3 - \frac{(2m+3)m}{2(m+1)(2m+1)} x \right] \leq \frac{2m^{\frac{1}{m}}(m+3)}{m} x^{1/m}$$

Bastará ver que es cierto sin restar al $m + 3$ de dentro de los corchetes lo que le estamos restando, en cuyo caso podemos eliminar $m + 3$ de la izquierda y la derecha, quedando lo siguiente:

$$\underline{\varphi}^{1-m}(x) \leq 2m^{\frac{1}{m}-1}x^{\frac{1}{m}-1}$$

Dado que

$$\underline{\varphi}(x) = \left[m \left(x - \frac{1}{2(m+1)}x^2 + \frac{m}{6(m+1)^2(2m+1)}x^3 \right) \right]^{1/m}$$

tenemos que demostrar que:

$$\left[m \left(x - \frac{1}{2(m+1)}x^2 + \frac{m}{6(m+1)^2(2m+1)}x^3 \right) \right]^{\frac{1}{m}-1} \leq 2m^{\frac{1}{m}-1}x^{\frac{1}{m}-1}$$

Puede eliminarse el factor $m^{\frac{1}{m}-1}x^{\frac{1}{m}-1}$ de ambos lados y queda lo siguiente:

$$\left(1 - \frac{1}{2(m+1)}x + \frac{m}{6(m+1)^2(2m+1)}x^2 \right)^{\frac{1}{m}-1} \leq 2$$

Si $m < 1$, entonces el exponente es positivo y podemos demostrar de hecho que eso es menor o igual que 1, pues bastará que se cumpla

$$1 - \frac{1}{2(m+1)}x + \frac{m}{6(m+1)^2(2m+1)}x^2 \leq 1$$

para lo cual, teniendo en cuenta que $x \geq 0$, bastará con que se cumpla esto otro:

$$\frac{m}{6(m+1)^2(2m+1)}x \leq \frac{1}{2(m+1)}$$

es decir:

$$\frac{m}{3(m+1)(2m+1)}x \leq 1$$

Como queremos ver que es cierto para $x \in [0, 1]$, bastará ver que es cierto para $x = 1$, en cuyo caso se tiene que:

$$\begin{aligned} m \leq 3(m+1)(2m+1) &\iff m \leq 3(2m^2 + 3m + 1) \\ &\iff m \leq 6m^2 + 9m + 3 \iff 0 \leq 6m^2 + 8m + 3 \end{aligned}$$

Lo cual es por fin evidente.

Veamos ahora el caso en que $m \geq 1$. Dado que la función x^α es decreciente si $\alpha < 0$, podemos escribir:

$$\begin{aligned} \left(1 - \frac{1}{2(m+1)}x + \frac{m}{6(m+1)^2(2m+1)}x^2 \right)^{\frac{1}{m}-1} &\leq \left(1 - \frac{1}{2(m+1)}x \right)^{\frac{1}{m}-1} \\ &\leq \left(1 - \frac{1}{2(m+1)} \right)^{\frac{1}{m}-1} \end{aligned}$$

Teniendo en cuenta que si $0 < x \leq 1$ y $\alpha \leq \beta$ entonces $x^\beta \leq x^\alpha$ se obtiene que:

$$\left(1 - \frac{1}{2(m+1)}\right)^{\frac{1}{m}-1} \leq \left(1 - \frac{1}{2(m+1)}\right)^{-1} \leq \left(1 - \frac{1}{4}\right)^{-1} = \frac{4}{3} \leq 2$$

como queríamos demostrar. \square

Corolario 29. Como $k(u) = (m+1)u^m$ entonces tenemos

$$0 \leq \frac{S(\underline{\varphi})(x)}{k(\underline{\varphi})(x)} \leq \frac{S(\underline{\varphi})(x)}{k(\underline{\varphi}_1)(x)} \leq \frac{4d_2(m)(m+3)m^{1/m}}{(2m+1)(3m+1)(4m+1)}x^{3+1/m} \quad (5.43)$$

para $x \in (0, 1]$.

Demostración. La primera desigualdad es inmediata ya que tanto numerador como denominador son positivos.

