

**TRABAJO FIN DE GRADO**

Facultad de Ciencias

Grado en Matemáticas

**Autor:** Héctor Sanz Herranz

**Tutor:** Javier de Frutos Baraja

*[El Método de  
Elementos Finitos en la  
Valoración de  
Opciones.]*



---

**Universidad de Valladolid**







Septiembre de 2016.



# PROLEGÓMENOS.

## LA FÁBRICA DE IRREALIDAD.

"El mundo es todo cuanto es el caso". Con esta sentencia comienza Ludwig Wittgenstein su *Tractatus logico-philosophicus*, una de las obras más influyentes en el pensamiento del siglo XX. Hasta donde quien escribe estas líneas alcanza a comprender, Wittgenstein trata de distinguir lo que son las antípodas de toda representación o figuración para hacer, después, un original análisis del lenguaje. Según él, existe un constructo figurativo de la realidad que se corresponde con lo que sería el mundo. Sin embargo, esta asociación es meramente supuesta pues no existe una justificación clara de ella. Así, cuando pensamos lo que estamos haciendo es modelar, figurar, de algún modo la realidad y es, precisamente, el objetivo último del *Tractatus* el tratar de encontrar ese algo ulterior que posibilita la figuración, es decir, eso común a realidad y pensamiento que hace posible tal modelado.

Seguramente, a estas alturas, el propio Ludwig se estaría rasgando las vestiduras o, mejor aún, empezaría a valorar (como, por otra parte, acostumbraba a hacer) el suicidio como única alternativa a la incompreensión de su *Tractatus*. Por tanto, dejaremos de comentar su obra en aras de su eterno descanso. En cambio, si que siento que existe una dualidad asombrosa entre lo real y lo irreal de la que, a menudo, no somos conscientes. En particular, las matemáticas son en mi opinión la mayor fábrica de irrealidad que el hombre ha construido. Mucho mayor incluso que el cine o la televisión, pues cuando uno ve una película es consciente de su posición relativa frente al aparato, tiene claro ser el *sujeto observante* y eso le confiere una posición de superioridad, dominante, frente al *objeto observado*. No ocurre así, en cambio, con las matemáticas. Cuando uno se sumerge en ellas, fruto de la admiración que suscitan, pierde habitualmente la percepción de que toda la matemática es en esencia una mera figuración, un modelo.

Y es que, en nuestra condición de contingentes, gestamos el anhelo melancólico de una palmadita en la espalda, de un cálido manto de camuflaje o de una armadura de cota de malla, que nos tranquilice y apacigüe la angustia que nos provoca la configuración de estados posibles que es la realidad. Cuando la figuramos llegan los problemas. Problemas sencillos, complejos, resolubles o lo contrario. Pero problemas. Y es que poco le importa a quien camina cansado la condición, la configuración efectiva, de tales modelos pues cree, y pronto sabrá cuando retorne a la duda, que no se podrá colorear nunca el universo figurativo con una cantidad numerable de tonos.

De vuelta a la realidad, no se puede gozar de esa tranquilidad. Lo que se tiene a este lado no son problemas. Son, cuanto menos, otras cosas. A veces, son desatinos de la fortuna y, otras veces, criaturas abandonadas. Pero no problemas. Por eso proyectamos

nuestra realidad en lo intangible. Por el descanso que otorga escapar de los monstruos allá donde (creemos) no llegan.

Aún así, hay incluso quienes se plantean si son las matemáticas una creación del hombre o si, por el contrario, son un *algo* que estaba esperando ser descubierto. En todo caso, son estas cuestiones las que dotan de sentido el estudio de esta ciencia pues, su existencia es en tanto en que, efectivamente, pueda ser comprendida. Y es que *"de lo que no se puede hablar, mejor es callar"*.



## RESUMEN.

En este trabajo se expone una introducción a los conceptos propios de la teoría de valoración de derivados financieros prestando especial atención a las opciones. Asimismo, se presenta el método de los elementos finitos para la resolución de ecuaciones en derivadas parciales justificando su uso desde el marco teórico. Finalmente, se muestran sendos algoritmos del método en algunos problemas de valoración de derivados financieros.

## ABSTRACT.

An introduction to proper concepts of the financial derivatives is exposed in this work, giving a special attention to the options. Likewise, finite element method is introduced in order to solve partial differential equations justifying its use from the theoretical frame. Finally, several algorithms of the method are shown into some financial derivatives valuation problems.

## TÉRMINOS CLAVE.

Derivados financieros, opción financiera, modelo de Black-Scholes, formulación variacional, espacio de Sobolev, método de elementos finitos.



## AGRADECIMIENTOS.

Resulta que agradecer públicamente lo que es evidentemente susceptible de ser agradecido constituye una exposición sublime del sentimentalismo que aniquila la intimidad del alma. Quien tiene confianza con todo el mundo es que, en realidad, no la tiene con nadie. Por eso, prefiero no enumerar aquí a las personas a las que, por supuesto, tengo algo que agradecer sino que, en cambio, mi reconocimiento será un mensaje sin acuse de recibo para que quien lo encuentre valore si merece mi gratitud y, si así lo estima, espero de buen grado que la acepte.



# CONTENIDOS.

<b>Introducción.</b>	<b>xiii</b>
<b>1 El Modelo de Black-Scholes.</b>	<b>1</b>
1.1 Introducción y Primeros Conceptos. . . . .	3
1.2 Tipos de Opciones. . . . .	5
1.3 Un Modelo para el Valor de Subyacente. . . . .	7
1.4 La Ecuación de Black-Scholes y su Entorno. . . . .	15
1.5 El Caso más Simple. . . . .	18
1.6 Referencias Bibliográficas. . . . .	22
<b>2 El Método de los Elementos Finitos.</b>	<b>27</b>
2.1 Panorama General. . . . .	29
2.2 Algunos Conceptos de Análisis Funcional. . . . .	32
2.3 Un Problema Bien Puesto. . . . .	34
2.4 Espacios de Sobolev. . . . .	47
2.5 El Método Galerkin. . . . .	49
2.6 El Método de los Elementos Finitos. . . . .	54
2.7 Referencias Bibliográficas. . . . .	57
<b>3 Valoración de Opciones.</b>	<b>61</b>
3.1 Consideraciones Previas. . . . .	63
3.2 El Problema Discreto. . . . .	63
3.3 El Problema de la Localización. . . . .	66
3.4 Experimentos Numéricos. . . . .	72
3.5 La Opción Americana. . . . .	76
3.6 Referencias Bibliográficas. . . . .	81
<b>Comentarios finales.</b>	<b>85</b>
<b>A Programas en MatLab.</b>	<b>89</b>
A.1 BSeuropea. . . . .	89
A.2 BSamericana. . . . .	91
<b>Bibliografía.</b>	<b>95</b>



# INTRODUCCIÓN.

El trabajo que presentamos está motivado por la aplicación del método de los elementos finitos a la resolución numérica efectiva de las ecuaciones en derivadas parciales que aparecen en el ámbito de la valoración de ciertos derivados financieros, concretamente, de las opciones.

El problema de valorar opciones se remonta a comienzos del siglo XX con la publicación de la tesis doctoral de L. Bachelier, *Théorie de la spéculation*. ([2]). Sin embargo, es a partir de la publicación de los artículos de F. Black, M. Scholes ([5]) y R.C. Merton ([16]) en 1973 cuando da comienzo la teoría moderna de la matemática financiera.

En los últimos años se ha estado trabajando en los problemas de valoración de tipos más complejos de contratos, las denominadas *opciones exóticas*. Asimismo, se han ido descubriendo otras aplicaciones de la teoría de valoración de opciones como es la valoración de *warrants*.

En cuanto a la estructura del trabajo cabe señalar que introducimos, en el capítulo 1, los conceptos financieros que albergarán nuestra atención a lo largo de toda la memoria, en particular, el de opción. Se ha tratado de realizar una exposición clara y amena de los mismos que resulte comprensible por cualquier persona sin nociones en este campo. Más adelante, se deduce la célebre ecuación de Black-Scholes de la manera en que fue presentada originalmente en [5]. También se resuelve de manera explícita para un tipo concreto de opción: la europea.

Por su parte, el capítulo 2 comienza con el planteamiento de un problema de contorno concreto que se va utilizando como ejemplo a lo largo de toda esta parte. Se presentan resultados propios del análisis funcional y se describe el paradigma propio del método de elementos finitos.

Asimismo, el capítulo 3 está enfocado a la valoración efectiva de algunos tipos de opciones. Se describen los procedimientos habituales y se implementan los métodos descritos mediante sendos programas en MatLab elaborados para la ocasión.









# 1

## EL MODELO DE BLACK-SCHOLES.

---

*"No hay nada que aprender del éxito.  
Todo se aprende del fracaso."  
— David Bowie.*

**Resumen:** Se introducen los conceptos propios del problema de valoración de opciones. A continuación, se describe una modelización del precio del activo suyacente y se deduce la ecuación de Black-Scholes. Finalmente, se resuelve tal ecuación para el caso de las opciones europeas.



## 1.1. INTRODUCCIÓN Y PRIMEROS CONCEPTOS.

David Robert Jones, más conocido como David Bowie, constituye, sin duda uno de los máximos exponentes musicales de las últimas cinco décadas. Es considerado un innovador sobre todo por haber servido, a través de sus personajes reunidos en su persona, de nexo común entre varios estilos y corrientes verdaderamente distintos. Sin embargo, lo que no es tan conocido para el gran público es que también fue un innovador de la relación entre los artistas y los bancos. En 1997, Bowie firmó un contrato por el cual vendía los derechos (*royalties*) de los beneficios futuros de sus canciones en la llamada *Bowie Bond*.

Este mismo ingenio había tenido François-René de Chateaubriand aproximadamente ciento cincuenta años antes al respecto de su obra *Memorias de ultratumba*. Tardó treinta y cinco años en escribirla y no quería que la publicaran hasta cincuenta años después de su muerte pues, de hecho, el libro está escrito como si se tratase de la voz de un difunto. El problema era que necesitaba dinero en vida con motivo de numerosas deudas que había ido contrayendo. Por ello, realizó unas cuantas lecturas de sus memorias ante un selecto público con objeto de hallar un editor dispuesto a pagarle un anticipo, esto es, dispuesto a darle dinero por una obra que no vería la luz hasta pasados bastantes años.

Al igual que hicieron Chateaubriand y Bowie, las empresas también han de gestionar el riesgo en el que incurren cuando toman decisiones estratégicas, y de tal gestión depende, en cierto modo, la calidad de vida de la sociedad. Pensemos, por ejemplo, en una empresa del sector farmacéutico que está valorando si adquirir un medicamento en desarrollo que podría no reportar beneficios hasta pasada una década, si es que lo hace. En tal situación, sería conveniente llegar a un acuerdo, a cambio quizás de una cantidad, que le permita evitar las grandes pérdidas que podría suponer el fracaso en la investigación del medicamento.

Tal acuerdo se materializa en las llamadas *opciones financieras*. Éstas son contratos que proporcionan a su propietario, el *holder*, el derecho, pero no la obligación, de comprar o vender un bien a un precio predeterminado en el futuro. La otra parte del contrato, el *writer*, está obligada a vender o comprar ese bien si el *holder* lo solicita. Debido a esta asimetría de condiciones, el *holder* debe compensar económicamente al *writer* por la obligación que éste adquiere cuando subscriben el contrato. Ahora bien, el *holder* se ve en la necesidad de estimar la cuantía de esa compensación económica, esto es, debe *valorar* la opción. Pero antes de insistir en este asunto, continuemos introduciendo más conceptos que vamos a necesitar posteriormente.

Cuando la opción se establece sobre una posibilidad de compra, se dice que es una *call*. Del mismo modo, cuando refiere a una venta, entonces se la llama *put*. Como las opciones apuntan a una elección que tiene el poseedor de las mismas, se dice que éste *ejerce* la opción si efectivamente compra o vende el activo por el precio acordado. A ese precio se lo conoce como *precio de ejercicio* o *strike price*. Asimismo, al instante en el que finaliza el contrato se le llama *fecha de vencimiento* o *expiry date*.

Las opciones se engloban dentro de un tipo de valores más genérico llamado *derivados financieros*, los cuales son títulos cuyo valor depende del precio de otros activos comercializados a los que se los conoce como *activos subyacentes*. Otros derivados financieros son, por ejemplo, los *forwards*, los *swaps* o los *CDOs*<sup>1</sup>.

Podemos encontrar opciones financieras sobre activos de todo tipo, desde materias primas hasta índices bursátiles (S&P 100, *Dow Jones Industrial*, NYSE), pasando incluso por otras opciones. Sin embargo, los contratos de opciones más habituales se realizan sobre acciones<sup>2</sup>.

Las opciones *put* se pueden utilizar para compensar las pérdidas de una cartera de inversión<sup>3</sup> provocadas por el desplome del valor de las acciones en el mercado. El uso de opciones para reducir el riesgo de poseer títulos cuyos beneficios se correlacionan negativamente con la exposición a ciertos riesgos se llama *cobertura de riesgo* o *hedging*. Por otra parte, dado que las opciones son muy sensibles a las variaciones del valor del activo subyacente, alguien que quiera apostar sobre la dirección que cree que el mercado va a seguir podría recurrir a estos derivados para hacerlo de una forma que le permita aumentar sus beneficios potenciales (y también sus posibles pérdidas). Como dice Paul Wilmott ([23]), "*options can be a cheap way of exposing a portfolio to a large amount of risk*"<sup>4</sup>. Cuando un agente actúa de esta forma se habla de *especulación*.

En 1973 se publica el artículo mundialmente aceptado como la piedra angular del desarrollo de la matemática financiera: el artículo de Robert Merton ([16]) en el que hacía referencia a los trabajos de Fischer Black y Myron Scholes ([5]) y que incluye una fórmula analítica cerrada para valorar un tipo específico de opciones. Esto permitió que éstas pudiesen negociarse de una manera notablemente más fácil y rápida al conocer de antemano las condiciones legítimas sobre su valor en detrimento de la necesidad de reunión entre las dos partes para negociarlas. De este modo se inauguró, en ese mismo año, el primer mercado moderno, organizado y regulado de opciones que aún sigue funcionando en la actualidad: el *Chicago Board Options Exchange* (CBOE).

Cabe señalar que el mérito y la importancia de las contribuciones de Black, Scholes y Merton fue reconocida en 1997 cuando se les concedió el Premio Nobel de Economía. En

---

<sup>1</sup>Los contratos *forward* son un acuerdo entre dos partes de modo que una se compromete ineludiblemente a comprar o vender un activo a un precio especificado. Asimismo, un *swap* no es más que un intercambio de capital entre dos partes en una cierta fecha. La definición de un *CDO* no es tan simple; limitémonos a señalar que fue uno de los actores principales en la burbuja inmobiliaria que condujo a la quiebra de varios bancos de inversión norteamericanos como *Merrill Lynch* o *Lehman Brothers* y a la posterior crisis económica internacional.

<sup>2</sup>Una acción es cada una de las partes en las que se divide el capital de una sociedad anónima. Las acciones confieren a su poseedor ciertos derechos tales como la intervención proporcional en la junta de accionistas y la participación en los beneficios de la empresa. A la parte correspondiente del beneficio por cada acción se la llama *dividendo*.

<sup>3</sup>Una *cartera de inversión* o *portfolio* es una combinación de productos financieros en los que se decide invertir. El éxito de una inversión depende en gran medida de la elección que se haga de los productos que componen la cartera.

<sup>4</sup>"*las opciones pueden ser una manera barata de exponer una cartera de inversión a una gran cantidad de riesgo*".

realidad, con motivo de su muerte en 1995, Fisher Black no recibió tal mención aunque sí que fue citado como contribuidor por la Academia Sueca.

Desde que, en 1973, las opciones comenzaran a cotizarse, éstas se han convertido en uno de los productos financieros más importantes y negociados. Tanto es así, que existe una gran cantidad de mercados repartidos por todo el mundo que organizan, registran y garantizan esta actividad. Algunos de los más conocidos son, junto al CBOE, el EUREX, el *London International Financial Futures and options Exchange* (LIFFE) y el NASDAQ-OMX.

## 1.2. TIPOS DE OPCIONES.

A medida que el uso de los derivados financieros se fue consolidando, estos comenzaron a aumentar en complejidad. En lo que a opciones se refiere, se puede decir que los dos tipos más importantes son las opciones *européas* y las opciones *americanas*. Se alude a las primeras, tanto si son de compra como si son de venta, como opciones *vanilla* mientras que al resto de tipos se las conoce como opciones *exotic*<sup>5</sup>.

La característica fundamental que diferencia a las opciones europea y americana estriba en que, mientras que las europeas solo pueden ser ejercidas en la fecha de vencimiento del contrato, las americanas pueden ejercerse en cualquier instante desde el momento de la firma hasta el *expiry date*. A la hora de valorar el precio de estas opciones, esta diferencia aparentemente menuda y asumible va a suponer, en cambio, que resolver el problema para las opciones americanas sea notablemente más complejo que para las opciones europeas. Tanto es así que, como veremos, existe una fórmula analítica cerrada para valorar las opciones europeas, pero no para las americanas. No obstante, en la práctica no es realmente posible ejercer la opción en cualquier instante pues los subscriptores de los acuerdos suelen exigir que se avise de la ejecución con, al menos, un pequeño plazo de antelación.

La clave en el problema de ajustar un precio para los distintos tipos de contrato va a ser, sin duda, el llamado *payoff*, que es el rédito que nos otorgan estos derivados a la expiración del contrato. Para introducir este concepto, consideremos como ejemplo una opción europea de compra. Llamemos  $K$  al precio de ejercicio,  $T$  a la fecha de vencimiento,  $S_T$  al valor del subyacente en  $T$ . Entonces, si  $S_T > K$ , tiene sentido que el *holder* ejerza la opción para adquirir el subyacente a un precio más interesante; el beneficio de la opción sería exactamente  $S_T - K$ . Por otra parte, si  $S_T < K$ , el poseedor no ejercerá la opción, pues estaría perdiendo dinero. En tal situación, el *payoff* de la opción sería nulo. Aunando ambas ideas, el *payoff* de la opción será

$$\phi(S_T) = \max(S_T - K, 0) = (S_T - K)^+.$$

<sup>5</sup>Algunos tipos de opciones exóticas pueden ser ejercidas en una cantidad finita de tiempos entre la fecha de compra y el día de vencimiento y se clasifican según sus derechos, según si dependen o no del histórico de precios del subyacente... Algunos tipos son, entre otros, las *asiáticas*, las de tipo *lookback* o las *barrier*.

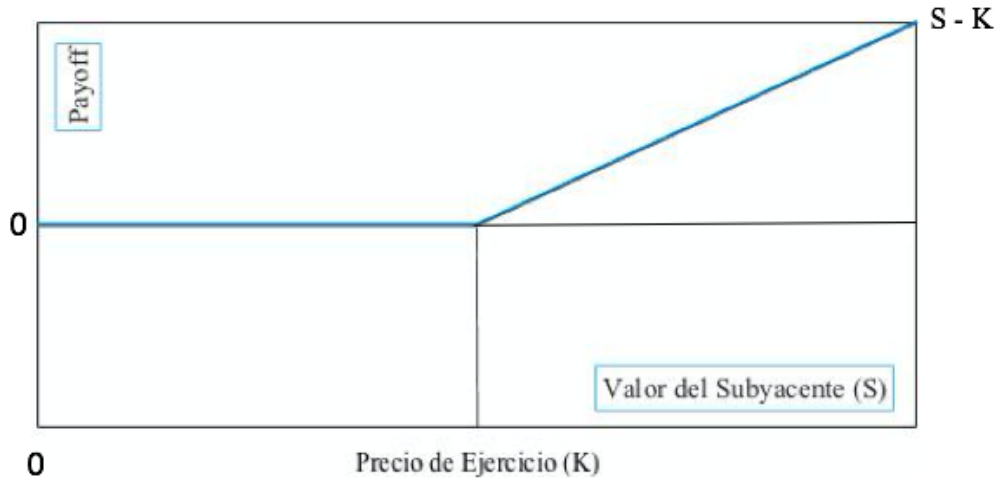


Figura 1.1: Opción europea de compra (*Call*).

A la diferencia entre el coste de la opción en el momento en el que se adquiere y su precio en  $T$  es lo que se conoce como valor del tiempo. Asimismo, al importe en la fecha de vencimiento se lo denomina valor intrínseco de la la opción.

Análogamente al caso de la opción *call* se deduce que el *payoff* de una opción europea de venta (*put*) es

$$\phi(S_T) = \max(K - S_T, 0) = (K - S_T)^+.$$

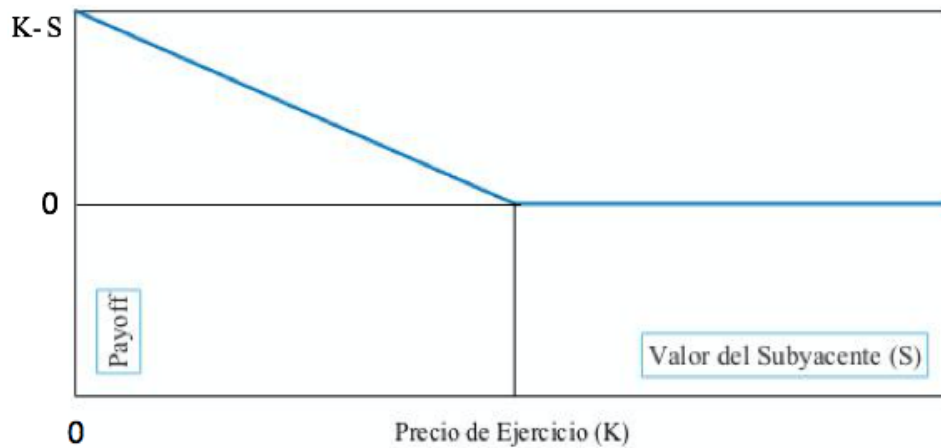


Figura 1.2: Opción europea de venta (*Put*).

Por su parte, los *payoffs* de las opciones exóticas, pese a ser más enrevesados, también admiten formulaciones cerradas <sup>6</sup>.

<sup>6</sup>De forma sucinta, las fórmulas del *payoff* de algunas opciones exóticas son:



En la jerga propia del ámbito financiero se dice que una opción está *at the money* cuando el *strike price* es igual al valor del activo subyacente. Asimismo, si el resultado de ejercer la opción inmediatamente es estrictamente positivo, se declara que la opción está *in the money*. Por su parte, cuando el *payoff* de la opción es nulo, entonces el término utilizado es *out the money*. En ocasiones, cuando existen grandes diferencias entre el valor del subyacente y el precio de ejercicio, se emplean expresiones como *deep in the money* o *deep out the money*.

### 1.3. UN MODELO PARA EL VALOR DE SUBYACENTE.

Remarquemos que nuestro objetivo no es, ni mucho menos, tratar de estimar el valor de los activos en el tiempo. Sin embargo, es conveniente dotarnos de un ligero conocimiento de este ámbito para comprender mejor la dificultad y los límites del problema que nos ocupa, que es el de valorar opciones.

Una constante en el ámbito de la modelización es asumir, *a priori*, ciertas hipótesis que nos permitan dotar de consistencia al modelo al mismo tiempo que nos simplifican su desarrollo; menos hipótesis conllevan modelos más realistas a la par que más complejos. Respecto al precio de los activos, se suele suponer que el *mercado es eficiente*, esto es, que responde de manera inmediata a las nuevas informaciones que atañen a cada activo y que el pasado histórico de su valor se refleja completamente en el precio actual.

Los precios y las rentabilidades se pueden considerar como objetos aleatorios que evolucionan conforme transcurre el tiempo. El marco de trabajo idóneo para estudiar la aleación de azar y tiempo a la que nos enfrentamos lo constituyen los procesos estocásticos y, en particular, las cadenas de Markov<sup>7</sup>. Vamos a formalizar algunos conceptos relativos a esta teoría al tiempo que remitimos al la obra de Øksendal ([17]) para un desarrollo más exhaustivo de los mismos.

Como acabamos de insinuar, los precios de los activos pueden ser modelados mediante el uso de procesos estocásticos en tiempo continuo  $t \in [0, T]$ . Para trabajar con tales procesos, se requiere el marco del *espacio de probabilidad*. Un espacio de probabilidad es una terna,  $(\Omega, \mathcal{F}, P)$ , donde  $\Omega$  es el conjunto de todos los sucesos elementales,  $\mathcal{F}$  es una  $\sigma$ -álgebra que contiene todos los subconjuntos de interés de  $\Omega$  y  $P$  es una función que

- Opción asiática *call fixed strike*:  $\phi(S_T) = \max\{\int_0^T S_t dt - K, 0\}$ .
- Opción de tipo *lookback call floating strike*:  $\phi(S_T) = S_T - \min\{S_t : t \in [0, T]\}$ .
- Opción de tipo *barrier up-and-in call*,  $\phi(S_T) = (S_T - K)^+$  si  $S_T \geq \beta$ ,  $\phi(S_T) = 0$ , en otro caso, donde  $\beta$  es la barrera acordada.

<sup>7</sup>Andrei Andreyevich Markov (14 de junio de 1856, Ryazan, Imperio ruso - 20 de julio de 1922, Leningrado (San Petersburgo), RSFSR (Rusia)) estudió matemáticas en la Universidad de San Petersburgo, en la que después ingresaría. Fue discípulo de Chebyshev e impartió cursos de teoría de la probabilidad hasta 1905, año en el que se retiró definitivamente de la universidad.

asigna una probabilidad a cada  $A \in \mathcal{F}$ .

