



FACULTADE DE MATEMÁTICAS

Traballo Fin de Grao

Fórmulas de cuadratura. Aplicación ao cálculo de coeficientes de Fourier

Aroa Rodríguez Rodríguez

2020/2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

GRAO DE MATEMÁTICAS

Traballo Fin de Grao

**Fórmulas de cuadratura.
Aplicación ao cálculo de coeficientes
de Fourier**

Aroa Rodríguez Rodríguez

2020/2021

UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Traballo proposto

Área de Coñecemento: Matemática Aplicada.
Título: Fórmulas de cuadratura. Aplicación ao cálculo de coeficientes de Fourier.
Breve descrición do contido
<ul style="list-style-type: none">• Concepto de fórmula de cuadratura. Grao de precisión.• Converxencia superalxébrica da fórmula do trapecio composta.• Aplicación ao cálculo dos coeficientes de Fourier.• Transformada de Fourier vía FFT.
Recomendacións
Outras observacións
Titor: Óscar López Pouso.

Índice xeral

Resumo	VIII
Introdución	XI
1. Concepto de fórmula de cuadratura. Grao de precisión	1
1.1. Fórmulas de tipo interpolatorio polinómico	2
1.1.1. Propiedades dos coeficientes de cuadratura cando os nodos están simetricamente distribuídos	5
1.2. Fórmulas de Newton-Cotes	6
1.2.1. Cálculo dos coeficientes	7
1.2.2. O erro das fórmulas de Newton-Cotes	9
1.3. Fórmulas compostas	10
2. Método de Romberg e casos de converxencia superalxébrica da fórmula do trapecio composta	13
2.1. Extraplación de Richardson	13
2.2. Fórmula de Euler-Maclaurin	16
2.2.1. Método de Romberg	17
2.3. Superconverxencia e converxencia superalxébrica da fórmula do trapecio composta	19
3. Aplicación ao cálculo de coeficientes de Fourier	23
3.1. Funcións periódicas	23
3.2. Series de Fourier	24
3.2.1. Series de Fourier para funcións de período 2π	24
3.2.2. Serie de Fourier en notación complexa	27
3.3. Converxencia das series de Fourier	28
3.3.1. Velocidade de converxencia das series de Fourier	31

3.4. Cálculo dos coeficientes de Fourier mediante integración numérica	32
4. Transformada discreta de Fourier vía FFT	35
4.1. Introducción á transformada de Fourier	35
4.2. Transformada discreta de Fourier (DFT)	36
4.2.1. Relación cos coeficientes de Fourier	37
4.3. Transformada rápida de Fourier (FFT)	38
4.4. Aproximación dos coeficientes de Fourier mediante a DFT vía FFT	41
5. Resultados numéricos	45
5.1. Aproximación dos coeficientes de Fourier dunha función discontinua	45
5.2. Aproximación dos coeficientes de Fourier dunha función continua pero non diferenciable	48
5.3. Aproximación dos coeficientes de Fourier dunha función de clase \mathcal{C}^1 pero non \mathcal{C}^2	51
5.4. Aproximación dos coeficientes de Fourier dunha función de clase \mathcal{C}^∞	53
5.5. Conclusións	56
Bibliografía	59

Resumo

Neste traballo presentaremos diferentes fórmulas de cuadratura, centrándonos na fórmula do trapecio composta e con especial interese nun tipo de converxencia que esta presenta, chamada converxencia superalxébrica, onde se mellora a orde 2 de converxencia. Tras iso, introducirmos nas series de Fourier, veremos como obter os coeficientes da citada serie empregando fórmulas de cuadratura e abordaremos algúns resultados en canto á converxencia. Máis tarde, pasaremos a estudar a transformada discreta de Fourier (DFT) e un algoritmo para calculala de maneira rápida e eficaz, chamado transformada rápida de Fourier (FFT). Finalmente, implementaremos un método que fará uso da FFT para calcular os coeficientes de Fourier mediante a fórmula de cuadratura do trapecio composta, rematando con varios casos prácticos nos que daremos fe da eficacia do método.

Abstract

In this work we will present different quadrature rules, focusing on the composite trapezoidal rule and paying special attention to the superalgebraic convergence, a type of convergence where the order 2 of convergence is improved. After this, we will introduce the Fourier series, we will see how to obtain the coefficients of such series using quadrature formulas and we will discuss some results about convergence. Afterwards, we will study the discrete Fourier transform (DFT) and an algorithm to compute it quickly and efficiently, called the fast Fourier transform (FFT). Finally, we will also implement a method that will make use of the FFT to calculate the Fourier coefficients by means of the composite trapezoidal rule, ending with several practical cases in which we will demonstrate the effectiveness of the algorithm.