La segunda desigualdad sale porque $\underline{\varphi}_1 \leq \underline{\varphi}$ (Corolario 5.34).

($\underline{\varphi}$ tiene un término positivo que $\underline{\varphi}_1$ no tiene).

Como

$$k(\underline{\varphi}_1)(x) = m \left(x - \frac{1}{2(m+1)}x^2 \right)$$

aplicando el Lema 28, bastará ver que:

$$\frac{2m^{1/m}d_2(m)\frac{(m+3)m}{(m+1)(3m+1)(4m+1)}x^{4+1/m}}{m \left(x - \frac{1}{2(m+1)}x^2 \right)} \leq \frac{4d_2(m)(m+3)m^{1/m}}{(2m+1)(3m+1)(4m+1)}x^{3+1/m}$$

Simplificando todo lo que se puede, lo que hay que ver es que:

$$\frac{1}{(m+1) \left(1 - \frac{1}{2(m+1)}x \right)} \leq \frac{2}{2m+1} \quad x \in (0, 1]$$

es decir:

$$2m+1 \leq 2(m+1) \left(1 - \frac{1}{2(m+1)}x \right) \quad x \in (0, 1]$$

Bastará con que veamos que es cierto para $x = 1$:

$$2m+1 \leq 2(m+1) \left(1 - \frac{1}{2(m+1)} \right)$$

O sea:

$$2m+1 \leq 2(m+1) \frac{2m+1}{2(m+1)}$$

Lo cual es cierto porque se da la igualdad. \square

Nota previa al Corolario siguiente: Si llamamos

$$\alpha = \frac{4d_2(m)(m+3)m^{1/m}}{(2m+1)(3m+1)(4m+1)}$$

entonces el Corolario 29 lo que dice es que

$$\frac{S(\underline{\varphi})(x)}{k(\underline{\varphi})(x)} \leq \alpha x^{3+1/m}$$

Corolario 30. Cabe destacar que para w_i ($i = 1, 2, 3$)

$$\left\| \frac{S(\underline{\varphi})}{k(\underline{\varphi})} \right\|_{w_i} \leq \frac{4d_2(m)(m+3)m^{1/m}}{(2m+1)(3m+1)(4m+1)} \equiv \alpha \quad (5.44)$$

(En el caso w_3 tenemos $m \geq 1$).

Demostración. Para w_1 :

$$\sup_{x \in (0,1]} \frac{\alpha x^{3+1/m}}{x^{2+1/m}} = \sup_{x \in (0,1]} \alpha x = \alpha$$

Para w_2 :

$$\sup_{x \in (0,1]} \frac{\alpha x^{3+1/m}}{x^{1/m}} = \sup_{x \in (0,1]} \alpha x^3 = \alpha$$

Para w_3 :

$$\sup_{x \in (0,1]} \frac{\alpha x^{3+1/m}}{x^{3/m}} = \sup_{x \in (0,1]} \alpha x^{\frac{3m-2}{m}} = \alpha$$

siempre que sea $3m - 2 > 0$, es decir $m > \frac{2}{3}$, lo cual es cierto si $m \geq 1$. □

Ahora, de acuerdo con el Lema 16 podemos formular:

Teorema 31. Si $u \in K_0$ es la (única) solución de la ecuación (5.27) entonces

$$\underline{\varphi}(x) \leq u(x) \leq \underline{\varphi}(x) + \bar{\lambda}_i w_i(x)$$

para $x \in [0, 1]$ ($i = 1, 2, 3$), donde $\underline{\varphi}$ está definido por (5.32), es decir:

$$\underline{\varphi}(x) = [m(x - d_1(m)x^2 + d_2(m)x^3)]^{1/m}$$

y

$$\begin{aligned} \bar{\lambda}_1 &= \frac{4d_2(m)(m+3)m^{1/m}}{3m(2m+1)(4m+1)} \\ \bar{\lambda}_2 &= \frac{4d_3(m)(m+3)(m+1)m^{1/m}}{m(2m+1)(3m+1)(4m+1)} \\ \bar{\lambda}_3 &= \frac{4d_3(m)(m+3)^2 m^{1/m}}{(m+2)(2m+1)(3m+1)(4m+1)} \quad (m \geq 1) \end{aligned}$$