Debemos asumir, además, que el espacio de probabilidad es *completo*, esto es, que si  $B \subset A \in \mathcal{F}$  y  $P(A) = 0$ , entonces  $B \in \mathcal{F}$ . Dotaremos entonces al espacio probabilístico de lo que se conoce como una *filtración*. Una filtración es una familia  $\mathbb{F} = \{\mathcal{F}_t : t \in [0, T]\}$  de  $\sigma$ -álgebras tales que  $\mathcal{F}_s \subset \mathcal{F}_t \subset \mathcal{F}_T \subset \mathcal{F}$ , si  $0 \leq s \leq t \leq T$ . Cada una de las  $\sigma$ -álgebras de la filtración,  $\mathcal{F}_t$ , puede ser interpretada en el ámbito financiero como la información disponible hasta cierto instante,  $t$ . Por eso, debido a que la información relativa a los eventos pasados se va acumulando a medida que avanza el tiempo, las  $\sigma$ -álgebras son monótonas respecto a  $t$ .

Un espacio de probabilidad dotado de una filtración es lo que se denomina *espacio de probabilidad filtrado* y se representa por  $(\Omega, \mathcal{F}, P, \mathbb{F})$ . Llegados a este punto, es habitual asumir ciertas hipótesis:

- $\mathcal{F}$  es completo respecto a  $P$ ,
- $\mathcal{F}_0$  contiene a todos los conjuntos de  $\Omega$  de probabilidad cero,
- $\mathbb{F}$  es continua por la derecha, esto es,  $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s$ .

Antes de seguir con las definiciones, debemos señalar que éstas están condicionadas al ámbito en el que las vamos a requerir. Por ello, quizás, en otros contextos puedan existir ligeras variaciones de las mismas.

**Definición 1.** (*Proceso estocástico.*)

Un proceso estocástico  $X = \{X_t : t \in [0, T]\}$  es una familia de variables aleatorias definidas en un cierto espacio de probabilidad filtrado  $(\Omega, \mathcal{F}, P, \mathbb{F})$  parametrizadas por la variable  $t$ . Para cada  $\omega \in \Omega$ , se conoce a la función de  $t$ ,  $X_t(\omega)$ , como *trayectoria* o *camino* de  $X$ . Asimismo, se dice que el proceso es  $\mathbb{F}$ -adaptado o *adaptado* respecto a  $\mathbb{F}$  si  $X_t \in \mathcal{F}_t$  para cada  $t$ .

Por otra parte, es habitual modelar el valor del activo subyacente a partir de un tipo especial de procesos estocásticos en los que el comportamiento de  $X$  después del instante  $t$  solo depende de  $X_t$ . A esta clase de procesos estocásticos se los denomina *procesos de Markov* y, respecto a ellos, dice Pierre Brémaud ([6]) que "*the probabilistic dependence on the past is only through the previous date, but this limited amount of memory suffices to produce a great diversity of behaviours. For this reason Markov chains have found applications in many domains*"<sup>8</sup>. Vamos a formalizar este concepto.

<sup>8</sup>"La dependencia probabilística del pasado solo se refleja a través del dato precedente, sin embargo, esta cantidad limitada de memoria es suficiente para reproducir una gran variedad de comportamientos. Por esta razón, las cadenas de Markov son aplicables en multitud de campos".

**Definición 2.** (Proceso de Markov.)

Se dice que un proceso estocástico  $X = \{X_t : t \in [0, T]\}$  verifica la propiedad de Markov respecto a la filtración  $\mathbb{F}$  si

$$\mathbb{E}[f(X_s)|\mathcal{F}_t] = \mathbb{E}[f(X_s)|X_t],$$

para cualquiera que sea la función medible acotada  $f$  y  $s \geq t$ . Un proceso estocástico que cumple la propiedad de Markov es un proceso de Markov.

Otro concepto que es preciso introducir cuando se trabaja en este campo es el de *martingala*.

**Definición 3.** (Martingala.)

Un proceso estocástico  $X = \{X_t : t \in [0, T]\}$  que verifica que

- $X$  es  $\mathbb{F}$ -adaptado,
- $\mathbb{E}[|X_t|] < \infty \forall t \geq 0$ ,
- $\mathbb{E}[X_s|\mathcal{F}_t] = X_t$  casi seguro, si  $s \geq t \geq 0$ ,

es una martingala respecto  $(P, \mathbb{F})$ .

Notemos que la tercera condición que aparece en la definición de martingala goza de una interpretación interesante: la mejor predicción de  $X_s$  que se puede obtener con la información acumulada hasta el instante  $t$  es precisamente  $X_t$ . Por su parte, es el momento de presentar el tipo de proceso más importante dentro de este ámbito: el *proceso de Wiener*<sup>9</sup>, que se utiliza para modelar la evolución del precio de los activos.

**Definición 4.** (Proceso de Wiener.)

Un proceso de Wiener, habitualmente denotado por  $W = \{W_t : t \geq 0\}$ , es un tipo de proceso estocástico en un espacio de probabilidad  $(\Omega, \mathcal{F}, P)$  que goza de las siguientes propiedades:

- $W_0 = 0$ , con probabilidad 1,
- $W_{t+s} - W_t \sim \mathcal{N}(0, s)$ ,
- $W_t - W_s$  es independiente de  $\mathcal{F}_s = \sigma(W_u, u \leq s)$ , si  $s \leq t$ ,

<sup>9</sup>Norbert Wiener (26 de noviembre de 1894, Columbia, Estados Unidos - 18 de marzo de 1964, Estocolmo, Suecia) fue un matemático estadounidense, conocido por ser el fundador de la cibernética, concepto que acuñó en su obra publicada en 1948, *Cibernética o el control y comunicación en animales y máquinas*. A los once años ingresó en la Universidad de Tufts y a los quince se licenció en matemáticas. Un año más tarde, comenzó a estudiar zoología en Harvard pero, otro año después, decidió trasladarse a la Universidad de Cornell para estudiar filosofía. Más tarde volvería a Harvard, donde se doctoró en 1912 con una tesis sobre lógica matemática.

- para cada  $\omega \in \Omega$ , la función  $X_t(\omega)$  es continua respecto de  $t$ .

La definición que hemos presentado es la formulación moderna de la generalización de la idea de paseo aleatorio. Este concepto aparece en la naturaleza en multitud de ámbitos completamente distintos. Supongamos que una determinada partícula se encuentra situada, en el instante  $t$ , en un punto  $x$  de la recta real y que, además, esa partícula salta una distancia  $h$  hacia adelante o hacia atrás, con igual probabilidad, cada vez que transcurre  $\tau$  tiempo. Cuando se hace  $h \rightarrow 0$  y  $\tau \rightarrow 0$ , el paseo aleatorio se convierte en un camino continuo que es el proceso de Wiener.

La clave del poder de esta teoría reside en que no hemos concretado qué tipo de partícula es la que estamos considerando. Así, el milagro se obra al revelar que el modelo funciona para una infinitud de posibilidades. Por ejemplo, el botanista Robert Brown descubrió, en 1827 <sup>10</sup>, que las partículas microscópicas de polen suspendidas en un medio fluido se desplazaban en movimientos aleatorios sin un patrón aparente. Debido a su hallazgo, a este fenómeno se lo denomina *movimiento browniano*. En la práctica, los conceptos de proceso de Wiener y movimiento browniano se utilizan indistintamente.

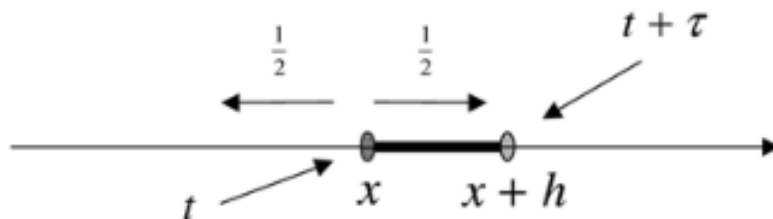


Figura 1.3: Paseo aleatorio simétrico. (Imagen tomada de [18]).

En realidad, no debe ser Wiener quien se lleve todo el mérito del asunto pues hubo varios avances precedentes a su obra debidos a otros insignes matemáticos y no nos gustaría deambular por los pasajes del camino aleatorio sin, al menos, mencionarlos. El primero de ellos es Louis Bachelier <sup>11</sup>. Él fue el primero en postular que los precios de los activos siguen un movimiento browniano pues la información que los hace variar, tanto al alza como a la baja, llega a los mismos de manera aleatoria. El otro al que nos gustaría mencionar es, posiblemente, el científico más célebre del siglo XX, Albert Einstein <sup>12</sup>.

<sup>10</sup>Anteriormente, en 1785, Jan Ingenhousz había descrito el mismo proceso valiéndose de partículas de carbón en alcohol. Es más, Tito Lucrecio en su obra *De rerum natura*, del siglo primero antes de Cristo, detalla la armonía y la belleza que residen en algunos sucesos cotidianos que podrían considerarse como paseos aleatorios.

<sup>11</sup>Louis Jean-Baptiste Alphonse Bachelier (11 de marzo de 1870, Le Havre, Francia - 28 de abril de 1946, Saint Servan sur Mer, Francia) fue un matemático francés. Defendió su tesis doctoral, *Théorie de la Spéculation* ([2]), en 1900, bajo la tutela del insigne Henri Poincaré, quien miraba con recelo la intromisión de las matemáticas en los asuntos financieros. Su tesis no fue lo suficientemente valorada hasta la década de los sesenta cuando Benoit Mandelbrojt, el artifice de los fractales, la rescató del olvido.

<sup>12</sup>Albert Einstein (14 de marzo de 1879, Ulm, Imperio alemán - 18 de abril de 1955, Princeton, Estados Unidos)

Resulta que sus trabajos sobre el movimiento browniano culminaron en el Premio Nobel que le fue concedido en 1905.

Como ya se ha podido intuir, estamos tratando con problemas en los que la suerte es protagonista. Esto resulta evidentemente necesario pues, si encontrásemos un modelo determinista válido para la cotización del subyacente, podríamos prever con total seguridad el comportamiento de su precio. Por ello, el sendero más fácil para empezar a caminar es el de considerar procesos que estén compuestos por una parte determinista y otra parte estocástica, esto es,

$$dX_t = \alpha(X_t, t)dt + \beta(X_t, t)dW_t.$$

Lo anterior es un ejemplo de lo que a su vez se denomina una *ecuación diferencial estocástica*. Esta expresión, junto con el hecho de que  $X_{t_0} = X_0$ , es un formato abreviado de la llamada *forma integral de la ecuación diferencial estocástica*, que es

$$X_t = X_0 + \int_{t_0}^t \alpha(X_s, s)ds + \int_{t_0}^t \beta(X_s, s)dW_s.$$

Al término  $\alpha(X_t, t)$  se lo denomina *tendencia*, *deriva* o *drift* mientras que a  $\beta(X_t, t)$  se lo conoce como *volatilidad*. La solución del modelo,  $X_t$ , es lo que se conoce como un *proceso de Ito*<sup>13</sup>. Un proceso de Wiener es un caso particular de proceso de Ito pues, al tomar  $\alpha \equiv 0$  y  $\beta \equiv 1$ , se obtiene la ecuación trivial  $dX_t = dW_t$ . El símbolo integral alude a la denominada *integral de Ito*. Sin embargo, el estudio de esta teoría no es objeto de este trabajo y, por tanto, no profundizaremos en este ámbito. Es posible encontrar suficiente documentación al respecto en [17].

El modelo más simple para el valor del subyacente es el que asume que los coeficientes  $\alpha$  y  $\beta$  son constantes. Esto ciertamente se encuentra bastante alejado de la realidad pero, como veremos, nos proporciona un marco de trabajo suficientemente amplio como para obtener algunos resultados interesantes. De ahora en adelante nos haremos partícipes de la notación más extendida que es la que llama  $\mu$  al *drift* y  $\sigma$  a la volatilidad.

Volviendo al asunto de encontrar un modelo para el precio del subyacente, reparemos en que lo realmente importante es la variación relativa de su precio; un cambio de 1€ es mucho más significativo si el precio de la acción era de 15€ que si era de 75€. Nace así el concepto de *tasa de retorno*, cuyo sentido matemático viene dado por la expresión  $dS/S$ .

---

fue un físico alemán de origen judío que a la postre sería nacionalizado suizo y estadounidense. En 1905 presentó su *teoría de la relatividad restringida* de la que se deduce la ecuación  $E = mc^2$ . Más tarde, en 1915, publicó la *teoría de la relatividad general*, la cual revolucionó totalmente las bases de la física sentadas hasta la fecha.

<sup>13</sup>Kiyoshi Ito (7 de septiembre de 1915, Hokusei, Japón - 10 de noviembre de 2008, Kioto, Japón) estudió matemáticas en la Universidad de Tokio, en la que se graduó a los 23 años. Sus contribuciones se centran en la teoría de la probabilidad y los procesos estocásticos. De hecho, su trabajo se llama ahora *cálculo de Ito*. En 2006 fue galardonado con el premio Carl Friedrich Gauss.

El modelo que resulta goza de nombre propio, *movimiento browniano geométrico*, y toma el aspecto siguiente,

$$dS_t = \mu S_t dt + \sigma S_t dW_t.$$

La aportación determinista no es sino la ecuación diferencial ordinaria

$$\frac{dS}{dt} = \mu S,$$

cuya solución, que se calcula fácilmente por integración directa, es

$$S_t = S_0 e^{\mu(t-t_0)}.$$

Reparemos en que la fórmula anterior es la celebrada capitalización continua que se utiliza para calcular el valor futuro de una cierta cantidad,  $S_0$ , cuyos intereses, al  $100\mu\%$ , se van acumulando instantáneamente. Ésta es, a su vez, una generalización de otra ecuación sobradamente conocida: la del interés simple.

Para poder seguir avanzando en la extracción de resultados en este terreno necesitamos enunciar el que es posiblemente el resultado fundamental de la teoría de procesos estocásticos que viene a ser una versión probabilística de la regla de la cadena.

**Lema 1.** (*Lema de Ito, versión general*)

Supongamos que un proceso estocástico  $X_t$  sigue  $dX_t = \alpha(X_t, t)dt + \beta(X_t, t)dW_t$ . Sea  $g(x, t)$  una aplicación tal que  $\partial g/\partial x$ ,  $\partial^2 g/\partial x^2$  y  $\partial g/\partial t$  son continuas. Entoces  $Y_t = g(X_t, t)$  verifica la ecuación diferencial estocástica

$$dY_t = \left( \frac{\partial g}{\partial t} + \frac{\partial g}{\partial x} \alpha + \frac{1}{2} \frac{\partial^2 g}{\partial x^2} \beta^2 \right) dt + \frac{\partial g}{\partial x} \beta dW_t$$

donde tanto todas las derivadas de  $g$  como  $\alpha$  y  $\beta$  dependen de los argumentos  $(X_t, t)$ .

Aunque no tenemos intención de dar una prueba rigurosa del resultado precedente (Ver [17]), si que vamos a realizar algunas consideraciones heurísticas que nos permitan entender su significado al menos de forma intuitiva. Como ya hemos comentado antes, si  $\beta \equiv 0$ , la ecuación es determinista y es viable resolverla valiéndose de los métodos analíticos usuales. Es más, es posible calcular de manera sencilla el diferencial de  $Y$ ,

$$dY_t = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t = \left( \frac{\partial g}{\partial t} + \alpha \frac{\partial g}{\partial x} \right) dt.$$

Supongamos ahora que  $\beta \neq 0$ . Si procediesemos del mismo modo, obtendríamos que

$$dY_t = \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t = \left( \frac{\partial g}{\partial t} + \alpha \frac{\partial g}{\partial x} \right) dt + \frac{\partial g}{\partial x} \beta dW_t,$$

pero esa fórmula dista bastante de ser correcta. Fijémonos en el desarrollo de Taylor de  $Y$ . Su diferencial se obtendría seleccionando los términos lineales respecto de  $dt$  y  $dW_t$ .

$$Y(X_t, t) = Y(X_0, 0) + \frac{\partial g}{\partial t} dt + \frac{\partial g}{\partial x} dX_t + \frac{1}{2} \left( \frac{\partial^2 g}{\partial t^2} (dt)^2 + 2 \frac{\partial^2 g}{\partial t \partial x} dt dX_t + \frac{\partial^2 g}{\partial x^2} (dX_t)^2 \right) + \dots$$

Está claro que los términos en  $dt$  y  $dX_t$  van a aparecer en el diferencial de  $Y_t$ . Sin embargo, y he aquí el hecho sorprendente, al examinar a fondo el término  $(dX_t)^2$  nos va a aparecer algún elemento que habíamos pasado por alto.

$$(dX_t)^2 = (\alpha dt + \beta dW_t)^2 = \alpha^2 (dt)^2 + 2\alpha\beta dt dW_t + \beta^2 (dW_t)^2.$$

Mientras que los dos primeros elementos no son lineales respecto de  $dt$  ni de  $dX_t$ , el tercero resulta ser

$$\beta^2 (dW_t)^2 = \beta^2 dt.$$

De forma intuitiva podríamos decir que es debido a la fórmula  $dW_t \sim \sqrt{dt} \mathcal{N}(0, 1)$ , que asigna el valor  $\sqrt{dt}$  como desviación típica de  $dW_t$ . De esta manera, el diferencial de  $Y_t$  a través de las trayectorias  $dt$  y  $dX_t$  viene dado, efectivamente, por la fórmula de Ito.

Como una situación particular se presenta el resultado siguiente, que no es más que una aplicación del lema de Ito al caso que nos ocupa.

**Lema 2.** (*Lema de Ito, aplicación al movimiento browniano geométrico*)

Supongamos que un proceso estocástico  $S_t$  sigue  $dS_t = \mu S_t dt + \sigma S_t dW_t$ . Entonces,  $\log(S_t)$  verifica la ecuación diferencial estocástica

$$d \log(S_t) = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dW_t.$$

Para la situación del lema previo se tiene que la esperanza de  $Y_t = \log S_t$  satisface la parte determinista, esto es,

$$\frac{d\mathbb{E}[Y_t]}{dt} = \mu - \frac{1}{2} \sigma^2.$$

Por lo tanto, si  $Y_{t_0} = S_0$ , la esperanza de  $Y_t$  es

$$\mathbb{E}[\log(S_t)] = \log(S_0) + \left( \mu - \frac{1}{2} \sigma^2 \right) (t - t_0).$$

Asimismo, se puede ver a partir de la ecuación diferencial que resulta para  $\mathbb{E}[Y_t^2]$  que la varianza de  $Y_t$  es  $\sigma^2(t - t_0)$ . Por lo tanto, aglutinando ambas ideas, se sigue que  $Y_t$  sigue una distribución normal con función de densidad

$$f(y) = \left(\sigma\sqrt{2\pi(t-t_0)}\right)^{-1} \exp\left\{-\frac{(y-y_0 - (\mu - \frac{1}{2}\sigma^2)(t-t_0))^2}{2\sigma^2(t-t_0)}\right\}.$$

Deshaciendo el cambio  $Y_t = \log(S_t)$ , y sin dejar de prestar atención a que  $\frac{dY_t}{dS_t} = \frac{1}{S}$ , se llega a que la densidad de  $S_t$  viene dada por

$$p(s) = \left(s\sigma\sqrt{2\pi(t-t_0)}\right)^{-1} \exp\left\{-\frac{(\log(s/S_0) - (\mu - \frac{1}{2}\sigma^2)(t-t_0))^2}{2\sigma^2(t-t_0)}\right\}.$$

Esta es la densidad de la *distribución lognormal* y se dice, por tanto, que el precio del subyacente,  $S_t$  se distribuye lognormalmente cuando se asume que sigue un movimiento browniano geométrico.

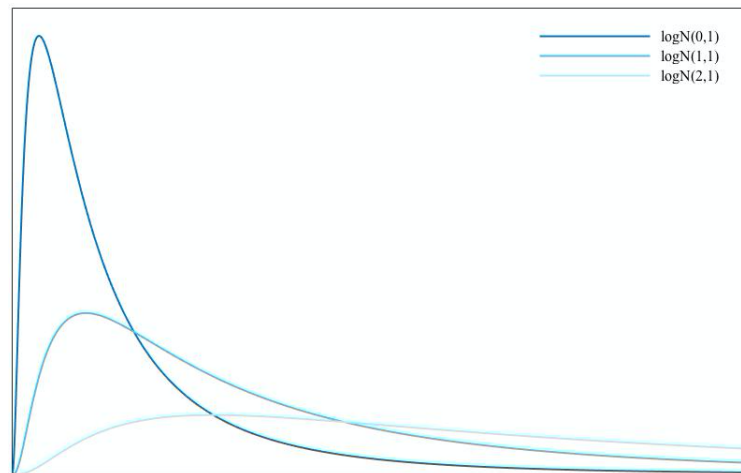


Figura 1.4: Distribución lognormal para distintos valores de  $\mu$ .

Advirtamos que el hecho de aceptar que tanto  $\mu$  como  $\sigma$  sean constantes no parece demasiado realista. Por ese motivo, se han desarrollado otros modelos que consideran que éstos varían con el tiempo.

El problema de estimar la volatilidad,  $\sigma$ , es particularmente complicado. Uno podría pensar en acudir a los valores históricos de  $\sigma$  para tratar de entender cómo evoluciona (volatilidad histórica o realizada). Sin embargo, no está claro que  $\sigma$  sea independiente de cada instante de tiempo en el que se mide y, por lo tanto, si realmente es útil para estimar su valor futuro.

Una idea muy original resulta al percatarse de que, aunque nosotros no podamos cuantificar el valor de  $\sigma$ , el mercado, como ente abstracto, sí que conoce tal valor pues,



como veremos más adelante, el valor de las opciones se puede calcular gracias a la fórmula de Black-Scholes. Por lo tanto, puesto que el resto de variables que aparecen en tal fórmula son fácilmente calculables, junto con el hecho de que el precio de una opción *call* crece monótonamente cuando lo hace la volatilidad, existe una correspondencia biyectiva entre  $\sigma$  y el valor de la opción. De este modo, al conocer el valor de la opción en un tiempo determinado se puede deducir el valor de la volatilidad en ese preciso instante. A esta manera de proceder se la denomina *modelo de volatilidad implícita*.

## 1.4. LA ECUACIÓN DE BLACK-SCHOLES Y SU ENTORNO.

Como ya hemos venido anunciando, el asumir ciertas hipótesis más o menos restrictivas es una pauta inherente a las entretelas de la modelación científica. En el marco de la ecuación de Black-Scholes, una de estas hipótesis pasa por un concepto financiero importante: el arbitraje.

En un sentido bastante amplio, podemos decir que el *arbitraje* denota cualquier situación que permite obtener un beneficio sin asumir ningún riesgo. Particularmente, se puede pensar en la práctica de comprar y vender el mismo producto en distintos mercados para aprovechar las pequeñas diferencias de precio.

Si pensamos en un banco que ofrece un cierto tipo de interés menor que el retorno que un inversor sensato obtendría haciendo uso de una oportunidad de arbitraje, resulta evidente que el banco no iba a resultar ser la opción más atractiva. Por ello, éste a su vez se vería forzado a elevar la cuantía que paga a los inversores a cambio de la cesión de su dinero. Esta elasticidad producida por la concatenación de las leyes de la oferta y de la demanda provoca que, en la práctica, las oportunidades de arbitraje se desvanezcan muy rápidamente una vez que son descubiertas.

Ahora que nos hemos podido hacer una ligera idea de lo que es el arbitraje, podemos volver al asunto de la ecuación de Black-Scholes y sus hipótesis de partida. Lo habitual en la literatura al respecto de este tema es presentar tales hipótesis de manera enumerativa. En aras de la claridad expositiva, prodecemos del mismo modo.

Las hipótesis que conducen al modelo de Black-Scholes son:

- El mercado está libre de fricción, esto es, se supone que no existen costes de transacción, ni impuestos, ni tasas, que afecten a la compra o la venta de los activos y los derivados. Más aún, se entiende que toda la información se distribuye de manera simétrica e inmediata entre todas las partes que intervienen en el mercado.
- No existen oportunidades de arbitraje.
- El activo subyacente sigue un movimiento browniano geométrico.
- La volatilidad  $\sigma$  es constante y conocida.

- Es posible comprar y vender cualquier cantidad de activos que se desee <sup>14</sup>.
- La tasa de interés libre de riesgo es  $r > 0$ , constante y conocida.
- El valor de las opciones solo depende del valor del activo subyacente y del tiempo.

En lo que sigue denotaremos por  $V = V(S, t)$  a la función que cuantifica el valor de las opciones y omitiremos, a menudo, la dependencia respecto de  $t$  en  $S_t$  y en  $W_t$ . En tales condiciones, a partir de la hipótesis que afirma que

$$dS_t = \mu S_t dt + \sigma S_t dW,$$

es posible hallar el diferencial de  $V$  haciendo uso del lema de Ito, esto es,

$$dV = \left\{ \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right\} dt + \sigma S \frac{\partial V}{\partial S} dW.$$

El procedimiento original presentado por Black y Scholes ([5]) consiste en lo siguiente: se construye una cartera de inversión compuesta por la opción que queremos valorar y una cantidad  $-\Delta$  de activos. El valor del *portfolio* es, por tanto,

$$\Pi = V - \Delta S.$$

No nos gustaría que pasase desapercibido el hecho de que acabamos de realizar la estrategia de cobertura (*hedging*) que es, *per se*, una de las justificaciones de la existencia de estos derivados que son las opciones.