Introdución

Unha serie de Fourier é unha serie infinita que xorde da necesidade de obter información dunha función periódica e continua a cachos. Estas series son unha ferramenta básica no campo da análise matemática; as súas aplicacións inclúen os ámbitos da enxeñaría eléctrica, a análise vibratoria, o procesamento de sinais e imaxes, a comprensión de datos e a resolución dalgunhas ecuacións diferenciais en derivadas parciais sumamente importantes como a ecuación da calor ou a ecuación de ondas, permitindo obter solucións prácticas na teoría da transmisión da calor, na teoría de placas ... , etc.

A idea subxacente a unha serie de Fourier sería, dada unha función $f(t)$ dependente da variable tempo, descompoñela nunha suma infinita de funcións sinusoidais simples. É dicir expresarase a función $f(t)$ en termos de senos e cosenos,

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nt) + b_n \sin(nt)), \quad (1)$$

onde $\{a_n\}_{n \in \mathbb{N}}$ e $\{b_n\}_{n \in \mathbb{N}}$ son os denominados coeficientes de Fourier. Durante este traballo desenvolveremos o tema empregando funcións periódicas definidas no intervalo $[-\pi, \pi]$, e veremos que estes coeficientes adoptan as expresións

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(nt) dt, \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(nt) dt \quad \text{para } n \in \mathbb{N}. \quad (2)$$

A problemática xorde cando estas integrais non son fáciles de calcular, ben porque non temos unha expresión analítica da función, ou porque simplemente as integrais non chegan a ter unha forma definida. Aquí é onde toma un papel fundamental a integración numérica para poder aproximar coa máxima precisión posible estes coeficientes.

Faremos uso das fórmulas de cuadratura para aproximar estas integrais. Introduciremos os diferentes tipos de fórmulas de cuadratura ca orde de converxencia correspondente, pero o noso interese estará enfocado na regra do trapecio composta, da cal estudaremos as propiedades máis importantes. Tras demostrar que a fórmula do trapecio composta ten orde 2 de converxencia, explicaremos un método que permite mellorar esta orde e daremos

conta de certos casos nos que a propia fórmula, sen aplicación de ningún método adicional, ten por si mesma orde maior que 2. En relación con isto, enunciaremos a fórmula de Euler-Maclaurin, e grazas a ela descubriremos a gran precisión que proporciona a fórmula do trapecio composta cando se integran funcións periódicas regulares nun intervalo de lonxitude o período. Afinando, poderemos chegar a concluír que esa fórmula de cuadratura ten orde de converxencia $2k + 2$ cando

$$f \in \mathcal{C}^{2k+2}([-\pi, \pi]) \quad \text{e} \quad f^{(2l-1)}(-\pi) = f^{(2l-1)}(\pi), \quad l = 1, \dots, k. \quad (3)$$

Isto denominarémolo superconverxencia da fórmula do trapecio composta, e será clave para crear o método que aproxime con precisión as integrais que dan lugar aos coeficientes de Fourier.

Seguidamente, introducirmonos na área da transformada de Fourier, concretamente na transformada discreta de Fourier (DFT), que non será máis que traballar cun conxunto de valores da función. Verase a relación que esta ten co cálculo dos coeficientes de Fourier, resultando que

$$c_k \approx \frac{e^{ik\pi}}{n} \hat{y}_k, \quad (4)$$

sendo \hat{y}_k a transformada discreta de Fourier e c_k os coeficientes de Fourier se expresamos a serie de Fourier (1) en notación complexa, é dicir

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{int}. \quad (5)$$

Analizaremos o custo computacional que supón calcular os coeficientes con este método para motivar o que se chama transformada rápida de Fourier (FFT), un algoritmo que nos permite calcular a DFT cun custo moito menor.

Finalmente, tras explicar como calcular os coeficientes de Fourier mediante a DFT vía FFT usando a fórmula de cuadratura do trapecio composta, recóllense unha serie de resultados numéricos que proban a eficacia do método implementado para aproximar os coeficientes de Fourier e, consecuentemente, a propia función f .

Capítulo 1

Concepto de fórmula de cuadratura. Grao de precisión

Introduciremos as fórmulas de cuadratura e analizaremos o grao de precisión que estas presentan facendo uso da referencia bibliográfica [1]. A integración numérica é unha ferramenta de grande utilidade para obter valores aproximados de integrais definidas

$$I(f) = \int_a^b f(x) dx \quad (1.1)$$

onde $(a, b) \subset \mathbb{R}$ é un intervalo acotado e non baleiro.