Demostración. Se hace aplicando el Lema 16, tomando $p = \underline{\varphi}$, el valor definido de α y tomando γ como cada uno de los λ_i , con lo cual se obtiene

$$\bar{\lambda}_i = \frac{\alpha}{1 - \lambda_i}$$

En efecto:

$$\bar{\lambda}_1 = \frac{\alpha}{1 - \lambda_1} = \frac{\alpha}{1 - \frac{1}{3m+1}} = \frac{(3m+1)\alpha}{3m}$$

$$\bar{\lambda}_2 = \frac{\alpha}{1 - \lambda_2} = \frac{\alpha}{1 - \frac{1}{m+1}} = \frac{(m+1)\alpha}{m}$$

$$\bar{\lambda}_3 = \frac{\alpha}{1 - \lambda_3} = \frac{\alpha}{1 - \frac{1}{m+3}} = \frac{(m+3)\alpha}{m+2}$$

□

Corolario 32. *Obviamente la mejor estimación se obtiene utilizando w_1 .*

Quiere decir, que de los tres λ_i , el más pequeño (que también es el que da el $\bar{\lambda}_i$ más pequeño) es λ_1 .

Que $\lambda_1 \leq \lambda_2$ es claro, ya que

$$\frac{1}{3m+1} \leq \frac{1}{m+1}$$

Si $m \geq 1$ se cumple también que $\lambda_1 \leq \lambda_3$, ya que

$$\frac{1}{3m+1} \leq \frac{1}{m+3} \iff m+3 \leq 3m+1 \iff 2 \leq 2m \iff 1 \leq m$$

Corolario 33. *Podemos estimar los errores de aproximación de u mediante $\underline{\varphi}$. Tenemos menos que*

6.2 % para $m = 0,5$, 0.06 % para $m = 2$,

1.7 % para $m = 0,75$, 0.02 % para $m = 3$,

0.6 % para $m = 1$, 0.005 % para $m = 4$.

El error decrece muy rápidamente conforme m crece.

Corolario 34. *Aproximando u por $\underline{\varphi}$, el error cometido es menor que 6,2 % para $m = 0,5$, etc.*

Corolario 35. *Si $m \gg 1$ entonces $u \sim \underline{\varphi}$ en $[0, 1]$.*

5.8. Nota final y comentarios.

Si regresamos al problema presentado en la Sección 4 entonces las funciones apropiadas $\underline{\varphi}_i$ ($i = 1, 2$) toman la forma

$$\underline{\varphi}_i(x) = \left[\frac{(A_i a + f_i(a))^2}{A_i c_i (m+1)} \right]^{1/m} \underline{\varphi} \left(\frac{A_i x}{A_i a + f(a)} \right) \quad (i = 1, 2) \quad (5.45)$$

donde $x \in [0, a]$ y $\underline{\varphi}$ están definidos por (5.32). Si $c_2 - c_1 \ll 1$ y $A_2 - A_1 \ll 1$ entonces por el punto 6. de la idea de la Sección 3 y por el Teorema 31 se infiere que $\underline{\varphi}_i$ ($i = 1, 2$) es una aproximación muy buena de la solución no trivial del problema considerado. En este caso podemos también formular el corolario similar al Corolario 35, a saber:

Corolario 36. *Sea u la solución no trivial del problema presentado en la Sección 4. Si $m \gg 1$ entonces $\underline{\varphi}_i \sim u$ ($i = 1, 2$) en $[0, a]$*

En este trabajo mostramos el método de aproximación de soluciones no triviales basándonos en un ejemplo. Podemos ver que la aproximación depende de la elección de la función de peso w y de la clase de funciones φ que satisfacen $S''(\varphi) \geq 0$.

En uno de los artículos recientes [28] se ha mostrado cómo encontrar cualquier función de peso w y el coeficiente de contracción relacionado en un caso más general que (5.27).

Pensamos que con la ayuda de los resultados de [28] será posible aplicar también el presente método de aproximación en algunos casos más generales que el considerado.