Pensemos ahora en la variación del valor de la cartera en un incremento  $dt$ . Si  $\Delta$  permanece constante y sin entrar en otras consideraciones técnicas más profundas, se tendría que

$$d\Pi = dV - \Delta dS.$$

Sustituyendo las expresiones de las que disponemos para  $dV$  y  $dS$ , la ecuación precedente toma el aspecto de

$$d\Pi = \left\{ \frac{\partial V}{\partial t} + \mu S \frac{\partial V}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} - \mu S \Delta \right\} dt + \sigma S \left( \frac{\partial V}{\partial S} - \Delta \right) dW.$$

<sup>14</sup>Una práctica habitual en el campo de las finanzas consiste en que los inversores vendan activos que han sido tomados en préstamo de un tercero. A esta estrategia se la denomina *venta en corto* y explica el hecho de que un inversor posea, en un instante determinado, una cantidad negativa de activos.

Si no hemos olvidado el hecho de que  $\Delta$  es una cantidad sobre la que no se ha impuesto ninguna restricción, nada nos impide *a priori* elegir

$$\Delta = \frac{\partial V}{\partial S}.$$

El valor  $\Delta$  juega un papel importante en este marco, pues se puede considerar como una suerte de medida de correlación entre el valor de la opción y el del activo subyacente. Volviendo al punto donde estábamos, notemos que, en tal situación, lo que se consigue es eliminar la componente estocástica de la variación de  $\Pi$ , esto es, la que afecta al término  $dW$ . De esta manera, la evolución del valor de la cartera es totalmente determinista y viene dada por la relación

$$d\Pi = \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right\} dt.$$

Fijémonos en que el *drift*  $\mu$  se ha esfumado lo cual viene a significar que el valor de una opción es independiente de lo rápido o despacio que evolucione el precio del subyacente.

Llega ahora el momento de aplicar la hipótesis que concierne a la inexistencia de oportunidades de arbitraje. Si el inversor deseara emplear el valor de su cartera,  $\Pi$ , en un producto libre de riesgo a un tipo  $r$ , lo que obtendría al cabo de un cierto tiempo,  $dt$ , sería  $r\Pi dt$ . Al no permitir oportunidades de arbitraje, esta expresión ha de ser igual a la que obtuvimos antes pues, en otro caso, un arbitrajista podría obtener un beneficio instantáneo sin riesgo alguno. De este modo, se tiene que

$$r\Pi dt = \left\{ \frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} \right\} dt,$$

y sustituyendo  $\Pi$  por su valor explícito, esto es,

$$\Pi = V - S \frac{\partial V}{\partial S},$$

se llega a la célebre *ecuación de Black-Scholes*:

$$\boxed{\frac{\partial V}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0}$$

Cabe mencionar que, además, se puede entender la ecuación de Black-Scholes como el núcleo de un operador lineal y continuo. Tal operador, al que llamaremos *operador de Black-Scholes*, viene dado por la expresión:

$$\mathcal{L}_{BS} = \frac{\partial}{\partial t} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2}{\partial S^2} + rS \frac{\partial}{\partial S} - r.$$

Se le puede conferir a este operador una interpretación financiera: es la diferencia entre los pagos correspondientes a la estrategia de cobertura y el interés libre de riesgo que otorga un banco.

Es posible encontrar una deducción de la ecuación de Black-Scholes desde el ámbito de la teoría de la probabilidad en [1] y en [13].

## 1.5. EL CASO MÁS SIMPLE.

Nos gustaría finalizar este capítulo resolviendo explícitamente el caso de la opción *vanilla* europea de compra. En cambio, para el resto de tipos de opciones no podemos dar una fórmula explícita y se debe, por tanto, recurrir a métodos numéricos.

Partamos de la ecuación de Black-Scholes con un *payoff* conocido: el de la opción europea de compra,

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \\ V(S_T, T) = \phi(S_T) = \max(S_T - K, 0) = (S_T - K)^+. \end{cases}$$

El procedimiento habitual para resolver este problema involucra varios cambios de variables para transformar el problema de Black-Scholes en otro sobradamente conocido: el de la ecuación del calor.

El primero de tales cambios de variables es  $\tau = T - t$ . De esta manera, lo que hacemos es invertir la secuencia temporal del problema. De esta manera,

$$\begin{cases} -\frac{\partial V}{\partial \tau} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + r \left( S \frac{\partial V}{\partial S} - V \right) = 0, \\ V(S, 0) = \phi(S). \end{cases}$$

El siguiente cambio de variable que debemos realizar es  $x = \log(S)$ . Así, derivamos haciendo uso de la regla de la cadena y de la fórmula de derivación del producto de funciones, esto es,

- $\frac{\partial V}{\partial S} = \frac{\partial V}{\partial x} \frac{\partial x}{\partial S} = \frac{\partial V}{\partial x} S^{-1} = \frac{\partial V}{\partial x} e^{-x};$
- $S \frac{\partial V}{\partial S} = e^x \left( \frac{\partial V}{\partial x} e^{-x} \right) = \frac{\partial V}{\partial x};$

- $\frac{\partial^2 V}{\partial S^2} = \frac{\partial}{\partial S} \left( \frac{\partial V}{\partial x} e^{-x} \right) = \frac{\partial}{\partial x} \frac{\partial x}{\partial S} \left( \frac{\partial V}{\partial x} e^{-x} \right) = e^{-x} \frac{\partial}{\partial x} \left( e^{-x} \frac{\partial V}{\partial x} \right) = e^{-2x} \frac{\partial^2 V}{\partial x^2} - e^{-2x} \frac{\partial V}{\partial x};$
- $S^2 \frac{\partial^2}{\partial S^2} = e^{2x} \left\{ e^{-2x} \frac{\partial^2 V}{\partial x^2} - e^{-2x} \frac{\partial V}{\partial x} \right\} = \frac{\partial^2 V}{\partial x^2} - \frac{\partial V}{\partial x}.$

Si ahora reemplazamos en la ecuación, obtenemos que,

$$-\frac{\partial V}{\partial \tau} + \frac{1}{2} \sigma^2 \frac{\partial^2 V}{\partial x^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial V}{\partial x} - rV = 0.$$

Es momento de realizar otro cambio de variable:

$$W(x, \tau) = e^{r\tau} V(x, \tau) \implies V(x, \tau) = e^{-r\tau} W(x, \tau).$$

Derivando otra vez de forma adecuada,

- $\frac{\partial V}{\partial \tau} = -r e^{-r\tau} W(x, \tau) + e^{-r\tau} \frac{\partial W}{\partial \tau};$
- $S \frac{\partial V}{\partial x} = e^{-r\tau} \frac{\partial W}{\partial x};$
- $S^2 \frac{\partial^2 V}{\partial x^2} = e^{-r\tau} \frac{\partial^2 W}{\partial x^2}.$

Y, otra vez, llevándolo a la ecuación y realizando las simplificaciones precisas,

$$-\frac{\partial W}{\partial \tau} + \frac{1}{2} \sigma^2 \frac{\partial^2 W}{\partial x^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial W}{\partial x} = 0.$$

Nos queda aún un cambio de variable más por llevar a cabo. Vamos con ello.

$$z = x + \left( r - \frac{1}{2} \sigma^2 \right) \tau.$$

Sea entonces  $U(z, \tau) = W(x, \tau)$ . Volvemos a requerir ciertas derivaciones.

- $\frac{\partial W}{\partial x} = \frac{\partial U}{\partial z} \frac{\partial z}{\partial x} = \frac{\partial U}{\partial z};$
- $\frac{\partial^2 W}{\partial x^2} = \frac{\partial^2 U}{\partial z^2};$
- $\frac{\partial W}{\partial \tau} = \frac{\partial U}{\partial \tau} + \frac{\partial U}{\partial z} \frac{\partial z}{\partial \tau} = \frac{\partial U}{\partial \tau} + \left( r + \frac{1}{2} \sigma^2 \right) \frac{\partial U}{\partial z}.$

Si ahora sustituimos en la ecuación que teníamos,

$$-\frac{\partial U}{\partial \tau} - \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial U}{\partial z} + \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial z^2} + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial U}{\partial z} = 0,$$

esto es,

$$\frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial z^2}.$$

Por supuesto,  $z \in \mathbb{R}$ . Hemos obtenido la ecuación del calor. Estudiemos ahora de qué manera afectan estos cambios de variable al *payoff* de la opción europea de compra que, como dijimos en una de las secciones previas, es

$$\phi(S) = \max(S - K, 0) = (S - K)^+.$$

Aplicando los distintos cambios de variables, se va obteniendo que

$$V(x, 0) = \phi(e^x) = \max(e^x - K, 0);$$

$$W(x, 0) = V(x, 0) = \phi(e^x) = \max(e^x - K, 0);$$

$$U(z, 0) = W(x, 0) = \phi\left(\exp\left(z - \left(r - \frac{1}{2}\sigma^2\tau\right)\right)\right) =: \phi(z).$$

De esta manera, el problema que ahora queremos resolver es

$$\begin{cases} \frac{\partial U}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 U}{\partial x^2}, & z \in \mathbb{R}, \\ U(z, 0) = \phi(z). \end{cases}$$

Pero el problema de la ecuación del calor está muy estudiado y conocemos su solución (ver, por ejemplo, [18]). Por tanto, afirmamos que

$$U(z, \tau) = \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} \exp\left(\phi(\xi) \frac{-|z - \xi|^2}{2\sigma^2\tau}\right) d\xi.$$

Si ahora deshacemos el último cambio de variable, llegamos a que

$$W(x, \tau) = \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} \phi(\xi) \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi.$$

Reescribimos la expresión anterior haciendo uso de la forma explícita del *payoff*,

$$\begin{aligned} W(x, \tau) &= \frac{1}{\sigma\sqrt{2\pi\tau}} \int_{-\infty}^{\infty} \max(e^{\xi} - K, 0) \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi \\ &= \int_{\log K}^{\infty} \frac{1}{\sigma\sqrt{2\pi\tau}} (e^{\xi} - K) \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi. \end{aligned}$$

Si evaluamos las dos integrales siguientes, habremos acabado. Por una parte, tenemos que,

$$\begin{aligned} &\int_{\log K}^{\infty} \frac{1}{\sigma\sqrt{2\pi\tau}} e^{\xi} \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi \\ &= \exp\left(x + \left(r - \frac{1}{2}\sigma^2\right)\tau + \frac{1}{2}\sigma^2\tau\right) \mathcal{N}\left(\frac{x + \left(r - \frac{1}{2}\sigma^2\right)\tau - \log K + \sigma^2\tau}{\sqrt{\tau}\sigma}\right) \\ &= e^{tr} S \mathcal{N}\left(\frac{\log \frac{S}{K} + \left(r + \frac{1}{2}\sigma^2\tau\right)}{\sqrt{\tau}\sigma}\right), \end{aligned}$$

donde  $\mathcal{N}$  representa la función de distribución normal. Nótese que  $x = \log S$ .

Aún nos queda la otra integral, esto es,

$$\begin{aligned} &\int_{\log K}^{\infty} \frac{1}{\sigma\sqrt{2\pi\tau}} K \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi \\ &= K \int_{\log K}^{\infty} \frac{1}{\sigma\sqrt{2\pi\tau}} \exp\left(\phi(\xi) \frac{-|x + (r - \frac{1}{2}\sigma^2) - \xi|^2}{2\sigma^2\tau}\right) d\xi \\ &= K \mathcal{N}\left(\frac{x + \left(r - \frac{1}{2}\sigma^2\right)\tau - \log K}{\sqrt{\tau}\sigma}\right) \\ &= K \mathcal{N}\left(\frac{\log \frac{S}{K} + \left(r - \frac{1}{2}\sigma^2\tau\right)}{\sqrt{\tau}\sigma}\right). \end{aligned}$$

Finalmente, se tiene que el valor de una opción europea de compra (*call*) viene dado por la expresión,

$$\begin{aligned} C(S, \tau) &= e^{-tr} \left\{ e^{tr} S \mathcal{N}\left(\frac{\log \frac{S}{K} + \left(r + \frac{1}{2}\sigma^2\tau\right)}{\sqrt{\tau}\sigma}\right) - K \mathcal{N}\left(\frac{\log \frac{S}{K} + \left(r - \frac{1}{2}\sigma^2\tau\right)}{\sqrt{\tau}\sigma}\right) \right\} \\ &= S \mathcal{N}(d_1) - K e^{-tr} \mathcal{N}(d_2), \end{aligned}$$

con

$$\begin{cases} d_1 = \frac{\log \frac{S}{K} + \left( r + \frac{1}{2} \sigma^2 \tau \right)}{\sqrt{\tau} \sigma}, \\ d_2 = d_1 - \sqrt{\tau} \sigma. \end{cases}$$

Existe un concepto financiero llamado *paridad call-put* que relaciona los valores de las opciones de compra y de venta. Por concisión, no incluiremos aquí una descripción de esta idea, la cual puede encontrarse en [4] y en [23], y nos limitaremos a señalar que el valor de una opción europea de venta se deduce de lo anterior y viene dado por

$$P(S, t) = K e^{-r\tau} \mathcal{N}(-d_2) - S \mathcal{N}(-d_1).$$

## 1.6. REFERENCIAS BIBLIOGRÁFICAS.

Debemos reparar en que a lo largo de estas páginas hemos tratado aspectos bastante diversos como son los conceptos propios del ámbito financiero o los procesos estocásticos. Describiremos en las líneas posteriores los aspectos más reseñables de la bibliografía que hemos utilizado.

- El libro de Berk y DeMarzo ([4]), pese a no estar destinado en absoluto a matemáticos, hace una exposición bastante clara y global de los conceptos propios del campo de las finanzas corporativas. En particular, el capítulo 20 versa sobre opciones financieras.
- Por su parte, no podemos dejar de mencionar la obra de Wilmott ([23]), la cual es de lectura ineludible para aquellos que, procediendo del ámbito de las ciencias o las ingenierías, desean introducirse en el terreno de los derivados financieros. La Parte 1 expone de manera clara a la par que rigurosa los conceptos y las ideas de la teoría de valoración de opciones. Por su parte, la Parte 2, aunque no trata con el método de los elementos finitos, sí que presenta la aplicación de algunas técnicas numéricas al problema de valoración.
- Asimismo, el libro de Salsa ([18]) estudia, en el capítulo segundo dedicado a la ecuación del calor, el problema de Black-Scholes. Además, presenta de forma suficientemente rigurosa pero sin perder la claridad expositiva el concepto de movimiento browniano.
- También merece ser mencionado el libro de Øksendal ([17]) que estudia las ecuaciones diferenciales estocásticas. Aunque este tema queda por completo fuera del alcance de este trabajo, en él pueden encontrarse con el debido rigor todas las demostraciones relativas al lema de Ito, integrales estocásticas... Además, en el



capítulo primero, se abordan aspectos de matemática financiera, eso si, desde el punto de vista de los procesos estocásticos.

- El capítulo 1 del libro de Hilber, Reichmann, Schwab y Winter ([13]) plantea de una forma muy concisa los aspectos y las definiciones más importantes de la matemática financiera así como de los procesos estocásticos.
- Finalmente, el libro de Seydel ([20]) da, en el capítulo 1, una visión general del problema de valoración de opciones que resulta ser bastante ilustrativa. Además, describe la cuestión de la *geometría de las opciones* de una manera fácilmente comprensible para el lector.







# 2

## EL MÉTODO DE LOS ELEMENTOS FINITOS.

---

*"Si la gente no cree que las matemáticas son sencillas,  
es solo porque no se da cuenta de lo complicada que es la vida."*

— John von Neumann.

**Resumen:** Se comienza presentando el problema de la formulación variacional de un problema de contorno. A continuación, se compendian algunas definiciones y conceptos del análisis funcional para, más adelante, estudiar varios resultados importantes. Finalmente, se exponen los fundamentos teóricos que avalan el uso de técnicas de aproximación, en particular, del método de los elementos finitos, a tales problemas variacionales.



## 2.1. PANORAMA GENERAL A TRAVÉS DE UN PROBLEMA I-DIMENSIONAL.

Existe una eterna discusión al respecto de si son las matemáticas las que se desarrollan por sí mismas o si, en cambio, evolucionan para dar respuesta a los problemas que van siendo planteados por otras disciplinas. Esta confrontación se puede apreciar incluso entre la manera de afrontar los problemas "matemáticos" de cada una de las civilizaciones antiguas; los babilonios desarrollaron una teoría esencialmente práctica mientras que los griegos se preocuparon por demostrar la veracidad de los enunciados relativos a números y figuras.

Jonh von Neumann <sup>1</sup>, sin duda, uno de los matemáticos más importantes del siglo XX, aseveró y defendió que la mayor parte de las ideas matemáticas habían sido concebidas sin valorar siquiera la existencia de una utilidad de la que dotarlas. Sin embargo, también creía que, transcurrido un tiempo, los modelos desarrollados por los matemáticos acabarían por dar respuesta a situaciones e interrogantes de los ámbitos más diversos del conocimiento.

Valga de ejemplo lo acontecido al respecto de los números complejos: hasta finales del siglo XVIII, éstos fueron usados con recelo y desconfianza pues no respondían al problema de la medida de magnitudes. En cambio, de la mano de Euler y Gauss se perdió esa preocupación por la naturaleza de tales elementos lo cual culminó con la prueba del *Teorema Fundamental del Álgebra* en 1799.

Por hechos como el anterior, Von Neumann finalizó una de sus más famosas conferencias, *The Role of Mathematics in Science and Society*, afirmando que el progreso de la ciencia será mucho mayor si los investigadores se dejan guiar por los derroteros de la belleza intelectual que si se limitan al ámbito de lo estrictamente útil. Quizás de esta manera estaba envolviendo de un halo "científico" la célebre cita de Dostoyevski, "*la belleza salvará al mundo*".

En particular, algo parecido ocurrió con el asunto que nos va a ocupar en este capítulo. Por una parte, las bases del método de los elementos finitos fueron asentadas por Boris Galerkin <sup>2</sup> en 1915 y fue en 1943 cuando Richard Courant <sup>3</sup> propuso formalmente el

<sup>1</sup>John von Neumann (28 de diciembre de 1903, Budapest, Imperio austro-húngaro - 8 de febrero de 1957, Washington D.C., Estados Unidos) fue un matemático que realizó contribuciones notables en computación, física cuántica, economía, cibernética y teoría de conjuntos, entre otros muchos campos. Una de sus aportaciones más reseñables es la denominada *arquitectura Von Neumann*, que describe la composición de un computador electrónico y que fue publicada en 1945.

<sup>2</sup>Boris Grigorievich Galerkin (4 de marzo de 1871, Polotsk, Rusia, actualmente Bielorrusia - 12 de junio de 1945, Moscú, URSS, actualmente Rusia) estudió matemáticas e ingeniería en el Instituto Tecnológico de San Petersburgo. Sus ideas al respecto de las soluciones aproximadas de una ecuación diferencial fueron publicadas en 1915. Desde 1940 hasta su muerte estuvo a la cabeza del Instituto de Mecánica de la Academia Soviética de las Ciencias.

<sup>3</sup>Richard Courant (8 de enero de 1888, Lublinitz, Prusia, actualmente Polonia - 27 de enero de 1972, New York, Estados Unidos) realizó grandes aportaciones en el ámbito de las ecuaciones diferenciales. Mientras trabajaba en un laboratorio con la formación de películas jabonosas, Courant se planteó la cuestión de probar matemáticamente la existencia de una solución a tal problema físico.

método. Sin embargo, es con la aparición posterior de las primeras computadoras cuando las bondades de esta metodología se vuelven tangibles.

El *método de los elementos finitos* es una herramienta numérica muy potente que permite aproximar soluciones de ecuaciones en derivadas parciales. Hablando con propiedad, sería más correcto decir que el método de los elementos finitos nos proporciona un algoritmo discreto (finito) para aproximar tales ecuaciones. En esta misma línea se manifiesta Susanne Brenner ([7]) cuando afirma que "*it [el método] should be thought of as a black box into which one puts the differential equation and out of which pops an algorithm for approximating the corresponding solutions*"<sup>4</sup>.

El desarrollo de un algoritmo de este tipo consta generalmente de cuatro fases. A saber,

- Primero, el problema de contorno<sup>5</sup> debe ser reformulado en forma débil.
- Después, se debe diseccionar el dominio espacial en subdominios que compongan una partición.
- Se debe elegir un *espacio de elementos finitos* que defina la aproximación, esto es, el conjunto de funciones mediante el que se va a aproximar: funciones lineales a trozos, cuadráticas a trozos... A los anteriores subdominios, dotados los ciertos grados de libertad escogidos, se los conoce como *elementos*.
- A continuación se toma la proyección del problema en forma débil sobre el espacio de los elementos finitos. De este modo se obtiene un sistema de ecuaciones de dimensión igual a la del espacio vectorial de elementos finitos que se haya empleado.
- Finalmente, se resuelve el sistema de ecuaciones dando lugar a la solución aproximada.

Es habitual cuando se exponen todas estas ideas comenzar con un problema diferencial concreto. Nosotros vamos a hacer lo propio en el caso uno dimensional. Consideremos el problema de contorno siguiente

$$[P] \quad \begin{cases} -\frac{d}{dx} \left( p(x) \frac{du}{dx} \right) + r(x)u = f(x), & a < x < b, \\ u(a) = \alpha, \quad u(b) = \beta, \end{cases}$$

donde  $\alpha, \beta \in \mathbb{R}$ , las funciones  $p, p', r \in C([a, b])$  y  $f \in L^2(a, b)$ . Además, se requiere que  $p(x) \geq p_0 > 0, r(x) \geq 0, \forall x \in [a, b]$ . En lo que sigue, supondremos que  $\alpha = \beta = 0$ , esto es,

<sup>4</sup>"El método puede ser entendido como una especie de caja negra en la cual uno pone la ecuación diferencial y al lado aparece un algoritmo para aproximar las soluciones correspondientes"

<sup>5</sup>En el contexto de las ecuaciones diferenciales, se denomina *problema de contorno* o *de frontera* a la reunión de una ecuación diferencial junto con unas condiciones en la frontera del dominio de definición. Cuando estas condiciones se establecen sobre los valores de la función implicada, se denominan *condiciones Dirichlet* o *esenciales*. Si atañen al valor de su derivada, entonces se conocen como *condiciones Neumann* o *naturales*. Si, por su parte, según qué puntos afectan, o bien a los valores de la función, o bien a los de su derivada, se dice que son *condiciones Robin* o *mixtas*.



condiciones Dirichlet homogéneas.

Si una función  $u$  de clase  $C^2([a, b])$  satisface  $[P]$ , entonces se dice que  $u$  es una *solución fuerte* o *clásica* del problema de contorno.

La *formulación débil* o *variacional* del problema  $[P]$  consiste en considerar, para cada función  $v$  de clase  $C^1([a, b])$  que se anula en la frontera del dominio, esto es, en  $a$  y en  $b$ , la identidad que se obtiene al multiplicar la ecuación diferencial por  $v$  a ambos lados e integrar en  $(a, b)$ ,

$$-\int_a^b \frac{d}{dx} \left( p(x) \frac{du}{dx} \right) v dx + \int_a^b r(x) u v dx = \int_a^b f(x) v dx.$$

Si ahora se integra por partes el primer término del lado izquierdo, se obtiene que

$$-\int_a^b \frac{d}{dx} \left( p(x) \frac{du}{dx} \right) v dx = p(x) \frac{du}{dx} v \Big|_a^b + \int_a^b p(x) \frac{du}{dx} \frac{dv}{dx} dx.$$

Pero recordemos que hemos exigido que  $v$  se anule en  $a$  y en  $b$ . Luego, por tanto, se tiene que

$$\int_a^b p(x) \frac{du}{dx} \frac{dv}{dx} dx + \int_a^b r(x) u v dx = \int_a^b f(x) v dx \text{ }^6.$$

Si  $u$  es una función absolutamente continua con derivada de cuadrado integrable y con  $u(a) = \alpha$  y  $u(b) = \beta$  que verifica la relación anterior para todas las funciones  $v$  que verifiquen las condiciones que impusimos antes, se dice que  $u$  es una *solución débil* del problema de contorno.

Notemos que en la formulación débil únicamente se exige que tanto  $u$  como su derivada sean de cuadrado integrables en el intervalo considerado. De hecho, la derivada de  $u$  no tiene por qué ser derivable ni continua.

<sup>6</sup>Todo lo anterior es generalizable a dimensión mayor que uno. Si  $\Omega$  es un abierto con frontera regular  $\partial\Omega$ , el problema correspondiente es

$$\begin{cases} -\operatorname{div}(p(\mathbf{x})\nabla u) + r(\mathbf{u}) = f(\mathbf{x}), & \mathbf{x} \in \Omega \\ u(\mathbf{x}) = g(\mathbf{x}), & \mathbf{x} \in \partial\Omega, \end{cases}$$

y la formulación variacional pasa a ser

$$\int_{\Omega} p(\mathbf{x})\nabla u^T \nabla v \, d\mathbf{x} + \int_{\Omega} r(\mathbf{x}) u v \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x}) v \, d\mathbf{x}.$$

Por su parte, es evidente que para formalizar esta teoría es necesario poner nombre a las cosas así como dotar de cierto rigor matemático a los resultados concernientes. A ello vamos a dedicar las secciones siguientes.