Esta necesidade xorde por distintos feitos: integrais que non poden calcularse explicitamente, ben porque o integrando non ten unha primitiva expresable analiticamente, ou tena pero o seu cálculo é dificultoso ou impreciso debido a erros de redondeo, ou tamén porque dito integrando só se coñece dunha forma discreta (por exemplo valores experimentais). Tendo en conta que unha integral é o límite de sumas cuxo número de sumandos tende a infinito, consideramos natural a idea de que a súa aproximación sexa unha suma finita de mostras do integrando, $f(x_i)$, ponderadas con pesos A_i .

Definición 1.1. Consideremos n puntos distintos $x_1 < \dots < x_n$ nos que f está definida e é avaliable. Unha fórmula do xeito

$$Q_n(f) = \sum_{i=1}^n A_i f(x_i) \quad (1.2)$$

denomínase **fórmula de cuadratura**. Os puntos x_i son os *nodos* ou *puntos de cuadratura*, e os A_i son os *pesos* ou *coeficientes de cuadratura*.

Como mencionamos, a fórmula de cuadratura $Q_n(f)$ ofrécenos unha aproximación da integral $I(f)$. Así, parece natural expresar o *erro de cuadratura* ou *de truncamento*, denotado por $E_n(f)$, como aquela cantidade que cumpre

$$I(f) = Q_n(f) + E_n(f). \quad (1.3)$$

O problema principal que se plantexa é como tomar os nodos e os coeficientes de forma axeitada para que $Q_n(f)$ sexa unha boa aproximación de $I_n(f)$, dito doutro xeito: como conseguimos o menor erro, $E_n(f)$, posible?

Debemos ter en conta que a maioría das regras de cuadratura baséanse na interpolación polinómica, substituíndo o integrando polo polinomio que interpola a $f(x)$ nos nodos. Polo tanto, a precisión da aproximación esta intimamente ligada á capacidade da regra de integrar exactamente polinomios. De aí xorde o concepto de *grao de precisión*.

Definición 1.2. Diremos que unha fórmula de cuadratura ten **grao de precisión m** ($m \in \mathbb{N} \cup \{0\}$) se:

$$E_n(x^k) = 0 \quad \text{para } k = 0, 1, \dots, m, \quad (1.4)$$

$$E_n(x^{m+1}) \neq 0. \quad (1.5)$$

O erro $E_n(f)$ é unha función lineal de f , isto é: $E_n(f+g) = E_n(f) + E_n(g)$ e $E_n(\lambda f) = \lambda E_n(f)$.

Extraemos ademais da definición que, se unha fórmula de cuadratura ten grao de precisión m , entón é exacta en \mathbb{P}_m , isto é, integra exactamente os polinomios de grao $\leq m$. Así mesmo, unha fórmula de cuadratura ten grao de precisión m se, e só se, m é o maior enteiro para o cal a fórmula é exacta en \mathbb{P}_m .

Observación 1.3. Se se cumpre a condición (1.4), é inmediato que o grao de precisión será $\geq m$. Neste caso diremos que a fórmula ten **polo menos grao de precisión m** . É claro, pola linealidade do erro, que unha fórmula ten polo menos grao de precisión m se, e só se, é exacta en \mathbb{P}_m .

1.1. Fórmulas de tipo interpolatorio polinómico

Sexan n puntos distintos

$$x_1 < \dots < x_n \quad (1.6)$$

nos que f está definida e é avaliable, e sexa p o *polinomio de interpolación* de f relativo a eses puntos. Entón p é o único polinomio de grao $\leq n-1$ tal que

$$p(x_i) = f(x_i) \quad \text{para } i = 1, \dots, n. \quad (1.7)$$

Definición 1.4. Chamamos **fórmulas de tipo interpolatorio polinómico** (*i. p.*) ás fórmulas obtidas ao aproximar $\int_a^b f(x) dx$ por $\int_a^b p(x) dx$, isto é:

$$\int_a^b f(x) dx \approx \int_a^b p(x) dx. \quad (1.8)$$

É coñecido o seguinte resultado.

Teorema 1.5 (**caracterización das fórmulas de tipo i. p.**). *Unha fórmula de cuadratura de n nodos, $\sum_{i=1}^n A_i f(x_i)$, é de tipo i. p. se, e só se, ten polo menos grao de precisión $n - 1$.*

Observación 1.6. En virtude do teorema anterior, sabemos que toda fórmula de cuadratura de tipo i. p. de n nodos ten, polo menos, grao de precisión $n - 1$. Ademais, debemos salienta que existen fórmulas de n nodos que teñen grao de precisión maior que $n - 1$; estas fórmulas son denominadas **superexactas**.