Bibliografía

- [1] D. Anderson and M. Lisak. Approximate solutions of some nonlinear diffusion equations. *Phys. Rev. A*, 22:2761–2768, 1980.
- [2] D. G. Aronson. The porous medium equation. In Antonio Fasano and Mario Primicerio, editors, *Nonlinear Diffusion Problems*, volume 1224 of *Lecture Notes in Mathematics*, pages 1–46. Springer Berlin Heidelberg, 1986.
- [3] D. G. Aronson and J. L. Vázquez. Calculation of anomalous exponents in nonlinear diffusion. *Phys. Rev. Lett.*, 72:348–351, Jan 1994.
- [4] G. I. Barenblatt. On some unsteady motions of a liquid and gas in a porous medium. *Akad. Nauk SSSR. Prikl. Mat. Meh.*, 16:67–78, 1952.
- [5] G. I. Barenblatt. Self-similar turbulence propagation from an instantaneous plane source. In G. I. Barenblatt, G. Iooss, and D. D. Joseph, editors, *Nonlinear Dynamics and Turbulence*, pages 48–60. Pitman, 1983.
- [6] J. Bear. *Dynamics of Fluids in Porous Media*. Number v. 1 in Dynamics of Fluids in Porous Media. American Elsevier Publishing Company, 1972.
- [7] M. Bertsch. A class of degenerate diffusion equations with a singular nonlinear term. *Nonlinear Anal.*, 7:117–127, 1983.
- [8] M. Bertsch and L. A. Peletier. Porous media equation: an overview. 1983.
- [9] P. J. Bushell. On a class of Volterra and Fredholm non-linear integral equations. *Math. Proc. Cambridge Philos. Soc.*, 79:329–335, 1976.
- [10] J. R. Cannon. The one-dimensional heat equation. In *Encyclopedia of Mathematics and its Applications*, volume 23. Addison-Wesley Publishing Co. Reading Mass, 1984.
- [11] D. S. Cohen and W. L. Kath. Waiting-time behaviour in a nonlinear diffusion problem. *Stud. Appl. Mat.*, 67:79–105, 1982.
- [12] H. Darcy. *Les fontaines publiques de la ville de Dijon: Exposition et application des principes a suivre et des formules a employer dans les questions de*

distribution d'eau; ouvrage terminé par un appendice relatif aux fournitures d'eau de plusieurs villes au filtrage des eaux et a la fabrication des tuyaux de fonte, de plomb, de tole et de bitume. Atlas. Victor Dalmont, Libraire des Corps imperiaux des ponts et chaussées et des mines, 1856.

- [13] R. B. Fair. Boron diffusion in silicon-concentration and orientation dependence, background effects and profile estimation. *J. Electrochem. Soc.*, 122:800–805, 1975.
- [14] A. Friedman. *Partial Differential Equations of Parabolic Type*. Dover Books on Mathematics. Dover Publications, 2013.
- [15] S. K. Ghandhi. *VLSI Fabrication Principles*. Wiley, New York, 1983.
- [16] B. H. Gilding and L. A. Peletier. On a class of similarity solutions of the porous media equation II. *J. Math. Anal. Appl.*, 57(3):522–538, 1977.
- [17] J. Goncerzewicz. Porous medium-type equation with irregular boundary data. In *Pitman Res. Notes in Math. Sci. Ser.*, volume 208, pages 59–67, 1989.
- [18] J. Goncerzewicz. Properties of solutions to the initial-boundary value problem for a porous media-type equation. *Math. Meth. in the Appl. Sci.*, 15:299–314, 1992.
- [19] J. M. Hill and N. F. Smith. High order nonlinear diffusion. *IMA J. Appl. Math.*, 40:73–86, 1988.
- [20] A. S. Kalashnikov. Some problems of the qualitative theory of non-linear degenerate second-order parabolic equations. *Russian Mathematical Surveys*, 42(2):169–222, 1987.
- [21] S. Kamin and J. L. Vázquez. The propagation of turbulent bursts. *Euro. J. of Appl. Math.*, 3:263–272, 1992.
- [22] J. R. King. High concentration arsenic diffusion in crystalline silicon: An asymptotic analysis. *IMA J. Appl. Math.*, 38(2):87–95, 1987.
- [23] J. R. King. Approximate solutions to a nonlinear diffusion equation. *J. Engng. Math.*, 22(1):53–72, 1988.
- [24] A. Luque, J. Martín, and G. L. Araújo. Zn diffusion in GaAs under constant As pressure. *J. Electrochem. Soc.*, 123(2):249–254, 1976.
- [25] M. Muskat and R. D. Wyckoff. *The Flow of Homogeneous Fluids Through Porous Media*. International series in physics. McGraw-Hill Book Company, Incorporated, 1937.