## 2.2. ALGUNOS CONCEPTOS DE ANÁLISIS FUNCIONAL.

Llegados a este punto, es necesario incluir en esta memoria una breve sección que compendie aquellas definiciones propias del análisis funcional que son de interés a la hora de exponer los resultados que aparecerán más adelante. Advertimos que, por brevedad, no vamos a incluir en esta sección demostraciones de los resultados.

**Definición 5.** (Norma y espacio normado.)

Dado un espacio vectorial  $V$  sobre un cierto cuerpo. Se dice que una aplicación,  $\|\cdot\|$ , definida sobre  $V$  con valores en  $\mathbb{R}$  es una norma si verifica que

- $\|v\| \geq 0 \quad \forall v \in V$ ;
- $\|v\| = 0$  si, y solo si,  $v = 0$ ;
- $\|\lambda v\| = |\lambda| \|v\| \quad \forall \lambda \in \mathbb{R}, \forall v \in V$ ;
- $\|v + w\| \leq \|v\| + \|w\| \quad \forall v, w \in V$ .

A un espacio vectorial  $V$  al que se le ha dotado de una aplicación norma se le llama espacio normado.

**Definición 6.** (Sucesión de Cauchy.)

Dado un espacio normado  $V$ . Se dice que una sucesión,  $\{v_n\}_{n \geq 0}$ , de elementos de  $V$  es de Cauchy o fundamental si se verifica que

$$\text{para cada } \epsilon > 0, \exists n_0 \in \mathbb{N} : \|v_n - v_m\| < \epsilon, \quad \text{si } n, m \geq n_0.$$

El concepto anterior da lugar a la siguiente propiedad de algunos espacios que son los que realmente nos van a interesar. También estos espacios gozan de nombre propio.

**Definición 7.** (Espacios completo y espacios de Banach.)

Se dice que un espacio normado es completo si en él toda sucesión de Cauchy es convergente. A los espacios normados completos se los denomina espacios de Banach.

De forma intuitiva podemos afirmar que cuando se trabaja en el ámbito de las ecuaciones diferenciales lo que se está haciendo es tratar con aplicaciones entre ciertos espacios. Una propiedad esencial de las aplicaciones es la *linealidad*.

**Definición 8.** (*Aplicación lineal.*)

Sean  $V$  y  $W$  dos espacios vectoriales sobre el mismo cuerpo. Se dice que una aplicación  $f: V \rightarrow W$  es lineal si se verifica que

$$f(\lambda v + \mu w) = \lambda f(v) + \mu f(w)$$

para cualesquiera que sean  $\lambda, \mu$  reales y  $v, w \in V$ .

Al trabajar con funciones lineales en espacios normados, los conceptos de continuidad y acotación se vuelven equivalentes.

**Proposición 1.** (*Continuidad de operadores lineales.*)

Dados dos espacios normados  $V, W$  sobre el mismo cuerpo y una aplicación lineal  $f: V \rightarrow W$ , entonces son equivalentes:

- $f$  es continua en  $V$ ,
- $f$  transforma conjuntos acotados en conjuntos acotados,
- $\exists M \geq 0: \|f(v)\| \leq M\|v\| \quad \forall v \in V$ ,
- $f$  es uniformemente continua en  $V$ .

Es posible encontrar una prueba detallada de estos resultados en [9].

En estas condiciones, se puede hablar del conjunto de aplicaciones lineales y continuas entre dos espacios normados dados,  $V$  y  $W$ . A este conjunto se lo denota habitualmente por  $\mathcal{L}(V, W)$ . A este conjunto, junto con las operaciones usuales de las funciones, se le puede dotar de una norma a fin de hacer de él un espacio normado. Tal norma se define de la siguiente manera:

$$\|f\| = \inf\{M \geq 0: \|f(v)\| \leq M\|v\| \quad \forall v \in V\} = \sup\{\|f(v)\|: \|v\| = 1\}.$$

Por supuesto que en un mismo espacio vectorial pueden definirse distintas normas. De esta manera, cabe preguntarse al respecto de la relación, si es que la hay, entre ellas, dando lugar a la siguiente definición.

**Definición 9.** (*Normas equivalentes.*)

Sea un espacio vectorial  $V$  sobre un cierto cuerpo. Dadas dos normas,  $\|\cdot\|_1, \|\cdot\|_2$ , definidas en él, entonces se dice que son equivalentes si

$$\exists N, M \geq 0: N\|v\|_1 \leq \|v\|_2 \leq M\|v\|_1 \quad \forall v \in V.$$

El que viene a continuación es un caso particular de la denominada *desigualdad de Hölder*, resultado muy conocido y útil del ámbito del análisis funcional.

**Proposición 2.** (*Desigualdad de Schwarz.*)

Sean  $f, g \in L^2(I)$ , con  $I \subseteq \mathbb{R}$ . Entonces,  $fg \in L^1(I)$  y

$$\int_I |f(x)g(x)| dx \leq \|f\|_{L^2(I)} \|g\|_{L^2(I)}.$$

Vamos a finalizar esta breve sección con las definiciones de *funcional lineal* y *espacio dual* junto con un resultado que enunciamos sin demostración relativo al análisis funcional.

**Definición 10.** (*Funcional lineal y espacio dual.*)

Sea un espacio normado  $V$  sobre un cierto cuerpo  $\mathbb{K}$ . Un funcional lineal es una aplicación lineal de  $V$  en  $\mathbb{K}$ . Al espacio de los funcionales lineales y continuos de  $V$  en  $\mathbb{K}$ , esto es, a  $\mathcal{L}(V, \mathbb{K})$ , se lo conoce como espacio dual de  $V$  y se denota por  $V'$ .

**Proposición 3.**

Dado el dual de un espacio normado  $V$  sobre un cierto cuerpo, esto es,  $V'$ . Entonces, si se le provee de la norma de los operadores que definíamos antes, se tiene que  $V'$  es un espacio de Banach.

Se puede encontrar información relativa a esta proposición en [9].

### 2.3. UN PROBLEMA BIEN PUESTO.

En este apartado pretendemos asentar los fundamentos teóricos suficientes para establecer bajo qué condiciones es posible garantizar la existencia y la unicidad del problema de contorno en forma débil.

Comencemos con una definición que constituye el eje fundamental de este contexto.

**Definición 11.** (*Forma bilineal.*)

Una aplicación  $\varphi(\cdot, \cdot)$  definida en  $V \times V$ , siendo  $V$  un espacio vectorial, a los números reales

tal que sus dos argumentos son lineales se dice que es una forma bilineal. Si, además, verifica que

$$\varphi(u, v) = \varphi(v, u) \quad \forall u, v \in V$$

entonces se dice que es simétrica.

Un producto interno,  $(\cdot, \cdot)$  es una forma bilineal simétrica en  $V$  que verifica que

- $(u, u) \geq 0 \quad \forall u \in V$ ;
- $(u, u) = 0$  si, y solo si,  $u = 0$ .

En los espacios dotados de un producto interno se verifica la celebrada *desigualdad de Schwarz*, esto es,

$$|(u, v)| \leq \sqrt{(u, u)}\sqrt{(v, v)}.$$

Del mismo modo, es conocido que  $\|u\| := \sqrt{(u, u)}$ , para cada  $u \in V$ , define una norma,  $\|\cdot\|$ , que se conoce como *norma asociada* al producto interno o *norma inducida* por él, en el espacio  $V$  correspondiente.

Es el momento de seguir con otra definición sobradamente conocida, la de *espacio de Hilbert*<sup>7</sup>.

**Definición 12.** (*Espacio de Hilbert.*)

Sea  $V$  un espacio dotado de un producto interno. Si el correspondiente espacio normado  $(V, \|\cdot\|)$  es completo, entonces se dice que es un espacio de Hilbert.

En lo que sigue, supondremos siempre que los espacios vectoriales que aparezcan son reales.

Además, es cierto que un subespacio cerrado de un espacio de Hilbert es también un espacio de Hilbert. Dado un subespacio  $M$  de un espacio de Hilbert,  $H$ , es posible definir su *ortogonal* como

$$M^\perp = \{u \in H : (u, v) = 0 \quad \forall v \in M\}.$$

<sup>7</sup>David Hilbert (23 de enero de 1862, Königsberg, Prusia - 14 de febrero de 1943, Gotinga, Alemania) es uno de los matemáticos más influyentes de todos los tiempos. Publicó, en 1899, *Fundamentos de geometría*, una obra que, a partir de nuevos axiomas, dota a la geometría de una base firme en una mejora manifiesta de los *Elementos* de Euclides. Es célebre su famosa conferencia de 1900, enmarcada en el Congreso Internacional de Matemáticas, en la cual enunció veintitrés grandes problemas sin resolver. De ellos, se han resuelto en la actualidad algunos pero otros, en cambio, no. De estos últimos, el más conocido es, sin duda, la *Hipótesis de Riemann*.

Más aún, dado cualquier elemento  $x \in H$  es posible definir también para el subespacio ortogonal,

$$x^\perp = \{u \in H : (u, x) = 0\}.$$

Este caso merece una especial atención pues es posible entender entonces el producto interno con  $x$  como la acción del funcional lineal definido para cada  $u \in H$  por

$$L_x(u) := (u, x).$$

De esta manera, es claro que el subespacio  $x^\perp$  coincide con  $\ker(L_x)$ . Además, según la desigualdad de Schwarz,

$$|L_x(u)| \leq \|x\| \|u\|,$$

y, por tanto,  $L_x$  es un operador acotado, esto es, continuo.

También, debemos reparar en que, puesto que

$$M^\perp = \bigcap_{x \in M} x^\perp$$

y que cada  $x^\perp$  es un subespacio cerrado de  $H$ , se tiene que  $M^\perp$  también es un subespacio de  $H$ .

Enunciemos ahora en forma de proposición algunos resultados relativos a espacios de Hilbert.

#### **Proposición 4.**

*Sea  $H$  un espacio de Hilbert. Entonces se verifican las siguientes propiedades:*

- *Si  $M, N$  son subespacios de  $H$  tales que  $M \subset N$ , entonces  $N^\perp \subset M^\perp$ .*
- *Para cada subespacio  $M$  se verifica que  $M \cap M^\perp = \{0\}$ .*
- *$\{0\}^\perp = H$ .*
- *$H^\perp = \{0\}$ .*

Es conveniente observar que en los espacios de Hilbert con la norma asociada al producto interno se cumple la denominada *regla del paralelogramo*, esto es,

$$\|u + v\|^2 + \|u - v\|^2 = 2(\|u\|^2 + \|v\|^2).$$

Uno de los puntos más importantes de los espacio de Hilbert trata con el concepto de descomposición del espacio en suma directa de subespacios ortogonales. De ahí se van a extraer otras afirmaciones interesantes.

**Proposición 5.** (Teorema de la proyección.)

Sean  $M$  un espacio cerrado de un espacio de Hilbert  $H$  y  $u \in H \setminus M$ . Se define

$$\delta = \inf \{ \|u - v\| : v \in M \}.$$

Entonces existe  $v_0 \in M$  tal que  $\|u - v_0\| = \delta$ , esto es, existe un punto,  $v_0 \in M$ , que es el más cercano a  $u$  y, además,  $u - v_0 \in M^\perp$ .

*Demostración.* Sea  $\{v_n\}$  una sucesión tal que

$$\lim_{n \rightarrow \infty} \|u - v_n\| = \delta. \quad (2.1)$$

Entonces, en virtud de la regla del paralelogramo, se tiene que

$$0 \leq \|v_n - v_m\|^2 = 2(\|v_n - u\|^2 + \|v_m - u\|^2) - 4 \left\| \frac{(v_n + v_m)}{2} - u \right\|^2.$$

Como  $\frac{(v_n + v_m)}{2} \in M$ , entonces, de acuerdo con la definición que se hizo de  $\delta$ ,

$$\left\| \frac{(v_n + v_m)}{2} - u \right\|^2 \geq \delta^2.$$

De este modo,

$$0 \leq \|v_n - v_m\|^2 \leq 2(\|v_n - u\|^2 + \|v_m - u\|^2) - 4\delta^2.$$

Utilizando (2.1) se tiene que  $\{v_n\}$  es una sucesión de Cauchy, luego convergente. En otros términos, existe  $v_0 \in \overline{M} = M$  tal que

$$\lim_{n \rightarrow \infty} \|v_n - v_0\| = 0.$$

Además, es claro que  $\|u - v_0\| = \delta$ .

Para ver que  $u - v_0 \in M^\perp$ , llamemos  $z = u - v_0$ . Sea  $v \in M$  y  $t$  un número real. Entonces,  $v_0 + tv \in M$  lo que implica que la función de la variable  $t$ ,

$$\|z - tv\|^2 = \|u - (v_0 + tv)\|^2$$

tiene un mínimo absoluto en  $t = 0$ . Por tanto,

$$0 = \left. \frac{d}{dt} \|z - tv\|^2 \right|_{t=0} = -2(z, v).$$

Esto supone que, para todo  $v \in M$ ,

$$(u - v_0, v) = (z, v) = 0.$$

Como  $v \in M$  es arbitrario, se sigue que  $u - v_0 \in M^\perp$ . ■

Por otra parte, dados un subespacio cerrado  $M \subset H$ ,  $H$  espacio de Hilbert, y  $u \in H$ , es posible descomponer  $u = v + w$ , con  $v \in M$  y  $w \in M^\perp$ . Esto es una mera consecuencia inmediata de la proposición anterior pues basta elegir  $w = u - v$ . Mostremos, además, que esta descomposición es única. Supongamos que

$$u = v_0 + v_1 = w_0 + w_1, \quad v_0, w_0 \in M, \quad v_1, w_1 \in M^\perp,$$

y entonces se tendría que  $v_0 - w_0 \in M$ ,  $-(v_1 - w_1) \in M^\perp$  y  $v_0 - w_0 = -(v_1 - w_1)$ . Como  $M \cap M^\perp = \{0\}$ , debe ser  $v_0 = w_0$  y  $v_1 = w_1$  y la descomposición es, efectivamente, única.

Se define el operador de proyección como

$$P_M: H \longrightarrow M, \quad P_M(u) = \begin{cases} u & \text{si } u \in M, \\ v_0 & \text{si } u \in H \setminus M, \end{cases}$$

donde  $v_0$  es el que habíamos considerado en la Proposición 4. Además, como lo que nos va a interesar involucra a espacios de Hilbert, en los cuales existe un producto interno,  $(\cdot, \cdot)$ , la anterior es equivalente a

$$(u, v) = (P_M(u), v), \quad \forall v \in M.$$

Finalmente, es claro a partir de lo anterior que

$$H = M \oplus M^\perp.$$

Es posible comprobar fácilmente que los operadores así definidos son lineales. Además, este tipo de aplicaciones pertenecen a una familia con nombre propio de funciones lo cual nos guía a la definición siguiente.



**Definición 13.** (*Operador de proyección.*)

Un operador lineal  $P : X \rightarrow X$ , con  $X$  un espacio vectorial, es un operador de proyección o, simplemente, una proyección si cumple que

$$P^2 \equiv P, \quad \text{esto es,} \quad P(u) = u \quad \forall u \in P(X).$$

Llegamos al momento de presentar el que es, sin lugar a duda, uno de los resultados más importantes y de mayor trascendencia del análisis funcional: el *teorema de representación de Riesz*.

Antes, hagamos un comentario sobre la notación escogida. Denotaremos en lo que sigue la acción de un determinado funcional  $L \in H'$ , con  $H$  un espacio de Hilbert, sobre un elemento  $u$  por  $\langle L, u \rangle$  en vez de  $L(u)$ .

Volviendo al asunto que nos ocupa, sabemos que siempre es cierto que, dado un elemento de un espacio con producto interno,  $X$ , podemos definir un funcional  $L$  por

$$\langle L, v \rangle = (u, v) \quad \forall v \in X,$$

donde se ha utilizado la notación

$$\langle L, v \rangle = L(v).$$

La linealidad del funcional así definido es inmediata. Por otra parte, observemos que

$$|\langle L, v \rangle| = |(u, v)| \leq \|u\| \|v\|,$$

lo que implica que

$$\|L\| \leq \|u\|.$$

Por otro lado,

$$|\langle L, u \rangle| = |(u, u)| = \|u\|^2,$$

y, entonces,

$$\|L\| \geq \frac{|\langle L, u \rangle|}{\|u\|} = \|u\|.$$

Luego,  $\|L\| = \|u\|$ .

Pues bien, el teorema de representación de Riesz nos dice que, para un funcional lineal continuo en un espacio de Hilbert,  $H$ , el recíproco también es cierto, esto es, dado  $L$ , siempre es posible hallar un elemento  $u \in H$  tal que  $\langle L, v \rangle = (u, v)$ , para todo  $v \in H$ . Enunciemos como es debido este resultado.

**Teorema 1.** (Teorema de representación de Riesz)

Sea  $H$  un espacio de Hilbert y  $L \in H'$ . Entonces existe un único elemento  $u \in H$  tal que

$$\langle L, v \rangle = (u, v) \quad \forall v \in H.$$

Además, se tiene que

$$\|L\|_{H'} = \|u\|_H.$$

*Demostración.* Sea  $M = \{v \in H : \langle L, v \rangle = 0\}$ . Según hemos comentado antes,  $M$  es un subespacio cerrado de  $H$ . Además,  $H = M \oplus M^\perp$ .

Si  $M^\perp = \{0\}$ , entonces  $M = H$  y se tiene que  $L \equiv 0$ . En este caso, podemos elegir  $u = 0$ .

En caso contrario, esto es, si  $M^\perp \neq \{0\}$ , tomemos un elemento  $x \in M^\perp, x \neq 0$ . Se sigue que  $\langle L, x \rangle \neq 0$  pues, en otro caso,  $x$  pertenecería también a  $M$  lo que, dado que  $M \cap M^\perp = \{0\}$ , supondría que  $x = 0$ . Entonces, para  $v \in H$  y  $\lambda = \frac{\langle L, v \rangle}{\langle L, x \rangle}$ , se tiene que

$$\langle L, v - \lambda x \rangle = \langle L, v \rangle - \lambda \langle L, x \rangle = 0. \quad (2.2)$$

En otros términos, se tiene que  $v - \lambda x \in M$  y como  $x \in M^\perp$ , se tiene que  $P_M(v) = v - \lambda x$ .

Definamos ahora

$$u = \frac{\langle L, x \rangle}{\|x\|_H^2} x. \quad (2.3)$$

Nótese que  $u \in M^\perp$ . Por otra parte,

$$\begin{aligned} (u, v) &= \left( u, (v - \lambda x) + \lambda x \right) \\ &= (u, v - \lambda x) + (u, \lambda x) \\ &= (u, \lambda x) \\ &= \lambda \frac{\langle L, x \rangle}{\|x\|_H^2} (x, x) \\ &= \lambda \langle L, x \rangle \\ &= \langle L, \lambda x \rangle \\ &= \langle L, v \rangle, \end{aligned}$$

donde se ha utilizado (2.2).

Luego el elemento  $u \in H$  así definido es el que estábamos buscando. Probaremos ahora que  $\|L\|_{H'} = \|u\|_H$ .

Así, se tiene que por una parte que

$$\begin{aligned}\|L\|_{H'} &= \sup_{0 \neq v \in H} \frac{|\langle L, v \rangle|}{\|v\|_H} \\ &= \sup_{0 \neq v \in H} \frac{|(u, v)|}{\|v\|_H} \\ &\leq \|u\|_H\end{aligned}$$

Por la otra, de (2.3),

$$\begin{aligned}\|u\|_H &= \frac{|\langle L, x \rangle|}{\|x\|_H} \\ &\leq \frac{\|L\|_{H'} \|x\|_H}{\|x\|_H} \\ &= \|L\|_{H'}\end{aligned}$$

Por tanto, se deduce que  $\|L\|_{H'} = \|u\|_H$ .

Queda aún por probar la unicidad. Sean  $u_1, u_2$  dos elementos verificando que  $\langle L, v \rangle = (u_1, v) = (u_2, v)$ . Por lo tanto,

$$(u_1 - u_2, v) = 0 \quad \forall v \in H.$$

En particular, se tiene que

$$(u_1 - u_2, u_1 - u_2) = 0 \implies u_1 - u_2 = 0 \implies u_1 = u_2.$$

■

Antes de seguir, necesitamos introducir dos conceptos relativos a formas bilineales.

**Definición 14.** (Continuidad y coercitividad.)

Una forma bilineal  $a(\cdot, \cdot)$ , definida en un cierto espacio de Hilbert,  $H$ , es continua (acotada) si existe un número real  $\beta > 0$  tal que

$$|a(u, v)| \leq \beta \|u\| \|v\| \quad \forall u, v \in H.$$

Asimismo, si existe un número real  $\alpha > 0$  de modo que

$$a(u, u) \geq \alpha \|u\|^2 \quad \forall u \in H,$$

entonces se dice que es coercitiva.

Nótese que, como consecuencia inmediata de la definición de coercitividad, si se tiene que una forma bilineal coercitiva  $a(\cdot, \cdot)$  en un espacio de Hilbert,  $H$ , verifica que  $a(u, u) \geq 0$ , para todo  $u \in H$ ; la igualdad se alcanza si, y solo si,  $u = 0$ .

El teorema de representación de Riesz admite una suerte de generalización a formas bilineales no necesariamente simétricas. Enunciémoslo.

**Teorema 2.** (Teorema de Lax-Milgram.)

Sea  $H$  un espacio de Hilbert y  $a(\cdot, \cdot)$  una forma bilineal continua y coercitiva en  $H$ . Entonces, dado un funcional continuo  $L$  en  $H$ , existe un único elemento  $u_L \in H$  tal que

$$a(u_L, v) = \langle L, v \rangle \quad \forall v \in H.$$

Además,

$$\|u_L\|_H \leq \frac{1}{\alpha} \|L\|_{H'},$$

donde  $\alpha$  es la constante de coercitividad de  $a(\cdot, \cdot)$ .

La demostración de este teorema resulta ser bastante densa y, por ello, tal como hace Reddy en [9], hemos decidido diseccionarla en varios lemas previos. Advertimos de que en los lemas venideros se suponen, además de las explicitadas en cada caso, las condiciones del teorema de Lax-Milgram. Vamos con ello.

**Lema 3.**

Dado un elemento  $u \in H$ , existe un único elemento  $w \in H$  tal que

$$a(u, v) = (w, v) \quad \forall v \in H. \quad (2.4)$$

*Demostración.* Fijado  $u \in H$ , la aplicación  $a_u(v) := a(u, v)$ , para todo  $v \in H$ , es un funcional acotado pues

$$|a_u(v)| \leq \beta \|u\| \|v\| \quad \forall v \in H.$$

Por tanto, de acuerdo con el teorema de representación de Riesz, existe un único elemento  $w \in H$  tal que  $a_u(v) = (w, v)$ , esto es, tal que  $a(u, v) = (w, v)$  y, además,

$$\|w\| = \|a_u\| \leq \beta \|u\|.$$

■

**Lema 4.**

Sea  $T$  el operador definido por

$$T(u) = w,$$

donde  $w$  es el único elemento en  $H$  tal que  $a(u, v) = (w, v)$  para todo  $v \in H$ . Entonces  $T$  es un operador lineal y continuo.

*Demostración.* Sean  $u_1, u_2$  dos elementos de  $H$  y  $w_1 = T(u_1), w_2 = T(u_2)$ . Veamos que  $T(\lambda u_1 + \mu u_2) = \lambda T(u_1) + \mu T(u_2)$ , para todo  $\lambda, \mu \in \mathbb{R}$ .

Como  $a(\cdot, \cdot)$  es una forma bilineal, entonces se verifica que

$$a(\lambda u_1 + \mu u_2, v) = \lambda a(u_1, v) + \mu a(u_2, v) = (\lambda w_1 + \mu w_2, v).$$

Además,  $\lambda T(u_1) + \mu T(u_2) = \lambda w_1 + \mu w_2$ . Así, por el primer lema, se sigue que

$$T(\lambda u_1 + \mu u_2) = \lambda w_1 + \mu w_2,$$

como consecuencia de (2.4), y por tanto, se prueba la linealidad de  $T$ .

Por su parte, si  $u \neq 0$ , sea  $w = T(u)$ . Ahora, puesto que  $a(\cdot, \cdot)$  es continua, se tiene que

$$\beta \|u\| \|T(u)\| \geq a(u, T(u)) = (w, T(u)) = \|T(u)\|^2,$$

lo cual implica que

$$\|T(u)\| \leq \beta \|u\|.$$

Si  $u = 0$ , la desigualdad es trivial ( $T(0) = 0$ , pues es lineal). Por tanto,  $T$  es un operador acotado y

$$\|T\| \leq \beta.$$

■

#### **Lema 5.**

El operador  $T$ , definido en el Lema 4, es inyectivo y su inverso,  $T^{-1}$  es acotado.

*Demostración.* Veamos que  $\ker(T) = \{0\}$ , esto es, que  $T(u) = 0$  solo si  $u = 0$  para concluir que el operador lineal  $T$  es inyectivo. Sea  $u$  un elemento tal que  $T(u) = 0$ . Entonces

$$a(u, v) = (T(u), v) = (0, v) = 0 \quad \forall v \in H.$$

En particular, para  $v = u$ , se obtiene que

$$0 = a(u, u) \geq \alpha \|u\|^2,$$

donde  $\alpha$  es la constante de coercitividad de la forma bilineal  $a(\cdot, \cdot)$ . Entonces,  $\|u\| = 0$  y, por tanto,  $u = 0$ . Luego, en efecto, el operador  $T$  es inyectivo y, restringido a la imagen de  $T$ , su inverso existe. Además, es un resultado conocido del álgebra lineal que el inverso de un operador lineal es también lineal. Para ver que es continuo reparemos en que, si  $T(u) = w$ ,

$$\alpha \|u\|^2 \leq a(u, u) = (w, u) \leq \|w\| \|u\|,$$

de donde

$$\|T^{-1}(w)\| = \|u\| \leq \frac{1}{\alpha} \|w\|,$$

lo cual implica, que

$$\|T^{-1}\| \leq \frac{1}{\alpha}.$$

■

### Lema 6.

Sea  $T$  el operador definido en el Lema 4. Entonces, la imagen de  $T$ ,  $\text{Im}(T)$ , es un espacio completo.

*Demostración.* Sea  $(w_n)_{n \in \mathbb{N}}$  una sucesión de Cauchy en  $\text{Im}(T)$ . Como  $\text{Im}(T) \subset H$ ,  $(w_n)_{n \in \mathbb{N}}$  también es una sucesión de Cauchy en  $H$ . Como  $H$  es un espacio completo, la sucesión converge en él, esto es, existe un elemento  $w \in H$  tal que

$$\lim_{n \rightarrow \infty} \|w_n - w\| = 0.$$

Necesitamos ver que  $w \in \text{Im}(T)$ . Para ello, definamos  $u_n = T^{-1}(w_n)$ , para cada  $n \in \mathbb{N}$ . Entonces,

$$\|u_n - u_m\| = \|T^{-1}(w_n) - T^{-1}(w_m)\| = \|T^{-1}(w_n - w_m)\| \leq \frac{1}{\alpha} \|w_n - w_m\|,$$

luego

$$\lim_{n, m \rightarrow \infty} \|u_n - u_m\| \leq \frac{1}{\alpha} \lim_{n, m \rightarrow \infty} \|w_n - w_m\| = 0.$$

Por lo tanto,  $(u_n)_{n \in \mathbb{N}}$  es también una sucesión de Cauchy en  $H$ . Luego es convergente. Sea entonces  $u \in H$  tal que

$$\lim_{n \rightarrow \infty} u_n = u.$$

Como el operador  $T$  es continuo, se tiene que

$$T(u) = \lim_{n \rightarrow \infty} T(u_n) = \lim_{n \rightarrow \infty} w_n = w,$$

con  $w \in \text{Im}(T)$ . Luego, como queríamos probar, el espacio  $\text{Im}(T)$  es completo.

■

### Lema 7.

El operador  $T$ , definido en el Lema 4, es biyectivo.

*Demostración.* Tenemos que ver que la imagen del operador  $T$  es todo el espacio  $H$ , esto es, que  $\text{Im}(T) = H$ . Sea  $v \in \text{Im}(T)^\perp$ . Entonces,

$$(T(u), v) = 0 \quad \forall u \in H.$$

Pero  $(T(u), v) = a(u, v)$ , luego, en particular,

$$(T(v), v) = a(v, v) = 0.$$

Como la forma bilineal  $a(\cdot, \cdot)$  es coerciva, lo anterior implica que  $v = 0$  y, por tanto, el operador  $T$  es, en efecto, biyectivo. ■

*Demostración del teorema de Lax-Milgram.* En primer lugar, como consecuencia del teorema de representación de Riesz, existe un único  $w_L \in H$  de manera que

$$(w_L, v) = \langle L, v \rangle \quad \forall v \in H.$$

Ahora bien, según hemos probado en los lemas precedentes, el operador  $T$  que envía a cada  $u \in H$  en el único elemento  $w \in H$  tal que  $a(u, v) = (w, v) \quad \forall v \in H$  es biyectivo. Por tanto, existe una única contraimagen del elemento  $w_L$  o, dicho de otra manera, existe un único  $u_L \in H$  tal que  $T(u_L) = w_L$ . De esta manera, se llega a que

$$(T(u_L), v) = \langle L, v \rangle \quad \forall v \in H.$$

Así, por la propia definición de la aplicación  $T$ , se obtiene que, efectivamente,

$$a(u_L, v) = \langle L, v \rangle \quad \forall v \in H.$$

Por su parte, si  $u_L \neq 0$ , entonces

$$\alpha \|u_L\|_H^2 \leq a(u_L, u_L) = \langle L, u_L \rangle \leq \|L\|_{H'} \|u_L\|_H,$$

es decir,

$$\|u_L\|_H \leq \frac{1}{\alpha} \|L\|_{H'}.$$

Finalmente, señalemos que si  $u_L = 0$ , la desigualdad es trivial. ■

Vamos a dedicar ahora una definición a asentar de manera clara lo que se entiende por *problema variacional*.

**Definición 15.** (*Problema variacional.*)

Sean  $V$  es un espacio de Hilbert con producto interno  $(\cdot, \cdot)$ . Supongamos que  $a(\cdot, \cdot)$  es una forma bilineal continua y coercitiva en  $V$ . Dado  $F \in V'$ , se define el problema variacional como la búsqueda de  $u \in V$  tal que

$$a(u, v) = \langle F, v \rangle \quad \forall v \in V. \quad (2.5)$$

Si, además, la forma bilineal es simétrica, entonces se dice problema variacional simétrico.

La expresión que da nombre a esta sección es debida a Hadamard<sup>8</sup> y no se trata tan solo de una expresión coloquial. En verdad, él creía que para que un modelo matemático gozase de relevancia dentro del ámbito de la física, éste debía satisfacer ciertas propiedades. Concretamente,

- debería existir (existencia) una solución del modelo
- y ésta habría de ser única (unicidad).
- Además, el comportamiento de la solución variaría de forma continua conforme lo hiciesen las condiciones iniciales (dependencia continua de los datos), esto es, si los datos cambiasen un "poco", la solución del modelo perturbado sería "próxima" a la original.

En tales condiciones, se dice que el modelo pertinente es un *problema bien puesto* o *bien planteado* en el sentido de Hadamard. Enunciemos en esta línea el que va a ser el último resultado de esta sección.

**Teorema 3.**

*El problema variacional que hemos definido es un problema bien puesto.*

La demostración de este teorema es el propio teorema de Lax-Milgram.

<sup>8</sup>Jacques Salomon Hadamard (8 de diciembre de 1865, Versailles, Francia - 17 de octubre de 1963, París, Francia) fue un matemático francés que trabajó en las universidades de Burdeos y París. Sucedió a Henri Poincaré al frente de la Academia de las Ciencias de Francia en 1912. En su libro *Psicología de la invención en el campo matemático* trata de describir los estados de la mente al hacer matemáticas. Consultó a numerosos científicos de la época obteniendo como resultado el que muchos desarrollaban las matemáticas sin utilizar palabras percibiendo en cambio imágenes o sensaciones, algo que algunos sitúan cerca de la teoría de los estadios del aprendizaje de Piaget.



## 2.4. ESPACIOS DE SOBOLEV.

Los espacios  $C^n(\Omega)$ , pese a estar formados por funciones  $n$ -diferenciables con continuidad, tienen por contraparte que no son espacios de Hilbert para la norma usual. Por tanto, no van a ser estos los espacios naturales en los que trabajar con ecuaciones en derivadas parciales.

Así, los espacios que nos van a interesar son los denominados *espacios de Sobolev* pues, en primer lugar, existe un conjunto de ellos que si que son espacios de Hilbert y, por otra parte, es posible llegar a resultados de existencia y unicidad del problema variacional cuando se trabaja en tales espacios.

La presentación de los espacios de Sobolev suele implicar el desarrollo de la teoría de las distribuciones o funciones generalizadas. Éstas constituyen, valga la redundancia, una generalización del concepto de función que pasa por la idea de la distribución de masa. En cambio, dado que la presente memoria viene a colación del problema de valorar derivados financieros y, en particular, de aproximar la ecuación de Black-Scholes, será suficiente con que nos ciñamos al caso 1-dimensional. Es posible encontrar una presentación clara y suficientemente rigurosa del caso general en los libros [7] y [9]. Nos restringiremos, por tanto, a los espacios de Sobolev en intervalos de  $\mathbb{R}$ .

**Definición 16.** (*Espacio de Sobolev en un intervalo.*)

Fijado un entero  $k$ , se define el espacio de Sobolev  $H^k(a, b)$  como el conjunto de las funciones reales  $v$  definidas en el intervalo  $[a, b]$  tales que todas sus derivadas de orden menor o igual que  $k - 1$  son absolutamente continuas en  $[a, b]$  y, además,

$$\left( \int_a^b \left( \frac{d^k v}{dx^k}(x) \right)^2 dx \right)^{1/2} < \infty.$$

A este conjunto se le dota de una norma, que se conoce como *norma de Sobolev*, de la siguiente manera,

$$\|v\|_{H^k(a, b)} = \left( \sum_{j=0}^k \int_a^b \left( \frac{d^j v}{dx^j}(x) \right)^2 dx \right)^{1/2}$$

Como subconjuntos de  $H^1(a, b)$  aparecen  $H_E^1$  y  $H_0^1$  formados por aquellas funciones  $v$  que cumplen  $v(a) = \alpha$ ,  $v(b) = \beta$ , para ciertos valores reales  $\alpha, \beta$ , y  $v(a) = 0$ ,  $v(b) = 0$ , respectivamente. Observemos que si  $v \in H^1(a, b)$ , entonces  $v$  es continua y tiene perfecto sentido evaluar la función.

Un resultado típico de la teoría de espacios de Sobolev y que permite obtener cotas para ciertas funciones a partir de sus derivadas es la *desigualdad de Poincaré*<sup>9</sup>. A continuación,

<sup>9</sup>Jules Henri Poincaré (29 de abril de 1854, Nancy, Francia - 17 de julio de 1912, París, Francia) fue un matemático, físico y filósofo de la ciencia francés considerado, junto a David Hilbert, uno de los matemáticos

enunciamos un caso particular de la misma que es el que se ajusta a nuestros intereses.

**Proposición 6.** (*Desigualdad de Poincaré.*)

Sea  $I$  un intervalo real. Entonces, existe una constante  $C_P \in \mathbb{R}$  tal que

$$\|u\|_{L^2(I)} \leq C_P \|u'\|_{L^2(I)} \quad \forall u \in H_0^1(I).$$

Adviértase que, para una cierta función  $u \in H_0^1(I)$ , las normas  $\|u'\|_{L^2(a,b)}$  y  $\|u\|_{H_0^1(a,b)}$  son equivalentes. En efecto, por una parte, es claro que

$$\|u'\|_{L^2(a,b)} \leq \|u\|_{H_0^1(a,b)} = \left\{ \|u\|_{L^2(a,b)}^2 + \|u'\|_{L^2(a,b)}^2 \right\}^{1/2}.$$

Por otra, valiéndonos de la desigualdad de Poincaré, se tiene que

$$\|u\|_{H_0^1(a,b)} \leq \left\{ \|u'\|_{L^2(a,b)}^2 + C_P^2 \|u'\|_{L^2(a,b)}^2 \right\}^{1/2} = \sqrt{1 + C_P^2} \|u'\|_{L^2(a,b)}.$$

Volviendo ahora al caso particular del problema  $[P]$  que hemos planteado en la Sección 2.1, el problema variacional asociado se establece en términos de

$$a(u, v) := \int_a^b \left\{ p(x) u' v' dx + r(x) uv \right\} dx,$$

y de

$$\langle F, v \rangle := \int_a^b f(x) v dx.$$

De esta manera, el problema variacional asociado a  $[P]$  se enuncia como la búsqueda de  $u \in H_E^1(a, b)$  tal que

$$a(u, v) = \langle F, v \rangle \quad \forall v \in H_0^1(a, b).$$

Veamos que la forma  $a(\cdot, \cdot)$  así definida es continua en  $H_0^1(a, b)$ .

---

más célebres de su tiempo. Se le ha denominado como el "último matemático universal", es decir, el último con conocimiento profundo al respecto de todas las materias de su época. Poincaré apostaba por una visión romántica de las matemáticas basada en la *intuición*, una percepción diametralmente opuesta al formalismo estricto que defendía Hilbert, llegando a afirmar incluso que *"la lógica es yerma si no se fertiliza con la intuición"*.

$$\begin{aligned}
|a(u, v)| &= \left| \int_a^b \{p(x)u'v' + r(x)uv\} dx \right| \\
&\leq \left| \int_a^b p(x)u'v' dx \right| + \left| \int_a^b r(x)uv dx \right| \\
&\leq \max_{x \in (a,b)} \{p(x)\} \int_a^b |u'| |v'| dx + \max_{x \in (a,b)} \{r(x)\} \int_a^b |u| |v| dx \\
&\leq \max_{x \in (a,b)} \{p(x)\} \left( \int_a^b |u'|^2 dx \right)^{1/2} \left( \int_a^b |v'|^2 dx \right)^{1/2} \\
&\quad + \max_{x \in (a,b)} \{r(x)\} \left( \int_a^b |u|^2 dx \right)^{1/2} \left( \int_a^b |v|^2 dx \right)^{1/2} \\
&\leq \max_{x \in (a,b)} \{p(x)\} \left( \int_a^b |u'|^2 dx \right)^{1/2} \left( \int_a^b |v'|^2 dx \right)^{1/2} \\
&\quad + C_P \max_{x \in (a,b)} \{r(x)\} \left( \int_a^b |u'|^2 dx \right)^{1/2} \left( \int_a^b |v'|^2 dx \right)^{1/2} \\
&= \left( \max_{x \in (a,b)} \{p(x)\} + C_P \max_{x \in (a,b)} \{r(x)\} \right) \left( \int_a^b |u'|^2 dx \right)^{1/2} \left( \int_a^b |v'|^2 dx \right)^{1/2}.
\end{aligned}$$

Hagamos lo propio con la coercitividad en  $H_0^1(a, b)$ .

$$\begin{aligned}
|a(u, u)| &= \left| \int_a^b \{p(x)u'u' + r(x)uu\} dx \right| \\
&= \left| \int_a^b p(x)(u')^2 + r(x)u^2 dx \right| \\
&= \int_a^b p(x)(u')^2 dx + \int_a^b r(x)u^2 dx \\
&\geq p_0 \int_a^b (u')^2 dx + \min\{r(x)\} \int_a^b u^2 dx \\
&\geq p_0 \int_a^b (u')^2 dx.
\end{aligned}$$

Por lo tanto, en virtud del teorema de Lax-Milgram, el problema variacional asociado al problema de contorno  $[P]$  posee una única solución débil.

## 2.5. EL MÉTODO GALERKIN.

En las secciones previas nos hemos dedicado en exclusiva a estudiar conceptos teóricos del análisis funcional descuidando por completo el hecho de que nuestro fin último es el

de hallar soluciones efectivamente computables. Además, es indudable que, salvo para las ecuaciones en derivadas parciales más sencillas en dominios adecuadamente escogidos, es imposible encontrar soluciones analíticas para los problemas de contorno. Si bien el enfoque clásico de estos problemas conduce a métodos de aproximación tan válidos como el de las *diferencias finitas*, el tratamiento variacional ha provocado que, a partir de mediados del siglo XX, métodos como el de Galerkin se popularizasen notablemente. Concretamos en qué consiste tal método.

**Definición 17.** (*Problema de Aproximación de Galerkin.*)

Sea  $V$  es un espacio de Hilbert con producto interno  $(\cdot, \cdot)$ . Supongamos que  $a(\cdot, \cdot)$  es una forma bilineal continua y coercitiva en  $V$ . Dado  $F \in V'$ , consideremos el problema variacional (2.5). Sea  $V_h$  un subespacio de dimensión finita de  $V$ , se define el problema de aproximación de Galerkin como la búsqueda de  $u_h \in V_h$  tal que

$$a(u_h, v_h) = \langle F, v_h \rangle \quad \forall v_h \in V_h.$$

En tal situación, diremos que  $u_h$  es una aproximación de Galerkin en el espacio  $V_h$ .

Notemos que lo que realmente tenemos es un problema variacional restringido a un espacio de dimensión finita. Como lo general implica lo particular, se tiene que, en virtud de los resultados de la sección previa, el problema de aproximación de Galerkin es un problema bien puesto. Por tanto, podemos hablar de la (única) aproximación de Galerkin en el espacio correspondiente.

La notación del subíndice  $h$  es la que aparece, de manera más o menos similar, en la mayoría de los textos que hemos consultado. Resulta ser muy ilustrativa pues, si bien  $V$  es un espacio muy 'grande' (de dimensión infinita), el valor de  $h$  nos va a mostrar en cierto modo cómo de 'cerca' está el subespacio  $V_h$  del propio  $V$ . Concretamente, si  $V_h$  es un espacio de dimensión  $(N + 1)$ , esto es,

$$V_h = \text{span} \{ \psi_j \}_{j=0}^N,$$

para ciertas funciones linealmente independientes de  $V$ , definimos  $h = \frac{c}{N}$ ,  $c \in \mathbb{R} \setminus \{0\}$ . De esta manera,  $h \in (0, 1]$  y cuando  $N \rightarrow \infty$  se tiene que  $h \rightarrow 0$  significando que cuanto menor es el índice  $h$ , mejor es la aproximación del subespacio  $V_h$  a  $V$ .

Para tratar de obtener  $u_h$  debemos reparar en que tanto  $u_h$  como  $v_h$  deben ser sendas combinaciones lineales de las funciones que conforman la base de  $V_h$ . De esta manera,

$$u_h = \sum_{j=0}^N d_j \psi_j \quad v_h = \sum_{k=0}^N b_k \psi_k.$$

Al llevar estas expresiones al problema de aproximación de Galerkin, como  $a(\cdot, \cdot)$  es bilineal y  $F$  es lineal, se obtiene que

$$\sum_{j=0}^N \sum_{k=0}^N d_j b_k a(\psi_j, \psi_k) = \sum_{k=0}^N b_k \langle F, \psi_k \rangle.$$

Si denotamos por  $K_{jk} = a(\psi_j, \psi_k)$  y  $F_k = \langle F, \psi_k \rangle$ , podemos reescribir lo anterior como

$$\sum_{k=0}^N b_k \left( \sum_{j=0}^N K_{jk} d_j - F_k \right) = 0.$$

Dado que los coeficientes  $b_k$  son arbitrarios, debe ser

$$\sum_{j=0}^N K_{jk} d_j = F_k, \quad k = 0, \dots, N.$$

En la literatura al respecto de este tema, lo común es denotar por  $\mathbf{K}$  a la matriz  $\{K_{jk}\}_{\substack{1 \leq j \leq N \\ 1 \leq k \leq N}}$  y por  $\mathbf{F}$  y  $\mathbf{d}$  a los vectores  $F_k$  y  $d_j$ , obteniéndose así la compacta expresión

$$\mathbf{Kd} = \mathbf{F}.$$

Nótese que este sistema de ecuaciones tiene solución única como consecuencia de la coercitividad de la forma bilineal  $a(\cdot, \cdot)$ .

Observemos que la potencia de estas ecuaciones reside en que hemos llevado el problema de aproximación de Galerkin a la resolución de un sistema de ecuaciones lineales. A partir de los coeficientes  $d_j$ , se halla la solución aproximada  $u_h$ .

En la sección anterior, cuando definíamos lo que era una forma bilineal, dijimos también que si ésta era simétrica y coercitiva, entonces definía un producto interno del espacio. Asimismo, si la forma es además continua, entonces la norma asociada a este producto interno,  $\|\cdot\|_a = \sqrt{a(\cdot, \cdot)}$ , es equivalente a la norma original del espacio. Esto es debido a que, como consecuencia de la coercitividad de  $a(\cdot, \cdot)$ , se tiene que

$$\|u\|_a = \sqrt{a(u, u)} \geq \sqrt{\alpha} \|u\| = \sqrt{\alpha} \|u\|,$$

donde  $\alpha$  es la constante de coercitividad de la forma. Del mismo modo, si  $\beta$  es la constante de continuidad de la forma  $a(\cdot, \cdot)$ , se obtiene que

$$\|u\|_a = \sqrt{a(u, u)} \leq \sqrt{\beta} \|u\| = \sqrt{\beta} \|u\|.$$

Por tanto, efectivamente,

$$\sqrt{\alpha} \|u\| \leq \|u\|_a \leq \sqrt{\beta} \|u\|,$$

y las normas son equivalentes.

De esta manera, si  $V \subset H$  era completo respecto de la norma original, también lo será respecto de  $\|\cdot\|_a$ . Al producto interno asociado a  $a(\cdot, \cdot)$  lo denotaremos por  $(\cdot, \cdot)_a$ .

Si ahora tomamos las funciones de la base del espacio  $V_h$ ,  $\{\psi_j\}_{j=0}^N$  de modo que conformen una base ortonormal respecto al producto interno  $(\cdot, \cdot)_a$ , el sistema de ecuaciones se simplifica considerablemente pues

$$K_{jk} = a(\psi_j, \psi_k) = (\psi_j, \psi_k)_a = 0 \quad \text{si, y solo si, } j \neq k,$$

y, por tanto,

$$K_{jj}d_j = F_j,$$

esto es,

$$d_j = \frac{F_j}{K_{jj}}.$$

Cuando definíamos lo que se entendía por problema de aproximación de Galerkin, también comentábamos que éste era un problema bien puesto. Pues bien, vamos a analizar ahora lo que ocurre con los errores de tales aproximaciones.

El error se define como la diferencia entre la solución exacta y la aproximación, esto es,

$$e = u - u_h.$$

Por su parte, como  $V_h \subset V$ , podemos restar la expresión del problema variacional restringido a  $V_h$  a la del problema de Galerkin. De esta manera, obtenemos que

$$a(u, v_h) = a(u_h, v_h) \implies a(u - u_h, v_h) = 0,$$

esto es,

$$a(e, v_h) = 0. \tag{2.6}$$

A la expresión anterior se la conoce como *ortogonalidad Galerkin* y no depende del hecho de que  $a(\cdot, \cdot)$  sea simétrica.

Si la forma  $a(\cdot, \cdot)$  es simétrica, este resultado goza de una interpretación geométrica muy interesante. En tal caso y como ya hemos comentado, la forma define un producto interno,  $(\cdot, \cdot)_a$ . Por ello, ahora entra en juego todo el compendio de resultados relativos a proyecciones que presentábamos en la sección anterior. Así, si  $\{\psi_j\}_{j=0}^N$  es una base ortonormal respecto de  $(\cdot, \cdot)_a$  en  $V_h$ , la proyección ortogonal de un elemento  $v$  viene dada por

$$P(v) = \sum_{j=0}^N (v, \psi_j)_a \psi_j.$$

Si ahora tenemos en cuenta que, en particular,

$$(u, \psi_j)_a = (u_h, \psi_j)_a,$$

llegamos a que

$$P(u) = \sum_{j=0}^N (u_h, \psi_j)_a \psi_j = P(u_h) = u_h.$$

Luego la proyección ortogonal de la solución exacta,  $u$ , sobre el espacio  $V_h$ , respecto del producto interno  $(\cdot, \cdot)_a$  es la solución aproximada,  $u_h$ . De esta forma, como  $e = u - u_h$ , se tiene que  $e = u - P(u) \in V_h^\perp$  y, así, el error es ortogonal al subespacio  $V_h$ . En consecuencia,

$$(e, u_h)_a = 0. \quad (2.7)$$

Además, si atendemos al comportamiento de  $\|\cdot\|_a$ , vamos a ver qué ocurre con la distancia entre  $u$  y  $v_h \in V_h$ .

$$\begin{aligned} \|u - v\|_a^2 &= a(u - v, u - v) \\ &= a(u - u_h + u_h - v, u - u_h + u_h - v) \\ &= a(e + (u_h - v), e + (u_h - v)) \\ &= a(e, e) + 2a(e, u_h - v) + a(u_h - v, u_h - v) \\ &= \|e\|_a^2 + \|u_h - v\|_a^2 \end{aligned}$$

pues, según hemos afirmado, el error es ortogonal a todos los elementos del espacio  $V_h$ . Por tanto, podemos observar como el mínimo se alcanza cuando  $v = u_h$ . Esto es,

$$\|u - u_h\|_a \leq \|u - v_h\|_a \quad \forall v_h \in V_h.$$

Lo que acabamos de ver es que la aproximación de Galerkin es la mejor aproximación a la solución dentro del espacio  $V_h$ . Señalemos que esto podía haber sido deducido directamente a partir del teorema de la proyección.

Presentemos ahora uno de los resultados de mayor relevancia en este contexto.

**Lema 8.** (*Lema de Cea.*)

Sea  $V$  un espacio de Hilbert. Sean también  $a(\cdot, \cdot)$  una forma bilineal continua y coercitiva y  $F$  un funcional lineal y acotado de  $V$ . Entonces

$$\|u - u_h\|_V \leq \frac{\beta}{\alpha} \inf \{\|u - v\|_V : v \in V_h\},$$

donde  $\alpha$  y  $\beta$  son las constantes de coercitividad y continuidad de  $a(\cdot, \cdot)$ , respectivamente.