- [26] Y. Nakajima, S. Ohkawa, and Y. Fukukawa. Simplified expression for the distribution of diffused impurity. *Jap. J. Appl. Phys.*, 10(1):162–163, 1971.
- [27] S. Staňek. Qualitative behaviour of a class of second order nonlinear differential equation on the halfline. *Ann. Polon. Math.*, 58:65–83, 1993.
- [28] J. J. Nieto and W. Okrański. Existence, uniqueness, and approximation of solutions to some nonlinear diffusion problems. *J. Math. Anal. Appl.*, 210(1):231–240, 1997.
- [29] W. Okrański and S. Vila. Approximations of solutions to some second order nonlinear differential equations. *Nonlinear Anal.*, 35(9):1061–1072, March 1999.
- [30] W. Okrański. Integral equations methods in the theory of the water percolation. In *Mathematical Methods in Fluids Dynamics, Oberwolfach, 1981*, volume 24, pages 167–176. P. Lang, Frankfurt, 1982.
- [31] W. Okrański. On a nonlinear ordinary differential equation. *Annal. Polon. Math.*, 49:237–245, 1989.
- [32] W. Okrański. On approximate solutions to some nonlinear diffusion problems. *Z. Angew. Math. Phys.*, 44:722–731, 1993.
- [33] W. Okrański and S. Vila. Power series solutions to some nonlinear diffusion problems. *Z. Angew. Math. Phys.*, 44:988–997, 1993.
- [34] W. Okrański and S. Vila. Determination of the interface position for some nonlinear diffusion problems. *Applied Mathematics Letters*, 11(4):85–89, 1998.
- [35] O. A. Oleńnik, A. S. Kalashnikov, and Yu lin Zhou. The Cauchy problem and boundary problems for equations of the type of non-stationary filtration. *Izv. Akad. Nauk SSSR, Ser. Mat.*, 22:667–704, 1958.
- [36] D. O'Regan. Existence theory for the equations $(G'(y))' = qf(t, y, y')$ and $(G'(y) - pH(y))' = -p'H(y) + qf(t, y)$. *J. Math. Anal. Appl.*, 183:435–470, 1994.
- [37] R. E. Pattle. Diffusion from an instantaneous point source with a concentration-dependent coefficient. *The Quarterly Journal of Mechanics and Applied Mathematics*, 12(4):407–409, January 1959.
- [38] W. H. Press, W. T. Vetterling, S. A. Teukolsky, and B. P. Flannery. *Numerical Recipes in C*. Cambridge University Press, Cambridge, 2nd edition, 1992.
- [39] A. Rybarski. The integral equation describing propagation of the ground water tongue. preprint in Polish, 1973.

- [40] S. M. Sze. *Semiconductor Devices, Physics and Technology*. Wiley, New York, 1985.
- [41] B. Tuck. Some explicit solutions to the nonlinear diffusion equation. *J. Phys. D.*, 9:1559–1569, 1976.
- [42] J. L. Vázquez. An introduction to the mathematical theory of the porous medium equation. In Michel C. Delfour and Gert Sabidussi, editors, *Shape Optimization and Free Boundaries*, volume 380 of *NATO ASI Series*, pages 347–389. Springer Netherlands, 1992.
- [43] L. R. Weisberg and J. Blanc. Diffusion with interstitial-substitutional equilibrium, Zinc in GaAs. *Phys. Rev.*, 131:1548–1552, 1963.
- [44] T. P. Witelski. Stopping and merging problems for the porous media equation. *IMA J. Appl. Math.*, 54:227–243, 1995.
- [45] Ya. B. Zel'dovich and A. S. Kompaneets. Theory of heat transfer with temperature dependent thermal conductivity. Collection in honour of the 70th birthday of academician A.F. Ioffe, pages 61–71. 1950.