*Demostración.* Como la forma  $a(\cdot, \cdot)$  es coercitiva, podemos escribir que

$$\begin{aligned} \alpha \|u - u_h\|_V^2 &\leq a(u - u_h, u - u_h) \\ &= a(u - u_h, u - v_h - u_h + v_h) \\ &= a(u - u_h, u - v_h) - a(e, u_h - v_h) \\ &= a(u - u_h, u - v_h), \end{aligned}$$

para todo  $v_h \in V_h$ , por la ortogonalidad Galerkin (2.6).

Asimismo, como también es continua, se tiene que

$$\alpha \|u - u_h\|_V^2 \leq a(u - u_h, u - v_h) \leq \beta \|u - u_h\|_V \|u - v_h\|_V.$$

Luego,

$$\|u - u_h\|_V \leq \frac{\beta}{\alpha} \|u - v_h\|_V \quad \forall v_h \in V_h.$$

Entonces,

$$\|u - u_h\|_V \leq \frac{\beta}{\alpha} \inf\{\|u - v_h\|_V : v_h \in V_h\}.$$

■

Advirtamos que el lema de Cea nos muestra que la aproximación  $u_h$  es *óptima* en el sentido de que el error,  $\|u - u_h\|_V$ , es proporcional a la mejor aproximación dentro del espacio  $V_h$ .

Debemos apreciar que el lema de Cea transforma el problema de estimar el error  $u - u_h$  en otro de calcular la distancia de  $u$  al subespacio  $V_h$ . Como  $\inf\{\|u - v\| : v \in V_h\} \leq \|u - v_0\|$ , para cualquier  $v_0 \in V_h$ , es posible obtener una buena estimación si se elige  $v_0$  de la manera adecuada. Por ejemplo, si se considera la proyección ortogonal, se obtiene la mejor aproximación posible.

## 2.6. EL MÉTODO DE LOS ELEMENTOS FINITOS.

A la hora de implementar el método de los elementos finitos, debemos escoger un subespacio de aproximación  $S_0^h \subset H_0^1$ . La elección más simple, que será la que sigamos, responde al caso de considerar espacios de *splines* y, dentre todos ellos, sin duda, los más sencillos son los de grado uno. Se comienza fijando una partición,  $P$ , del intervalo  $[a, b]$  de manera que

$$a = x_0 < x_1 < \dots < x_N = b,$$

con  $N \geq 2$ . A menudo, dentro del ámbito de la teoría de los elementos finitos, al conjunto de puntos  $P = \{x_k\}_{k=0}^N$  se lo conoce como *mall*a. Llamando  $I_i := [x_{i-1}, x_i]$ ,  $i = 1, \dots, N$ , el espacio de *splines* de grado uno consiste en

$$S_P^1 := \{u(x) \in C(a, b) : u|_{I_i} \text{ es afín en } I_i\}. \quad (2.8)$$

El subespacio de elementos finitos que se utiliza para aproximar funciones con condiciones de contorno homogéneas es

$$S_{P,0}^1 := S_P^1 \cap H_0^1(a, b) = \{u_h \in S_P^1 : u_h(a) = u_h(b) = 0\}. \quad (2.9)$$



Una base,  $\{\psi_i\}_{i=0}^N$ , del espacio  $S_p^1$  puede ser la compuesta por las denominadas *funciones sombrero* o, si se prefiere, *funciones hat*.

Las funciones sombrero,  $\psi_i$ , son funciones lineales a trozos que se anulan en todos los puntos de la malla salvo en  $x_i$ , donden vale 1. De esta manera, si  $h_i = x_i - x_{i-1}$ , podemos dar una expresión analítica para tales funciones:

$$\psi_i(x) = \begin{cases} (x - x_{i-1})/h_i & \text{si } x \in [x_{i-1}, x_i] \\ (x_{i+1} - x)/h_{i+1} & \text{si } x \in [x_i, x_{i+1}] \\ 0 & \text{en otro caso,} \end{cases}$$

junto con

$$\psi_0(x) = \begin{cases} (x_1 - x)/h_1 & \text{si } x \in [x_0, x_1] \\ 0 & \text{en otro caso,} \end{cases}$$

y

$$\psi_N(x) = \begin{cases} (x - x_{N-1})/h_N & \text{si } x \in [x_{N-1}, x_N] \\ 0 & \text{en otro caso.} \end{cases}$$

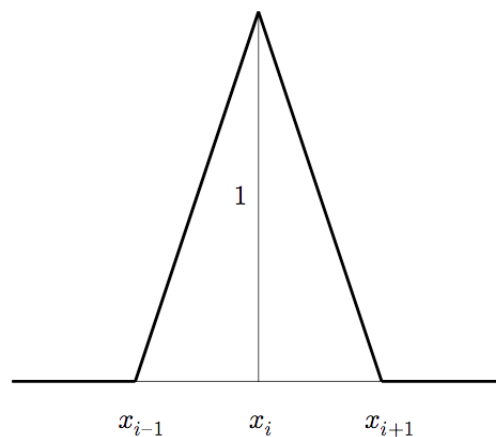


Figura 2.1: Grafo de una función *hat*.

Notemos que estas funciones  $\psi_i$  tienen soporte  $[x_{i-1}, x_{i+1}]$ . Ésta va a ser una propiedad muy importante de las funciones sombrero, pero antes recordemos el marco

que describíamos en la sección previa. Formábamos una matriz  $K_{i,j}$  cuyas entradas venían dadas por el producto interno de los elementos de la base, esto es,

$$K_{i,j} = a(\psi_i, \psi_j).$$

Si ahora consideramos para el problema particular  $[P]$  de la Sección 2.1 la forma bilineal

$$\begin{aligned} a(\psi_i, \psi_j) &= \int_a^b \left\{ p(x) \psi'_i(x) \psi'_j(x) dx + r(x) \psi_i(x) \psi_j(x) \right\} dx \\ &= \int_a^b p(x) \psi'_i(x) \psi'_j(x) dx + \int_a^b r(x) \psi_i(x) \psi_j(x) dx, \end{aligned}$$

para  $1 \leq i, j \leq n-1$ , se observa la propiedad capital de tomar como base las funciones sombrero. En efecto, como consecuencia de sus respectivos soportes, se obtiene que

$$K_{i,j} = 0 \quad \text{si} \quad |i-j| > 1.$$

Por lo tanto, la matriz  $K$  resulta ser tridiagonal. Además, en este caso, es simétrica y definida positiva. Sabemos entonces que el sistema lineal asociado va a ser fácilmente resoluble mediante el uso adecuado de técnicas del análisis numérico tales como la factorización LU.

Notemos que hemos escrito la matriz  $K$  como la suma de otras dos matrices las cuales, por motivos históricos, se denominan *matriz de rigidez* y *matriz de masa*, respectivamente. Asimismo, repararemos también en que los terminos  $K_{i,j}$  son sumamente sencillos pues, de hecho, las derivadas de las funciones  $\psi$  son funciones constantes a trozos en el intervalo  $[a, b]$ .

Aunque pueda parecer que las integrales que han aparecido son fácilmente calculables de forma analítica, debemos reparar en que dependen de  $p(x)$  y  $r(x)$ . Además, en la práctica, pueden no ser calculables explícitamente y, como hemos venido anunciando, habremos de recurrir al análisis numérico para hallar rutas alternativas a fin de superar tales escollos. En particular, vamos a comprobar cómo al utilizar ciertas reglas de cuadratura se obtienen los mismos sistemas de ecuaciones que se obtendrían a partir de otras técnicas, esto es, vamos a aproximar las integrales por la regla del punto medio y obtendremos el sistema propio del método de las diferencias finitas.

Supongamos que los puntos de la malla están equiespaciados. De esta manera,  $x_i = a + ih$ ,  $i = 0, \dots, n$ , si  $h = (b-a)/n$ . Si llamamos  $p_{i-1/2} := p(x_i - h/2)$  y  $p_{i+1/2} := p(x_i + h/2)$ , obtenemos que

$$\int_{x_{i-1}}^{x_i} p(x) \psi'_{i-1}(x) \psi'_i(x) dx = \frac{-1}{h^2} \int_{x_{i-1}}^{x_i} p(x) dx \simeq -p_{i-1/2},$$

y análogamente para el resto de las integrales. Para las relativas a la matriz de masa, valiéndonos en esta ocasión de la regla del trapecio, se tiene que

$$\int_{x_{i-1}}^{x_i} r(x)\psi_{i-1}(x)\psi_i(x)dx \simeq 0,$$

pues  $\psi_i$  se anula en  $x_{i-1}$  y  $\psi_{i-1}$  hace lo propio en  $x_i$ . Asimismo,

$$\int_{x_{i-1}}^{x_i} r(x)\psi_i(x)^2 dx \simeq \frac{hr_i}{2},$$

donde  $r_i := r(x_i)$ , pues  $\psi_i$  vale cero en un extremo del subintervalo y uno en el otro. Así, de la misma forma ocurre que

$$\int_{x_i}^{x_{i+1}} r(x)\psi_i(x)^2 dx \simeq \frac{hr_i}{2}.$$

Si  $f$  es una función continua en el intervalo  $[a, b]$ , aproximando su integral mediante la regla del trapecio y reuniendo todo lo anterior, encontramos que el problema de aproximación de Galerkin toma el aspecto siguiente:

$$-\frac{p_{i-1/2}}{h}u_{i-1} + \frac{p_{i-1/2} + p_{i+1/2}}{h}u_i - \frac{p_{i+1/2}}{h}u_{i+1} + hr_i u_i = hf_i,$$

donde, por supuesto,  $i = 1, \dots, N-1$ ,  $u_0 = \alpha$ ,  $u_N = \beta$  y  $f_i = f(x_i)$ .

Es evidente que lo que hemos obtenido es el mismo sistema de ecuaciones que el de un método en diferencias finitas.

## 2.7. REFERENCIAS BIBLIOGRÁFICAS.

Los temas presentados a lo largo de este capítulo se pueden abordar desde el más severo análisis matemático hasta la rama de la matemática aplicada que es el cálculo de variaciones pasando, por supuesto, por las técnicas de aproximación numérica. En consecuencia, la bibliografía al respecto es sobradamente extensa e inconmensurable. Sin embargo, no por ello vamos a dejar de apuntar aquí los libros que hemos consultado a lo largo de la realización de este trabajo.

- El libro de Suli y Mayers ([21]) constituye una valiosa a la par que clara exposición de prácticamente todos los aspectos del análisis numérico elemental. En concreto, los temas 11 a 14 tratan con el asunto de aproximar ecuaciones en derivadas parciales y han sido una referencia habitual a lo largo del estudio del tema.
- Asimismo, el libro de Daya Reddy ([9]) ha supuesto un gran regalo a la hora de clarificar los aspectos teóricos más importantes y dotar de rigor los resultados expuestos. En él se pueden encontrar prácticamente todas las demostraciones que aparecen en este bloque de la memoria.

- Por supuesto, no podemos dejar de mencionar la obra de Brenner y Scott ([7]). Hemos encontrado en los temas 0 a 4 una exposición clara y muy didáctica de los fundamentos del método de los elementos finitos.
- Finalmente, debemos mencionar que los apuntes de "Complementos de análisis matemático" e "Introducción a los espacios de funciones", suministrados por los profesores de las asignaturas, se han erigido como sendas guías de rápida consulta de prácticamente toda la teoría de análisis funcional, junto con la teoría de distribuciones y espacios de Sobolev que, aunque no se presenta en esta memoria, no hemos dejado de examinar.





# 3

## VALORACIÓN DE OPCIONES.

---

*"No hay ninguna rama de la matemática,  
por abstracta que sea, que no pueda aplicarse  
algún día a los fenómenos del mundo real."*

— Nikolai Lobachevsky.

**Resumen:** Se presenta uno de los tratamientos habituales a la hora de valorar opciones de tipo europeo y americano. Se justifica, a partir de la teoría previa, el uso de métodos de aproximación y se describen los problemas matriciales discretos asociados.





### 3.1. CONSIDERACIONES PREVIAS.

En los capítulos anteriores se han introducido los conceptos relativos a los productos financieros y a la teoría de los elementos finitos. Es ahora el momento de aunar ambos aspectos bajo la búsqueda del valor de ciertos tipos de opciones financieras. Asumiremos para ello que se siguen verificando las mismas hipótesis que supusimos en el capítulo primero de las cuales, la más importante es, sin duda, la que acepta que el precio del activo subyacente sigue un movimiento browniano.

Nos gustaría señalar aquí que, aunque el capítulo que ahora comenzamos constituye el nexo entre los otros dos precedentes, dada la complejidad de esta materia, solo pretendemos ilustrar algunas aplicaciones de la teoría expuesta sin entrar en otros derroteros más profundos. Al final de los resultados que expondremos, indicaremos algunas referencias en las cuales se puede profundizar en este campo.

Si bien en la Sección 1.4 dedujimos la ecuación de Black-Scholes,

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0,$$

lo que nos interesa ahora es hallar el valor de una opción dada una cierta función *payoff*  $\phi$ , esto es, computar efectivamente el valor de

$$V(s, t) = \mathbb{E} \left[ e^{-r(T-t)} \phi(X_T) | X_t = s \right], \quad (3.1)$$

donde  $X_t$  satisface la ecuación diferencial estocástica

$$dX_t = rX_t dt + \sigma X_t dW_t.$$

Se puede encontrar información relativa a la expresión (3.1) en [13] y en [17]. Asimismo, consideraremos los cambios de variables  $\tau := T - t$ , para obtener un problema progresivo, y  $x := \log(s)$  (nótese que estos cambios ya los realizábamos en la Sección 1.5 cuando se resolvía la ecuación de Black-Scholes para la opción europea). Llamaremos entonces

$$u(x, \tau) := V(s, t) \quad (3.2)$$

para tratar así el valor de la opción. De esta manera,  $u(x, 0) = \phi(e^x)$ , y  $u(x, \tau)$  verifica

$$\frac{\partial u}{\partial \tau} = \frac{1}{2}\sigma^2 \frac{\partial^2 u}{\partial x^2} + \left( r - \frac{\sigma^2}{2} \right) \frac{\partial u}{\partial x} - ru, \quad x \in (-\infty, \infty).$$

### 3.2. EL PROBLEMA DISCRETO.

Las funciones valor,  $u$ , de prácticamente la totalidad de los tipos de opciones que son negociados con mayor asiduidad verifican una ecuación en derivadas parciales de la forma

$$\frac{\partial u}{\partial \tau} - \mathcal{A}u = f, \text{ en } I \times (0, T), \quad (3.3)$$

junto con una condición inicial del tipo

$$u(x, 0) = u_0(x), \quad (3.4)$$

y condiciones frontera que, acorde a nuestros intereses, supondremos Dirichlet homogéneas, donde  $I \subseteq \mathbb{R}$  y siendo  $\mathcal{A}$  un operador diferencial lineal de segundo orden.

La formulación variacional del problema (3.3-3.4) se presenta como el problema de

$$\left\{ \begin{array}{l} \text{encontrar } u \in V \text{ tal que} \\ \left( \frac{\partial u}{\partial \tau}, v \right) + a(u, v) = (f, v) \quad \forall v \in V, \\ u(0) = u_0, \end{array} \right. \quad (3.5)$$

donde  $a(u, v) := -(\mathcal{A}u, v)$  es una forma bilineal generalmente no simétrica y  $f \in L^2(I \times (0, T))$ .

Para poder aplicar el método de los elementos finitos va a resultar preciso el diseccionar adecuadamente el dominio espacial y el dominio temporal, esto es, es necesario *discretizar* el problema. Para ello, se utiliza el *método de líneas* el cual consta, a su vez, se dos etapas. En la primera de ellas, solo se discretiza la componente espacial obteniéndose así un sistema de ecuaciones diferenciales ordinarias que es resuelto en la segunda de tales etapas.

Sea  $V_h \subset V$  un subespacio de dimensión  $(N - 1)$ . Recordemos que, como señalábamos en el capítulo 2, es posible elegir  $h = \frac{c}{N}$ ,  $c \in \mathbb{R}$ . Entonces, para cada  $\tau \in (0, T)$  fijo, se aproxima la solución  $u(\tau, x)$  de (3.5) por una función  $u_h(\tau) \in V_h$ . Además, se toma  $u_{h,0}$  como una aproximación de  $u_0$ . De esta forma, se obtiene lo que se conoce como la forma *semidiscreta* del problema (3.5), esto es, dado  $f \in L^2(I \times (0, T))$ ,

$$\left\{ \begin{array}{l} \text{encontrar } u_h \in V_h \text{ tal que para cada } \tau \in (0, T) \\ \left( \frac{\partial u_h}{\partial \tau}, v_h \right) + a(u_h, v_h) = (f, v_h) \quad \forall v_h \in V_h, \\ u_h(0) = u_{h,0}. \end{array} \right. \quad (3.6)$$

Particulatizando  $V_h$  al espacio de elementos finitos lineales a trozos, esto es,  $V_h = \text{span} \{ \psi_i \}_{i=1}^{N-1}$ , donde las  $\psi_i$  son las funciones sombrero que considerábamos en el capítulo 2, escribimos  $u_h$  en términos de ellas,

$$u_h(\tau, x) = \sum_{i=1}^{N-1} u_{h,i}(\tau) \psi_i(x).$$

De la misma forma,

$$u_{h,0}(x) = \sum_{i=1}^{N-1} u_0(x_i) \psi_i(x).$$

Notemos que lo que hemos hecho podría denominarse como *separación de variables* en cierto modo. Así, si  $\mathbf{u}_h$  y  $\mathbf{u}_0$  son los vectores de coeficientes de  $u_h$  y  $u_{h,0}$ , respectivamente, y las matrices de masa y rigidez y el vector  $\mathbf{f}$  vienen dados por

$$\mathbf{M} = (\psi_j, \psi_i), \quad \mathbf{A} = a(\psi_j, \psi_i), \quad \mathbf{f}_i = (f, \psi_i), \quad i, j = 1, \dots, N-1,$$

llegamos de esta manera a la forma matricial de la semidiscretización,

$$\left\{ \begin{array}{l} \text{encontrar } \mathbf{u}_h \in \mathbb{R}^{N-1} \text{ tal que para cada } \tau \in (0, T) \\ \mathbf{M} \dot{\mathbf{u}}_h(\tau) + \mathbf{A} \mathbf{u}_h(\tau) = \mathbf{f}(\tau), \\ \mathbf{u}_h(0) = \mathbf{u}_0. \end{array} \right. \quad (3.7)$$

Sean ahora  $k_m$ ,  $m = 1, \dots, M$ , una secuencia de saltos o pasos, no necesariamente equiespaciados, tales que  $\tau_0 = 0$  y  $\tau_m = \sum_{i=1}^m k_i$  de forma que se cumpla que  $\tau_M = T$ . Aplicando el denominado *método- $\theta$* , obtenemos la forma totalmente discreta del problema,

$$\left\{ \begin{array}{l} \text{encontrar } u_h^m \in V_h \text{ tal que para cada } \tau_m, m = 1, \dots, M, \\ \frac{1}{k_m} (u_h^m - u_h^{m-1}, v_h) + a(u_h^{m-1+\theta}, v_h) = (f^{m-1+\theta}, v_h), \forall v_h \in V_h, \\ u_h^0 = u_{h,0}, \end{array} \right. \quad (3.8)$$

donde  $u_h^{m+\theta} := \theta u_h^{m+1} + (1-\theta) u_h^m$  y  $f^{m+\theta} := \theta f(\tau_{m+1}) - (1-\theta) f(\tau_m)$ . Asimismo,  $\theta$  es una constante que toma valores en  $[0, 1]$  y que depende del método elegido. Los casos habituales corresponden a  $\theta = 0$  (método de Euler implícito),  $\theta = \frac{1}{2}$  (método de Crank-Nicolson),  $\theta = 1$  (método de Euler explícito).

Finalmente, es posible reescribir el problema (3.8) en forma matricial como

$$\left\{ \begin{array}{l} \text{encontrar } \mathbf{u}_h^m \in \mathbb{R}^{N-1} \text{ tal que para cada } m = 1, \dots, M, \\ (\mathbf{M} + k_m \theta \mathbf{A}) \mathbf{u}_h^m = (\mathbf{M} - k_m (1-\theta) \mathbf{A}) \mathbf{u}_h^{m-1} + k_m (\theta \mathbf{f}^m + (1-\theta) \mathbf{f}^{m-1}), \\ \mathbf{u}_h^0 = \mathbf{u}_0. \end{array} \right. \quad (3.9)$$

donde

$$\mathbf{f}^m = [(f(\cdot, \tau_m), \psi_1), \dots, (f(\cdot, \tau_m), \psi_{N-1})]^t$$

y

$$f^{m-1} = [(f(\cdot, \tau_{m-1}), \psi_1), \dots, (f(\cdot, \tau_{m-1}), \psi_{N-1})]^t.$$

En la sección siguiente particularizaremos este esquema general a nuestro problema concreto.

### 3.3. EL PROBLEMA DE LA LOCALIZACIÓN.

Cuando se impone la condición relativa al *payoff*,  $u_0(x) = \phi(e^x)$ ,  $x \in \mathbb{R}$ , el problema ha de ser reformulado, *localizado*, en un dominio acotado donde esta condición pueda ser efectivamente restringida.

Además, es habitual requerir una condición de crecimiento de orden polinómico para el *payoff*, esto es, que se verifique que

$$\phi(s) \leq C(s+1)^q, \quad s \geq 0 \quad (3.10)$$

para cada  $t \in [0, \infty)$ , donde  $C \in \mathbb{R}$  y  $q \in \mathbb{N}$ . La mayor parte de los tipos habituales de opciones verifican esta condición. En particular, las opciones *vanilla*, tanto de compra como de venta.

Partamos del problema de Black-Scholes para  $s \in [0, \infty)$ , esto es,

$$\begin{cases} \frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS \frac{\partial V}{\partial S} - rV = 0, \\ V(s, T) = \phi(s). \end{cases}$$

Si  $s = 0$ , entonces  $V(0, T) = \phi(0)$  y, además,

$$\frac{\partial V}{\partial t}(0, t) = rV.$$

Luego,

$$V(0, t) = Ce^{rt}.$$

Por su parte,

$$V(0, T) = Ce^{rT} = \phi(0),$$

y, por tanto,

$$C = \phi(0)e^{-rT}.$$

Así, se llega a que

$$V(0, t) = \phi(0)e^{-r(T-t)}.$$

El dominio,  $\mathbb{R}$ , del precio del subyacente,  $x = \log(S)$ , se restringe a un intervalo acotado,  $I$ . Sea  $I = (-\delta, \delta)$ <sup>1</sup>,  $\delta > 0$ . Entonces, acorde a (3.2), en los extremos del intervalo  $I$  se tiene que

$$u(-\delta, T - \tau) = V(e^{-\delta}, t) \simeq V(0, t) = \phi(0)e^{-r(T-t)},$$

y

$$u(\delta, \tau) = V(e^{\delta}, T - t) \simeq \begin{cases} \phi(e^{\delta}), & \text{si la opción es de tipo } call, \\ 0, & \text{si la opción es de tipo } put. \end{cases}$$

De esta forma, se plantea el problema en  $I = (-\delta, \delta)$ , para un cierto *payoff*  $\phi$  dado. Utilizaremos  $\tilde{u}$  para aludir a la solución del problema localizado. El aspecto que cobra tal problema es

$$\left\{ \begin{array}{l} -\frac{\partial \tilde{u}}{\partial \tau} + \frac{1}{2}\sigma^2 \frac{\partial^2 \tilde{u}}{\partial x^2} + \left(r - \frac{1}{2}\sigma^2\right) \frac{\partial \tilde{u}}{\partial x} - r\tilde{u} = 0, \quad x \in (-\delta, \delta), \\ \tilde{u}(x, 0) = \phi(e^x), \quad x \in (-\delta, \delta) \\ \tilde{u}(-\delta, \tau) = \phi(0)e^{-r\tau} \\ \tilde{u}(\delta, \tau) = \phi(e^{\delta}). \end{array} \right. \quad (3.11)$$

Al problema (3.11) le denominaremos *problema de Black-Scholes localizado*. Llegados a este punto, una manera habitual de proseguir consiste en transformar el problema anterior en otro equivalente pero con condiciones Dirichlet homogéneas. Para ello, comencemos denotando por  $u_{-\delta}$  y  $u_{\delta}$  a las condiciones de contorno, esto es,

$$u_{-\delta}(\tau) := \phi(0)e^{-r\tau}$$

y

$$u_{\delta} := \phi(e^{\delta}).$$

Sea ahora  $\tilde{W} = \tilde{u} - \left\{ \frac{x + \delta}{2\delta} u_{\delta} - \frac{x - \delta}{2\delta} u_{-\delta} \right\}$ . Derivemos y llevemos las nuevas expresiones a la fórmula original.

<sup>1</sup>Esta elección no constituye ni mucho menos la única posibilidad y su conveniencia depende del tipo de opción que se esté considerando.

- $\frac{\partial \tilde{W}}{\partial \tau} = \frac{\partial \tilde{u}}{\partial \tau} + \frac{x-\delta}{2\delta} \tilde{v}'_{-\delta},$
- $\frac{\partial \tilde{W}}{\partial x} = \frac{\partial \tilde{u}}{\partial x} - \left\{ \frac{u_\delta - u_{-\delta}}{2\delta} \right\},$
- $\frac{\partial^2 \tilde{W}}{\partial x^2} = \frac{\partial^2 \tilde{u}}{\partial x^2}.$

Por su parte, se tiene también que

$$\tilde{u} = \tilde{W} + \left\{ \frac{x+\delta}{2\delta} u_\delta - \frac{x-\delta}{2\delta} u_{-\delta} \right\},$$

y, así, llegamos a que

$$\begin{aligned} & -\frac{\partial \tilde{W}}{\partial \tau} + \frac{x-\delta}{2\delta} \tilde{u}'_{-\delta} + \frac{1}{2} \sigma^2 \frac{\partial^2 \tilde{W}}{\partial x^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial \tilde{W}}{\partial x} - r \tilde{W} \\ & + \left( r - \frac{1}{2} \sigma^2 \right) \left\{ \frac{u_\delta - u_{-\delta}}{2\delta} \right\} - r \left\{ \frac{(x+\delta)}{2\delta} u_\delta - \frac{(x-\delta)}{2\delta} u_{-\delta} \right\} = 0. \end{aligned}$$

Podemos ahora hablar del *problema de Black-Scholes localizado con condiciones Dirichlet homogéneas*, esto es,

$$\left\{ \begin{array}{l} -\frac{\partial \tilde{W}}{\partial \tau} + \frac{1}{2} \sigma^2 \frac{\partial^2 \tilde{W}}{\partial x^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial \tilde{W}}{\partial x} - r \tilde{W} = F(x, \tau), \\ \tilde{W}(x, 0) = \phi(e^x) - \left\{ \frac{(x+\delta)}{2\delta} u_\delta - \frac{(x-\delta)}{2\delta} u_{-\delta} \right\}, \\ \tilde{W}(-\delta, \tau) = 0, \\ \tilde{W}(\delta, \tau) = 0, \end{array} \right. \quad (3.12)$$

donde

$$F(x, \tau) = -\frac{(x-\delta)}{2\delta} \tilde{u}'_{-\delta} - \left( r - \frac{1}{2} \sigma^2 \right) \left\{ \frac{u_\delta - u_{-\delta}}{2\delta} \right\} + r \left\{ \frac{(x+\delta)}{2\delta} u_\delta - \frac{(x-\delta)}{2\delta} u_{-\delta} \right\}. \quad (3.13)$$

En cuanto a la cuestión de la convergencia, es posible probar (ver [14]) que  $\tilde{u} \rightarrow u$  cuando  $\delta \rightarrow \infty$  uniformemente en los intervalos compactos. Al intervalo  $I = (-\delta, \delta)$  se lo denomina como *dominio computacional*. En la práctica, lo que se hace es tomar un  $\delta_1 < \delta$

fijo y considerar el intervalo  $[-\delta_1, \delta_1]$ . A este nuevo conjunto se lo conoce como *dominio de aproximación*. Entonces, según se desprende del artículo [14], se tiene que

$$|u(\tau, x) - \tilde{u}(\tau, x)| \leq D \exp\{-\gamma_1\delta + \gamma_2|x|\} \leq D \exp\{-\gamma_1\delta + \gamma_2\delta_1\}, \quad d \in \mathbb{R},$$

para todo  $x \in [-\delta_1, \delta_1]$ . De esta manera,

$$\max_{x \in [-\delta_1, \delta_1]} |u(\tau, x) - \tilde{u}(\tau, x)| \leq D \exp\{-\gamma_1\delta + \gamma_2\delta_1\} \rightarrow 0, \text{ si } \delta \rightarrow \infty.$$

Además, tal convergencia es uniforme en  $[-\delta_1, \delta_1]$  y con velocidad exponencial, lo cual posibilita la efectiva computabilidad del método. Pasamos ahora a formular en términos variacionales el problema de Black-Scholes localizado con condiciones Dirichlet homogéneas.

Definamos para nuestros intereses la forma bilineal  $a_{BS}(\cdot, \cdot) : H_0^1(I) \times H_0^1(I) \rightarrow \mathbb{R}$ , con  $I = (-\delta, \delta)$ , dada por

$$a_{BS}(\psi, \phi) = \frac{1}{2}\sigma^2(\psi', \phi') + (\sigma^2/2 - r)(\psi', \phi) + r(\psi, \phi). \quad (3.14)$$

Notemos que esta forma bilineal no es simétrica debido a la presencia del término asociado al *drift*.

La forma bilineal  $a_{BS}$  que acabamos de definir es continua pues, como consecuencia de la desigualdad de Schwarz, se tiene que

$$\begin{aligned} |a_{BS}(\psi, \phi)| &\leq \frac{1}{2}\sigma^2\|\psi'\|_{L^2(I)}\|\phi'\|_{L^2(I)} + |\sigma^2/2 - r|\|\psi'\|_{L^2(I)}\|\phi\|_{L^2(I)} + r\|\psi\|_{L^2(I)}\|\phi\|_{L^2(I)} \\ &\leq c\|\psi\|_{H_0^1(I)}\|\phi\|_{H_0^1(I)}, \end{aligned}$$

con  $c \in \mathbb{R}$ , donde se ha utilizado la desigualdad de Poincaré.

Por otra parte, se tiene que también es coercitiva en  $H_0^1(I)$  pues, si reparamos primero en que

$$\int_I \psi' \psi dx = \frac{1}{2} \int_I (\psi')^2 dx = 0, \quad \forall \psi \in H_0^1(I),$$

entonces se tiene que

$$\begin{aligned} |a_{BS}(\psi, \psi)| &= \frac{1}{2}\sigma^2\|\psi'\|_{L^2(I)}^2 + r\|\psi\|_{L^2(I)}^2 \\ &\geq \frac{1}{2}\sigma^2\|\psi'\|_{L^2(I)}^2 = \frac{1}{2}\sigma^2\|\psi\|_{H_0^1(I)}^2. \end{aligned}$$

Presentemos la formulación débil del problema que estamos estudiando.

**Definición 18.** (Formulación variacional del problema de Black-Scholes localizado.)

Sean  $a_{BS}(\cdot, \cdot)$  la forma bilineal definida en (3.14),  $I = (-\delta, \delta)$ ,  $\delta > 0$ , y  $F(\cdot, \cdot) \in L^2(I \times (0, T))$  dada por (3.13). Se define la formulación variacional del problema de Black-Scholes localizado como la búsqueda de  $w \in H_0^1(I)$  tal que

$$\left( \frac{\partial w}{\partial \tau}, v \right) + a_{BS}(w, v) = (F, w) \quad \forall v \in H_0^1(I),$$

$$y w(0) = w_0 = \phi(e^x) - \left\{ \frac{(x+\delta)}{2\delta} u_\delta - \frac{(x-\delta)}{2\delta} u_{-\delta} \right\}.$$

Por su parte, existe un resultado que avala el uso del método de Galerkin a la hora de buscar soluciones aproximadas del problema de Black-Scholes localizado.

**Teorema 4.**

La formulación variacional del problema de Black-Scholes es un problema bien puesto.

El teorema anterior es consecuencia de que la forma bilineal  $a_{BS}(\cdot, \cdot)$  es continua y coercitiva en el espacio correspondiente. Es posible encontrar un estudio detallado al respecto de la existencia y la unicidad de la solución de ecuaciones parabólicas en [18].

En virtud del Teorema 4, sabemos que la formulación variacional en un dominio acotado que acabamos de describir es un problema bien puesto. Por tanto, en particular existe una única solución y es posible aplicar el método de Galerkin.

A partir de la formulación variacional del problema, se aproxima  $\tilde{W}$ , para cada instante  $\tau$  fijo, por una función  $w_h \in V_h$  donde  $V_h = S_P^1 \cap H_0^1(-\delta, \delta)$ , donde  $S_P^1$  es como en (2.8) y  $h = \frac{2\delta}{N}$ . Si, además, establecemos que  $w_0(x) = \phi(e^x) - \left\{ \frac{(x+\delta)}{2\delta} u_\delta - \frac{(x-\delta)}{2\delta} u_{-\delta} \right\}$ , que la malla sea uniforme de amplitud  $h$  y en instantes de tiempo equiespaciados, dígame,  $\tau_0 < \tau_1 < \dots < \tau_M$ , con  $|\tau_{i+1} - \tau_i| = k$ , procediendo como en la sección previa, obtenemos el problema matricial computable siguiente,

$$\left\{ \begin{array}{l} \text{encontrar } \mathbf{w}_h^{m+1} \in \mathbb{R}^{N-1} \text{ tal que para } m = 0, \dots, M-1 \\ (\mathbf{M} + k\theta \mathbf{A}^{BS}) \mathbf{w}_h^{m+1} = (\mathbf{M} - k(1-\theta) \mathbf{A}^{BS}) \mathbf{w}_h^m + k(\theta \mathbf{F}^{m+1} + (1-\theta) \mathbf{F}^m), \\ \mathbf{w}_h^0 = \mathbf{w}_0, \end{array} \right. \quad (3.15)$$

donde  $\mathbf{w}_0$  es el vector de las evaluaciones de  $w_0(x)$  en los nodos, con  $\mathbf{M} = \left\{ (\psi_i, \psi_j) \right\}_{\substack{1 \leq i \leq N-1 \\ 1 \leq j \leq N-1}}$ ,  $\mathbf{A}^{BS} = \left\{ a_{BS}(\psi_i, \psi_j) \right\}_{\substack{1 \leq i \leq N-1 \\ 1 \leq j \leq N-1}}$  y

$$\mathbf{F}^m = [(F(\cdot, \tau_m), \psi_1), \dots, (F(x, \tau_m), \psi_{N-1})]^t,$$



$$\mathbf{F}^{m-1} = [(F(\cdot, \tau_m), \psi_1), \dots, (F(\cdot, \tau_m), \psi_{N-1})]^t,$$

donde  $F(x, \tau)$  viene dado por (3.13).

Por su parte, los valores  $F(x, \tau)$  se aproximan por

$$F(x, \tau) \simeq \sum_{j=0}^N F(x_j, \tau) \psi_j(x),$$

y, a su vez,

$$(F(\cdot, \tau_m), \psi_k) \simeq \sum_{j=0}^N F(x_j, \tau_m) (\psi_j, \psi_k), \quad (3.16)$$

para  $k = 1, \dots, N-1$ .

En el caso que estamos considerando, esto es, donde las funciones base son las funciones sombrero, la matriz  $\mathbf{M}$  cobra el siguiente aspecto

$$\mathbf{M} = \frac{h}{6} \begin{bmatrix} 4 & 1 & & & \\ 1 & \ddots & \ddots & & \\ & \ddots & \ddots & 1 & \\ & & & 1 & 4 \end{bmatrix}. \quad (3.17)$$

Asimismo, de acuerdo con (3.16), requerimos el uso de otra matriz adicional que incorpore los extremos del intervalo en el que estamos trabajando. Esta matriz es

$$\mathbf{M}^{ext} = \frac{h}{6} \left[ \begin{array}{c|ccc|c} 1 & 4 & 1 & & 0 \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & & & 1 & 4 & 1 \end{array} \right]. \quad (3.18)$$

Así, los vectores  $\mathbf{F}^{m+1}$  y  $\mathbf{F}^m$  que efectivamente se implementan vienen dados por

$$\mathbf{F}^{m+1} = \mathbf{M}^{ext} \begin{bmatrix} F(x_0, \tau_{m+1}) \\ \vdots \\ F(x_N, \tau_{m+1}) \end{bmatrix}$$

y

$$\mathbf{F}^m = \mathbf{M}^{ext} \begin{bmatrix} F(x_0, \tau_m) \\ \vdots \\ F(x_N, \tau_m) \end{bmatrix}.$$

Además, es posible escribir  $A^{BS}$  de forma explícita como

$$A^{BS} = \frac{\sigma^2}{2} \mathbf{K} + \left( \frac{\sigma^2}{2} - r \right) \mathbf{B} + r \mathbf{M}, \quad (3.19)$$

donde

$$\mathbf{K} = \frac{1}{h} \begin{bmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{bmatrix},$$

y

$$\mathbf{B} = \frac{1}{2} \begin{bmatrix} 0 & 1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & 1 \\ & & -1 & 0 \end{bmatrix}.$$

Una vez hallado  $\tilde{W}$ , se deshace el cambio para recuperar  $\tilde{u}$ , esto es,

$$\tilde{u} = \tilde{W} + \left\{ \frac{x + \delta}{2\delta} u_{\delta} - \frac{x - \delta}{2\delta} u_{-\delta} \right\}.$$

### 3.4. EXPERIMENTOS NUMÉRICOS.

En esta sección vamos a mostrar y a comentar algunos resultados que han sido obtenidos para la ocasión con el programa *BSeuropea*, el cual puede encontrarse en el Apéndice A.

El programa está implementado siguiendo el esquema (3.15) para  $\theta = \frac{1}{2}$ , esto es, el método de Crank-Nicolson. El método de Crank-Nicolson es un método estable de orden  $\mathcal{O}(k^2 + h^2)$  que requiere prácticamente el mismo trabajo que el método de Euler implícito, el cual es también estable pero de orden  $\mathcal{O}(k + h^2)$ . En lo que sigue,  $N$  será el número de subintervalos considerado, mientras que  $k$  será el incremento temporal a cada etapa.

Fijados los valores  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$ ,  $K = 15$ ,  $\delta = 10$  para el caso de la opción *put* (el de la *call* es totalmente análogo), puede comprobarse en la Figura 3.1 como la aproximación generada por nuestro programa *BSeuropea* (en azul) resulta, hasta para un número de subintervalos relativamente pequeño ( $N = 100$ ) y un incremento temporal relativamente grande ( $k = 0.5$ ), prácticamente indistinguible de la solución exacta (en magenta).

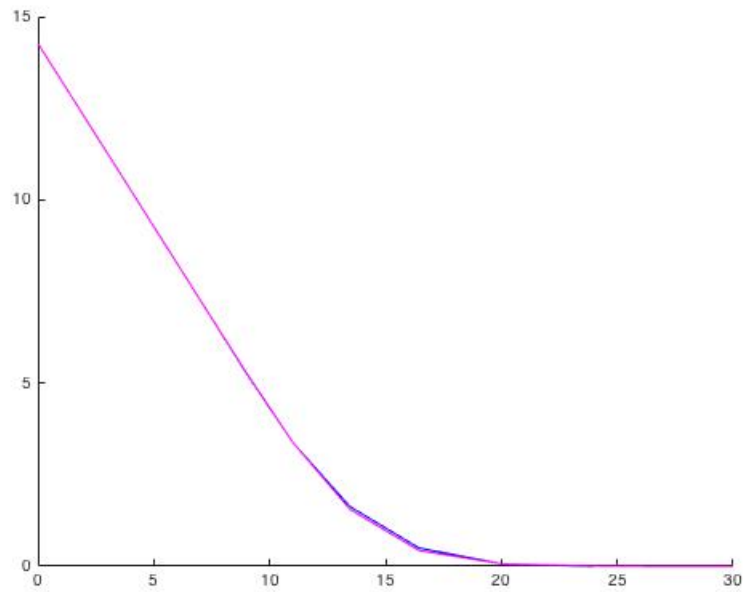


Figura 3.1: Aproximación opción *put* europea ( $N = 100$ ,  $k = 0.5$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$ ,  $K = 15$ ,  $\delta = 10$ ).

Necesitamos aumentar la imagen notablemente para poder apreciar el error cometido (Figura 3.2).

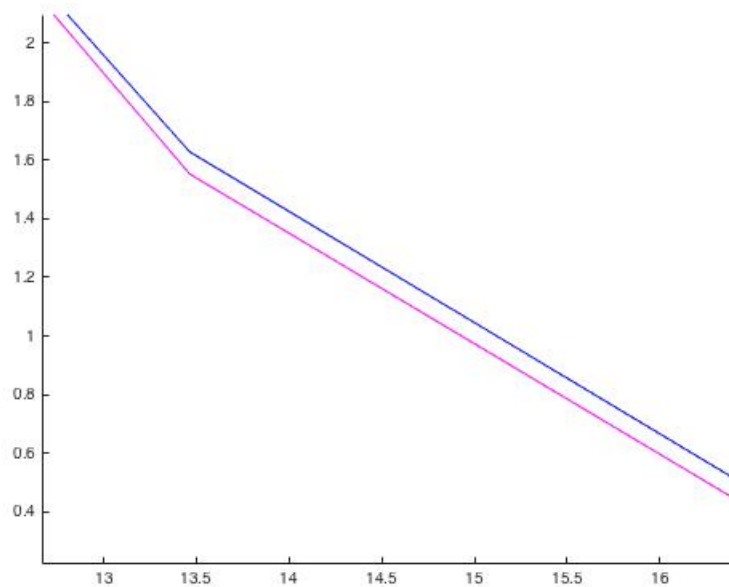


Figura 3.2: Aproximación opción *put* europea ( $N = 100$ ,  $k = 0.5$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$ ,  $K = 15$ ,  $\delta = 10$ ) (Extracto aumentado).

Al aumentar el número de subintervalos ( $N = 400$ ) y disminuir el incremento temporal ( $k = 0.0125$ ), la aproximación mejora muy rápidamente. Podemos ver como las curvas se suavizan en la Figura 3.3. Reparemos, además, en la escala de la Figura 3.4.

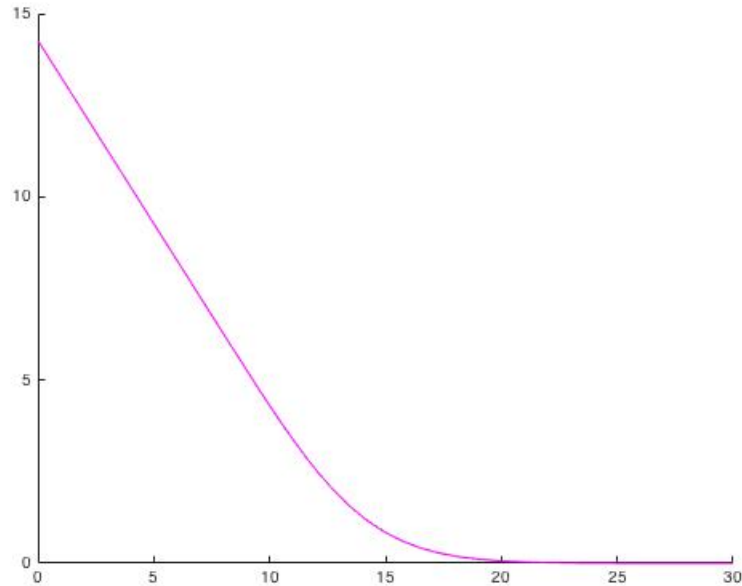


Figura 3.3: Aproximación opción *put* europea ( $N = 400$ ,  $k = 0.0125$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$ ,  $K = 15$ ,  $\delta = 10$ ) (Extracto aumentado).

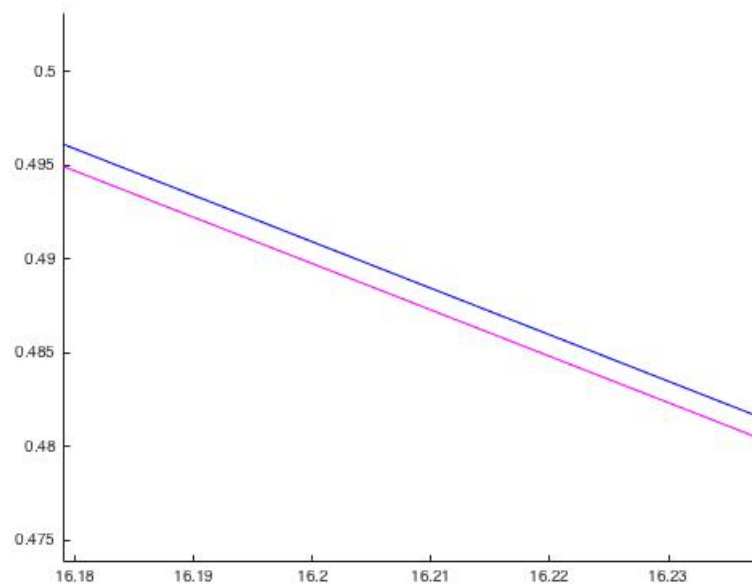


Figura 3.4: Aproximación opción *put* europea ( $N = 400$ ,  $k = 0.0125$ ,  $\sigma = 0.2$ ,  $r = 0.05$ ,  $T = 1$ ,  $K = 15$ ,  $\delta = 10$ ) (Extracto aumentado).

A continuación, mostramos en una tabla los errores, medidos en la norma del máximo, entre la aproximación calculada por *BSeuropea* y la función *blsprice* de MatLab para distintos valores de los incrementos espaciales y temporales en el intervalo  $\left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$ . Al observar el comportamiento diagonalmente, se comprueba que la tendencia es la correspondiente al orden del método,  $\mathcal{O}(\Delta\tau^2 + \Delta x^2)$ , esto es, que al dividir  $\Delta\tau$  y  $\Delta x$  a la mitad, el error se divide entre 4.

	0.2	0.1	0.05	0.025	0.0125	0.00625	$\Delta x$
0.5	7.5667e-02	5.3509e-02	8.3597e-02				
0.25	7.3514e-02	1.3539e-02	2.4945e-02				
0.125	7.3173e-02	1.1527e-02	3.5808e-03				
0.0625				1.8142e-03			
0.03125					5.7187e-04		
0.015625						2.2263e-04	
$\Delta\tau$							

Tabla 3.1: Errores medidos en la norma del máximo para distintos incrementos espaciales y temporales en el intervalo  $\left[-\frac{\delta}{2}, \frac{\delta}{2}\right]$  ( $\sigma = 0.2, r = 0.05, T = 1, K = 15, \delta = 10$ ).

En la Figura 3.5 se muestra, en escala doblemente logarítmica la evolución del error al dividir a la mitad los incrementos espacial y temporal. Además, se ha dibujado en línea discontinua la pendiente que cabe esperar.

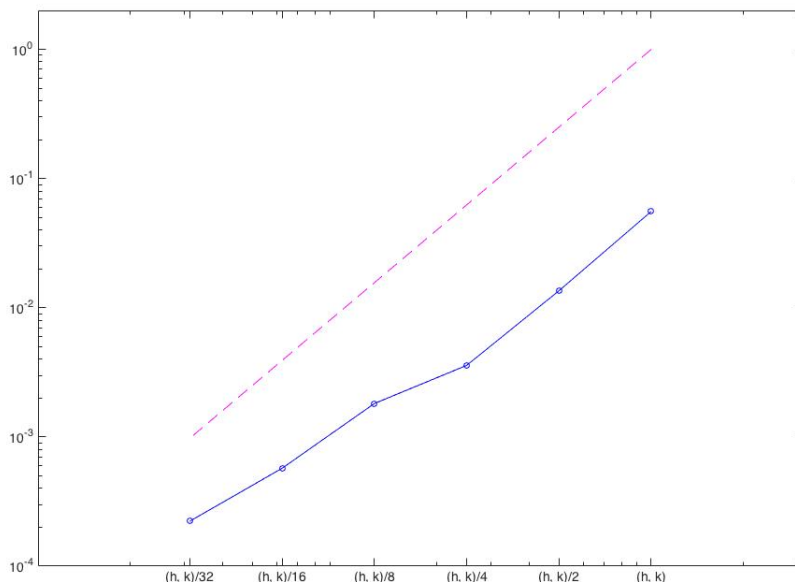


Figura 3.5: Gráfica en escala doblemente logarítmica de la error conforme se dividen  $\Delta\tau$  y  $\Delta x$  a la mitad ( $\sigma = 0.2, r = 0.05, T = 1, K = 15, \delta = 10$ ).

### 3.5. LA OPCIÓN AMERICANA.

Valorar opciones americanas requiere de la solución de problemas de *parada óptima* sobre el proceso del activo subyacente. Al igual que ocurría en el caso de la opción europea, la solución de tales problemas goza de una caracterización a partir de la teoría de la probabilidad. En este caso, el valor cumple un sistema de inecuaciones en derivadas parciales.

Por su parte, si asumimos como venimos haciendo que no se abonan dividendos al poseedor de la opción, se puede probar que el valor de una opción americana de compra ha de ser el mismo que el correspondiente a otra opción europea de compra. Por ello, los métodos que describiremos están orientados a la opción de venta.

Dada una filtración,  $\mathcal{F}_t$ , un *tiempo de parada*,  $\tau$ , es una variable aleatoria que toma valores en  $(0, \infty)$  y que satisface que

$$\{\tau \leq t\} \in \mathcal{F}_t \quad \forall t \geq 0.$$

Si ahora denotamos por  $\mathcal{T}_{t,T}$  al conjunto de todos los tiempos de parada para el proceso  $S_t$ , con valores en el intervalo  $(t, T)$ , el valor de la opción americana viene dado por

$$V(t, s) := \sup_{\tau \in \mathcal{T}_{t,T}} \mathbb{E} [e^{-r(\tau-t)} \phi(S_\tau) | S_t = s].$$

Se puede encontrar un desarrollo más exhaustivo de estos conceptos en [13] y en [20].

Al igual que en el caso de las opciones europeas, existe una relación muy estrecha entre la representación probabilística del valor de la opción americana y una ecuación determinista en forma de EDP. Para facilitar la escritura, definamos el operador  $A_{BS}$  para funciones  $f \in C^2(\mathbb{R})$  como

$$A_{BS}(f) := \frac{1}{2} \sigma^2 \frac{\partial^2 f}{\partial x^2} + \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial f}{\partial x}. \quad (3.20)$$

**Teorema 5.** (*Sistema de inecuaciones para la opción americana.*)

Sea  $A_{BS}$  el operador que hemos definido en (3.20). Sea también  $u(\tau, x)$  una función que satisface el sistema

$$\left\{ \begin{array}{l} \frac{\partial u}{\partial \tau} - A_{BS}(u) + r u \geq 0 \\ u(\tau, x) \geq \phi(e^x) \\ \left( \frac{\partial u}{\partial \tau} - A_{BS}(u) + r u \right) (\phi - u) = 0 \\ u(0, x) = \phi(e^x). \end{array} \right. \quad (3.21)$$

Entonces,  $V(T - t, e^x) = u(\tau, x)$ .

Es posible encontrar una prueba detallada de este resultado en [3]. Para cada  $t \in (0, T)$  existe lo que se denomina como *precio óptimo de ejercicio*,  $s^*(t)$ , que toma valores en  $(0, K)$  y que, para  $s \leq s^*(t)$ , se tiene que  $V(t, s) = \phi(s)$ . La región

$$\mathcal{C} := \{(t, s) : s > s^*(t)\}$$

se denomina *región de continuación*. Por su parte, su complementario es lo que se conoce como *región de ejercicio*. Como el precio óptimo de ejercicio no es conocido *a priori*, se dice que es una *condición de contorno libre* y el problema de determinar el valor de la opción es un *problema de contorno libre*. Nótese que esta condición relativa a  $s^*(t)$  no aparece en las condiciones del teorema previo (3.21) lo cual justifica esta reformulación.

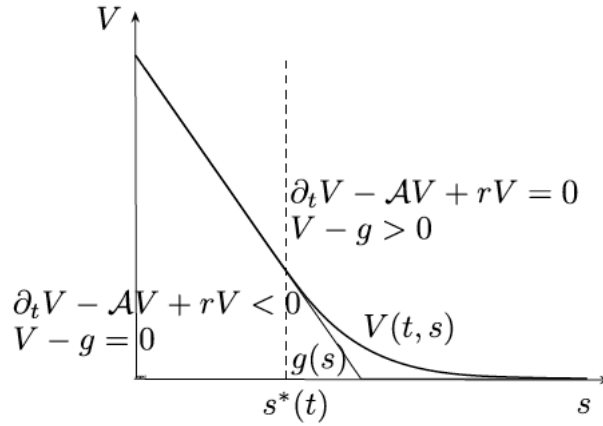


Figura 3.6: Opción americana de venta. (Imagen tomada de [13]).

Al igual que se hizo en la sección precedente, en el caso de las opciones americanas también se ha de localizar el problema, esto es, restringirlo a un intervalo  $I = (-\delta, \delta)$ ,  $\delta > 0$ . Vamos a presentar la formulación variacional del problema de valoración de este tipo de opciones. En estas condiciones, el conjunto de funciones admisibles para el problema débil asociado a (3.21) es el convexo

$$\mathcal{K}_\phi := \{v \in H_0^1(I) : v \geq \phi \text{ para todo } x \in I\}.$$

**Definición 19.** (Formulación variacional del problema localizado.)

Sea  $a_{BS}(\cdot, \cdot)$  la forma bilineal definida en (3.14). Se define la formulación variacional del problema de valoración de una opción americana en el dominio acotado  $I = (-\delta, \delta)$  como la búsqueda de  $u \in \mathcal{K}_\phi$  tal que

$$\left( \frac{\partial u}{\partial t}, v - u \right) + a_{BS}(u, v - u) \geq -a_{BS}(\phi, v - u) \quad \forall v \in \mathcal{K}_\phi,$$

y  $u(0) = u_0 = \phi(e^x)$ .

Puesto que la forma bilineal  $a_{BS}(\cdot, \cdot)$  definida en (3.14) es continua y coercitiva en  $H_0^1(I)$ , según el teorema 4, la formulación variacional para la opción americana admite una única solución.

Atendiendo al artículo [14], es posible obtener una cota para el error entre la solución del problema original y la del problema localizado válida siempre que la función *payoff* cumpla la condición de crecimiento (3.10).

Asimismo, igual que en el caso de la opción europea, denotaremos por  $\tilde{u}$  a la solución del problema localizado. De esta manera, al restringir el problema al intervalo  $I$  es posible fijar las condiciones frontera y, así, el sistema de inecuaciones (3.21) cobra el aspecto siguiente,

$$\left\{ \begin{array}{l} \frac{\partial \tilde{u}}{\partial \tau} - A_{BS}(\tilde{u}) + r\tilde{u} \geq 0, \quad x \in (-\delta, \delta), \\ \tilde{u}(x, \tau) \geq \phi(e^x), \quad x \in (-\delta, \delta), \\ \left( \frac{\partial \tilde{u}}{\partial \tau} - A_{BS}(\tilde{u}) + r\tilde{u} \right) (\phi - \tilde{u}) = 0, \quad x \in (-\delta, \delta), \\ \tilde{u}(x, 0) = \phi(e^x), \quad x \in (-\delta, \delta), \\ \tilde{u}(\pm\delta, \tau) = \phi(e^{\pm\delta}). \end{array} \right. \quad (3.22)$$

Podemos proceder ahora como lo hicimos en el caso de la opción europea, esto es, transformando el problema en otro equivalente pero con condiciones Dirichlet homogéneas. Para ello, denotamos por  $u_\delta$  y  $u_{-\delta}$  a las condiciones de contorno. Puesto que las opciones *call* no presentan interés en este tipo de contratos, nos ceñiremos al caso de las *put*. De esta manera, se tiene que

$$u_\delta := 0$$

y

$$u_{-\delta} := K.$$

Sea ahora  $\tilde{W} = \tilde{u} - \left\{ \frac{x+\delta}{2\delta} u_\delta - \frac{x-\delta}{2\delta} u_{-\delta} \right\} = \tilde{u} + \frac{(x-\delta)}{2\delta} K$ . Realicemos las derivaciones pertinentes.

- $\frac{\partial \tilde{W}}{\partial \tau} = \frac{\partial \tilde{u}}{\partial \tau},$
- $\frac{\partial \tilde{W}}{\partial x} = \frac{\partial \tilde{u}}{\partial x} + \frac{K}{2\delta},$
- $\frac{\partial^2 \tilde{W}}{\partial x^2} = \frac{\partial^2 \tilde{u}}{\partial x^2}.$



También se tiene que

$$\tilde{u} = \tilde{W} - \frac{(x - \delta)}{2\delta} K,$$

y, así, llegamos a que

$$\begin{aligned} \frac{\partial \tilde{W}}{\partial \tau} - \frac{1}{2} \sigma^2 \frac{\partial^2 \tilde{W}}{\partial x^2} - \left( r - \frac{1}{2} \sigma^2 \right) \frac{\partial \tilde{W}}{\partial x} + r \tilde{W} \\ + \left( r - \frac{1}{2} \sigma^2 \right) \frac{K}{2\delta} - r K \frac{(x - \delta)}{2\delta} \geq 0. \end{aligned}$$

Si llamamos

$$F(x) := - \left( r - \frac{1}{2} \sigma^2 \right) \frac{K}{2\delta} + r K \frac{(x - \delta)}{2\delta}, \quad (3.23)$$

el problema con condiciones Dirichlet homogéneas pasa a ser

$$\left\{ \begin{array}{l} \frac{\partial \tilde{W}}{\partial \tau} - A_{BS}(\tilde{W}) + r \tilde{W} \geq F(x), \quad x \in (-\delta, \delta), \\ \tilde{W}(x, \tau) \geq \phi(e^x) + \frac{x - \delta}{2\delta} K, \quad x \in (-\delta, \delta), \\ \left( \frac{\partial \tilde{W}}{\partial \tau} - A_{BS}(\tilde{W}) + r \tilde{W} - F(x) \right) \left( \phi(e^x) + \frac{x - \delta}{2\delta} K - \tilde{W} \right) = 0, \quad x \in (-\delta, \delta), \\ \tilde{W}(x, 0) = \phi(e^x) + \frac{x - \delta}{2\delta} K, \quad x \in (-\delta, \delta), \\ \tilde{W}(\pm\delta, \tau) = 0. \end{array} \right. \quad (3.24)$$

Como en el caso de la opción europea, vamos a considerar para discretizar el problema una malla uniforme con  $h = \frac{2\delta}{N}$  e instantes de tiempo equiespaciados, dígase,  $\tau_0 < \tau_1 < \dots < \tau_M$ , con  $|\tau_{i+1} - \tau_i| = k$ . Además, establecemos también que  $w_0(x) = \phi(e^x) + \frac{x - \delta}{2\delta} K$ .

$$\left\{ \begin{array}{l} \text{encontrar } \mathbf{w}_h^{m+1} \in \mathbb{R}^{N-1} \text{ tal que para } m = 0, \dots, M-1 \\ (\mathbf{M} + k\theta \mathbf{A}^{BS}) \mathbf{w}_h^{m+1} \geq (\mathbf{M} - k(1 - \theta) \mathbf{A}^{BS}) \mathbf{w}_h^m + k\theta \mathbf{F}, \\ \mathbf{w}_h^{m+1} \geq \mathbf{w}_h^0, \\ [(\mathbf{M} + k\theta \mathbf{A}^{BS}) \mathbf{w}_h^{m+1} - (\mathbf{M} - k(1 - \theta) \mathbf{A}^{BS}) \mathbf{w}_h^m - k\theta \mathbf{F}] (\mathbf{w}_h^0 - \mathbf{w}_h^m) = 0, \\ \mathbf{w}_h^0 = \mathbf{w}_0, \end{array} \right. \quad (3.25)$$

donde  $w_0$  es el vector que contiene las evaluaciones de  $w_0(x)$  en los nodos y las matrices  $M$  y  $A^{BS}$  vienen dadas por (3.17) y (3.19), respectivamente. Por su parte, el vector  $F$  viene dado por

$$F = M^{ext} \begin{bmatrix} F(x_0) \\ \vdots \\ F(x_N) \end{bmatrix},$$

donde  $F(\cdot)$  está definido como en (3.23) y la matriz  $M^{ext}$  es (3.18).

Una vez calculado  $\tilde{W}$  es preciso deshacer el cambio para recuperar la solución buscada, es decir,

$$\tilde{u} = \tilde{W} - \frac{(x - \delta)}{2\delta} K.$$

Una posibilidad para aproximar efectivamente el valor de una opción americana pasa por valerse de las opciones bermúdeas. Estas opciones pueden ser ejercidas en una cantidad finita de instantes intermedios antes de la fecha de expiración. En concreto, se fijan un conjunto de fechas

$$0 = \tau_0 < \dots < \tau_N = T$$

en los que es posible ejercer la opción. Para aproximar el valor de la americana, se toman los incrementos temporales relativamente pequeños. Por ejemplo, se pueden elegir como el paso temporal en el problema (3.25). En esa situación, habría que tomar, a cada iteración,

$$w_h^{m+1} = \max\{w_0, w_h^m\},$$

tomando el máximo componente a componente.

En el Apéndice A puede encontrarse este método programado en MatLab. Al ejecutarlo para los valores  $\sigma = 0.2$ ,  $r = 0.1$ ,  $\delta = 10$ ,  $K = 12$  y  $T = 1$ , con  $N = 400$  y un incremento temporal  $inct = \frac{1}{365}$ , correspondiente, en términos financieros, a una posibilidad diaria de ejercer la opción, se obtiene la Figura 3.7, en la cual se ha representado, además, el la función *payoff* en azul.

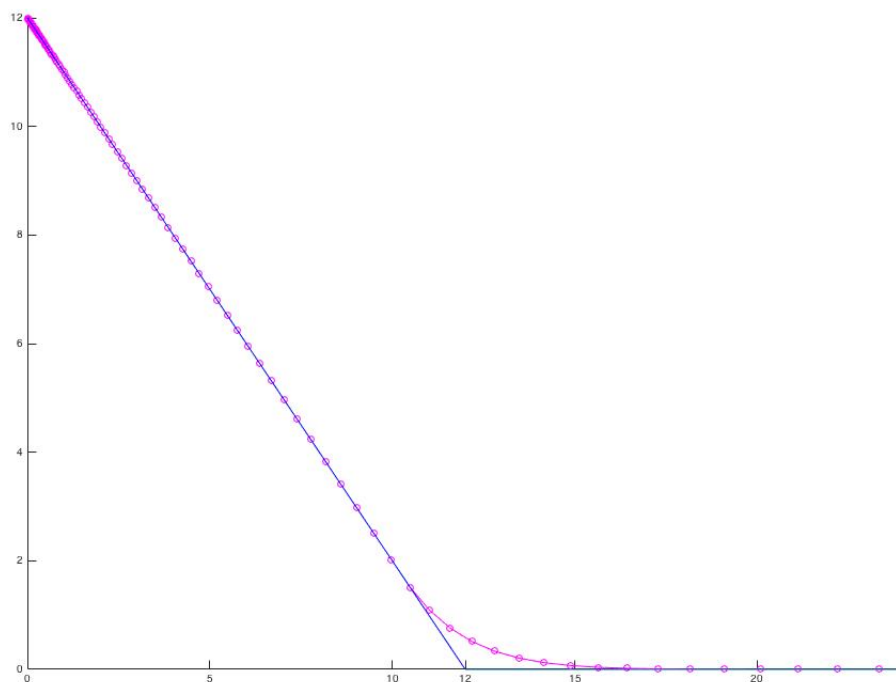


Figura 3.7: Aproximación opción americana ( $N = 400$ ,  $k = 0.0125$ ,  $\sigma = 0.2$ ,  $r = 0.1$ ,  $\delta = 10$ ,  $K = 12$ ,  $T = 1$ ).

Otras posibilidades recurren al denominado *método SOR proyectado (PSOR)* o al *método de la penalización (penalty)*. Respecto al primero, puede encontrarse una descripción del mismo en [13]. Por su parte, respecto al método de la penalización, el artículo [11] contiene un análisis detallado del mismo.

### 3.6. REFERENCIAS BIBLIOGRÁFICAS.

El problema de aproximar la solución de la ecuación de Black-Scholes para distintas opciones engloba, según hemos podido comprobar durante el estudio de tal materia, una gran amalgama de matices que pasan por adaptar los métodos numéricos a la idiosincracia particular de cada tipo de opción. Si bien, lo que mayormente hemos expuesto ha sido fruto del trabajo realizado en numerosas sesiones de tutoría, no por ello hemos dejado de consultar abundantes referencias bibliográficas de las cuales las más reseñables enumeramos a continuación.

- El libro de Hilber, Reichmann, Schwab y Winter ([13]) hace una clara exposición de los procedimientos habituales cuando se trata el problema de la valoración de opciones. En particular, conciernen a esta parte de la memoria los capítulos 2 a 5.
- Por su parte, la obra de Seydel ([20]) también presenta el método de elementos finitos aplicado al problema de la valoración de opciones. En concreto, es de interés el capítulo 5.

- Asimismo, la obra de Achdou y Pironneau ([1]) constituye una célebre referencia al respecto de los distintos métodos numéricos y, en particular, del método de elementos finitos, utilizados para valorar opciones financieras.
- Por supuesto, las notas proporcionadas por el tutor de este trabajo, así como distintos artículos ([11],[12],[14],[8]) han constituido el eje central de esta última parte de la memoria y componen, sin lugar a dudas, las mejores y más adecuadas referencias respecto a la materia que hemos estudiado.





# COMENTARIOS FINALES.

Aunque existen fórmulas cerradas para los modelos más simples de valoración de derivados financieros, la mayor parte de las ocasiones es preciso recurrir a métodos de aproximación numérica. Dado que existe una amplia amalgama de técnicas distintas, es necesario saber discernir la conveniencia de unos métodos a otros según cada problema concreto. Y solo mediante un conocimiento profundo de tales métodos es posible llevar a cabo tal tarea.

Por otra parte, en las últimas décadas se ha desarrollado un excelso campo de investigación en derredor a la teoría de valoración de opciones. Tal como hemos podido comprobar a lo largo de este trabajo introductorio, las matemáticas requeridas por este ámbito presentan una alta complejidad que crece muy rápidamente a medida que uno se sumerge en él. Por ello, buena parte de los investigadores de esta materia son matemáticos.

Asimismo, fuera del ámbito académico, las empresas que ejercen su actividad en el sector financiero están demandando, en la actualidad, profesionales con una elevada formación específica en matemática aplicada a las finanzas.

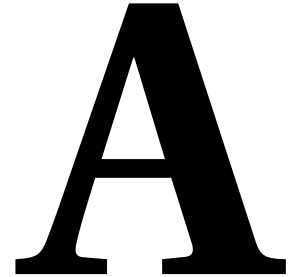
Por ello, la Matemática Financiera se erige como una rama en auge de las matemáticas que, previsiblemente, será notablemente demandada en los próximos años.











# PROGRAMAS EN MATLAB.

## A.1. BSEUROPEA.

```
function [V, error] = BSeuropea(sigma, r, delta, inct, T, N, strike, tipoopcion)
%La funcion BSeuropea aproxima el valor de una opcion europea mediante el
%MEF y compara el resultado con la funcion blsprice de MatLab en el dominio
%(-delta/2,\delta/2).
%N := numero de subintervalos, inct := incremento temporal a casa etapa.

incx = (2*delta)/N;
nodos = [-delta:incx:delta]';

switch tipoopcion
    case 'put'
        payoff = max(strike - exp(nodos), 0);
        phi0 = strike;
        phiedelta = 0;
    case 'call'
        payoff = max(exp(nodos) - strike, 0);
        phi0 = 0;
        phiedelta = strike;
end

%Generamos las distintas matrices dispersas que son requeridas.
M = (incx/6)*(gallery('tridiag', N-1,1,4,1));
K = (1/incx)*(gallery('tridiag', N-1,-1,2,-1));
B = (gallery('tridiag',N-1,-1/2,0,1/2));

aux1 = zeros(N-1, 1);
aux2 = zeros(N-1, 1);
```

```

aux1(1,1) = incx/6;
aux2(N-1,1) = incx/6;
Mext = [aux1,M,aux2];

Abs = ((sigma^2)/2)*K + (((sigma^2)/2)-r)*B +r*M ;

%El vector inicial es el payoff sin los extremos del intervalo considerado.
u = payoff(2:(end -1)) ...
    - ( ((nodos(2:(end-1)) + delta)/(2*delta))*phiedelta ...
    -   (((nodos(2:(end-1)) - delta)/(2*delta))*phi0) ) ;

%Realizamos la factorizacion LU de la matriz del sistema fuera del bucle
%para evitar que MatLab lo haga en cada iteracion.
[L,U] = lu(M+((inct/2)*Abs));

for i = 0:inct:(T-inct)
    Fm = Mext*terminofuente(nodos, delta, phi0, phiedelta, r, sigma, i);
    Fm1 = Mext*terminofuente(nodos, delta, phi0,...
        phiedelta, r, sigma, i+inct);
    b = (M-(inct/2)*Abs)*u + (inct/2)*(Fm+Fm1);

    u = U\(L\b);
end

%Recuperamos el problema original.
V = u + ( ((nodos(2:(end-1)) + delta)/(2*delta))*phiedelta ...
    -   (((nodos(2:(end-1)) - delta)/(2*delta))*phi0*exp(-r*T)) ) ;

nodos = nodos(2:(end-1));
[a, b] = blsprice(exp(nodos), strike, r, T, sigma);

hold on
plot(exp(nodos),V, 'b')
axis([0 2*strike 0 strike])

V1 = V(N/4:3*N/4);
a1 = a(N/4:3*N/4);
b1 = b(N/4:3*N/4);

switch tipoopcion
    case 'put'
        error = max(abs(V1 - b1));
        plot(exp(nodos),b, 'm')
    case 'call'
        error = max(abs(V1 - a1));
        plot(exp(nodos), a, 'm')
end

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

function f = terminofuente(nodos, delta, phi0, phiedelta, r, sigma, tau)

```

```

usubd = phiedelta;
usubmenosd = phi0*exp(-r*tau);
usubmenosdprima = -r*phi0*exp(-r*tau);

f = (usubmenosdprima/(2*delta))*(nodos - delta) ...
    + (r-(sigma^2)/2)*(1/(2*delta))*(usubd-usubmenosd) ...
    - (r/(2*delta))*((nodos+delta)*(usubd) - (nodos-delta)*(usubmenosd));

end

%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%

```

## A.2. BSAMERICANA.

```

function V = BSamericana(sigma, r, delta, inct, T, N, strike)
%La funcion BSeuropea aproxima el valor de una opcion europea mediante el
%MEF y compara el resultado con la funcion blsprice de MatLab.
%N := numero de subintervalos, inct := incremento temporal a casa etapa.

incx = (2*delta)/N;
nodos = [-delta:incx:delta]';

payoff = max(strike - exp(nodos), 0);
phi0 = strike;
phiedelta = 0;

f = -(strike/(2*delta))*(r - (sigma^2)/2) + ...
    ((r*strike)/(2*delta))*(nodos - delta);

%Generamos las distintas matrices dispersas que son requeridas.
M = (incx/6)*(gallery('tridiag', N-1,1,4,1));
K = (1/incx)*(gallery('tridiag', N-1,-1,2,-1));
B = (gallery('tridiag', N-1,-1/2,0,1/2));

aux1 = zeros(N-1, 1);
aux2 = zeros(N-1, 1);
aux1(1,1) = incx/6;
aux2(N-1,1) = incx/6;
Mext = [aux1,M,aux2];

Abs = ((sigma^2)/2)*K + (((sigma^2)/2)-r)*B + r*M ;
F = Mext*f;

%El vector inicial es el payoff sin los extremos del intervalo considerado.
u0 = payoff(2:(end -1)) + (strike/(2*delta))*(nodos(2:(end-1))-delta);
u = u0;

%Realizamos la factorizacion LU de la matriz del sistema fuera del bucle
%para evitar que MatLab lo haga en cada iteracion.
[L,U] = lu(M+((inct/2)*Abs));

```









# BIBLIOGRAFÍA.

- [1] Y. Achdou & O. Pironneau, *Computational Methods for Option Pricing*, Society for Industrial and Applied Mathematics(SIAM), Philadelphia, 2005.
- [2] L. Bachelier, *Théorie de la spéculation.*, Ann. Sci. Éc. Norm. Super., 1900.
- [3] A. Bensoussan & J.L. Lions, *Applications of Variational Inequalities in Stochastic Control.*, Studies in Mathematics and Its Applications, North-Holland, Amsterdam, 1982.
- [4] J. Berk & P. DeMarzo & J. Hardford, *Fundamentos de finanzas corporativas.*, Pearson Educacion S.A., Madrid, 2010.
- [5] F. Black & M. Choles, *The pricing of options and corporate liabilities.*, J. Polit. Econ., 1973.
- [6] P. Brémaud, *Markov chains: Gibbs fields, Montecarlo simulation and queues.*, Springer Science + Business Media, New York, 1999.
- [7] S. C. Brenner & L. R. Scott, *The Mathematical Theory of Finite Element Methods.*, Springer-Verlag, New York, 2002.
- [8] M. Breton & J. de Frutos, *Option Pricing Under GARCH Processes Using PDE Methods.*, Operations Research, 2010.
- [9] B. Daya Reddy, *Introductory Functional Analysis.*, Springer-Verlag, New York, 1998.
- [10] L. C. Evans, *Partial Differential Equations.*, American Mathematical Society, Providence, 1998.
- [11] J. de Frutos, *Implicit-explicit Runge-Kutta methods for financial derivatives pricing models.*, European Journal of Operational Research, 2005.
- [12] J. de Frutos, *Numerical Methods for the Solution os Partial Differential Equations Arising in the Valuation of Financial Derivatives.*, Departamento de Matemática Aplicada, Universidad de Valladolid, Spain, 2004.
- [13] S. Hilber & O. Reichmann & C. Schwab & C. Winter, *Finite Element Methods for Derivative Pricing.*, Springer-Verlag Berlin, Heidelberg, 2013.
- [14] R. Kangro & R. Nicolaidis, *Far Field Boundary Conditions for Black-Scholes Equations.*, Society for Industrial and Applied Mathematics(SIAM), journals, 2000.
- [15] A.M. Matache & P.A. Nitsche & Ch. Schwab, *Wavelet Galerkin pricing of American options on Lévy driven assets.*, Quant, Finance, 2005.

- 
- [16] R. C. Merton, *Theory of rational option pricing.*, Bell J. Econ. Manag. Sci., 1973.
- [17] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications.*, Springer-Verlag Berlin, Heidelberg, 2013.
- [18] S. Salsa, *Partial Differential Equations in Action: From Modelling to Theory.*, Springer-Verlag Italia, Milano, 2008.
- [19] K. Sato, *Lévy processes and infinitely divisible distributions.*, Cambridge University Press, Cambridge, 1999.
- [20] R. Seydel, *Tools for Computational Finance.*, Springer-Verlag Berlin, Heidelberg, 2002.
- [21] E. Süli & D. Mayers, *An Introduction to Numerical Analysis.*, Cambridge University Press, New York, 2003.
- [22] P. Wilmott & S. Howison & J. Dewynne, *Option pricing: mathematical models and computation.*, Oxford Financial Press, Oxford, 1995.
- [23] P. Wilmott & S. Howison & J. Dewynne, *The mathematics of financial derivatives: a student introduction.*, Cambridge University Press, Oxford, 1993.