Nótese que deste teorema podemos extraer un método para o cálculo dos coeficientes A_i da nosa fórmula de tipo i. p. Será o chamado *método dos coeficientes indeterminados*.

Método dos coeficientes indeterminados

Sexan os nodos $x_1 < \dots < x_n$. Lembremos que unha fórmula de cuadratura ten polo menos grao de precisión $n - 1$ se integra exactamente os polinomios de \mathbb{P}_{n-1} ou, equivalentemente, os monomios x^k onde $k = 0, \dots, n - 1$. Agora, teñamos en conta que

$$\int_a^b x^k dx = \frac{b^{k+1} - a^{k+1}}{k + 1}. \quad (1.9)$$

Para que a nosa fórmula de cuadratura (1.2) sexa de tipo i. p. é necesario e suficiente (polo Teorema 1.5) que teña polo menos grao de precisión $n - 1$. Así pois, os coeficientes serán solución do sistema lineal

$$\sum_{i=1}^n A_i x_i^k = \frac{b^{k+1} - a^{k+1}}{k + 1} \quad \text{con } k = 0, \dots, n - 1, \quad (1.10)$$

isto é,

$$\left\{ \begin{array}{l} A_1 + A_2 + \dots + A_n = b - a, \\ A_1 x_1 + A_2 x_2 + \dots + A_n x_n = \frac{b^2 - a^2}{2}, \\ \vdots \\ A_1 x_1^{n-1} + A_2 x_2^{n-1} + \dots + A_n x_n^{n-1} = \frac{b^n - a^n}{n}. \end{array} \right. \quad (1.11)$$

Por ser os nodos n puntos distintos, este sistema sempre ten solución única. En poucas palabras, cando hai que determinar n pesos, plantexamos un sistema impondo a integración exacta dos n monomios de grao 0 ata grao $n - 1$, obtendo así un sistema lineal $n \times n$. Así e todo, a matriz asociada ao sistema (1.11) adoita volverse mal condicionada a medida que o número de nodos aumenta, o cal provoca unha mala precisión no método dos coeficientes indeterminados. En realidade, a problemática non son as aproximacións dos coeficientes, senón o tipo de fórmulas de cuadratura que vimos ata agora, aquelas que se basean na integración do polinomio de interpolación en todo o intervalo (*fórmulas simples*).

As mellores aproximacións de integrais definidas mediante métodos de cuadratura acádanse dende as fórmulas de baixo grao de precisión, isto é: de poucos nodos. Así, as fórmulas que teñen un verdadeiro valor práctico son aquelas que empregan repetidamente fórmulas de baixo grao de precisión en pequenas subdivisións do intervalo, ás que chamaremos *fórmulas compostas*.

Exemplo 1.7 (Fórmula do trapecio). A fórmula do trapecio é unha fórmula de cuadratura de tipo i. p. de dous nodos, $n = 2$, empregando así unicamente os valores extremos do intervalo (a, b) , $(x_1 = a, x_2 = b)$. Fixándonos na fórmula (1.2), será do xeito:

$$\int_a^b f(x) dx \approx A_1 f(a) + A_2 f(b). \quad (1.12)$$

Imos calcular explicitamente os coeficientes.

▷ *Método dos coeficientes indeterminados.*

En virtude do Teorema 1.5 sabemos que a fórmula (1.12), por ser de tipo i. p., ten polo menos grao de precisión $n - 1 = 1$. Polo tanto, debe integrar exactamente os polinomios de grao 0 e de grao 1. Impondo esta exactitude obtemos un sistema do tipo (1.11):

$$\begin{cases} A_1 + A_2 = b - a, \\ A_1 a + A_2 b = \frac{b^2 - a^2}{2}. \end{cases} \quad (1.13)$$

Onde os coeficientes A_1 e A_2 son solución do sistema. Resolvéndoo, vemos que

$$A_1 = A_2 = \frac{b - a}{2}. \quad (1.14)$$

Polo tanto, volvendo a (1.12), obtense a fórmula

$$\int_a^b f(x) dx = \frac{b - a}{2} [f(a) + f(b)]. \quad (1.15)$$

Chegando así ao mesmo resultado, a **fórmula do trapecio**.

Observación 1.8. Esta fórmula ten, por construción, polo menos grao de precisión 1. É máis, pódese comprobar facilmente que ten exactamente grao de precisión 1 (ao non ser exacta para x^2). Comprobarémolo de novo máis adiante, a partir da expresión do erro.

Exemplo 1.9. Vexamos máis fórmulas simples de cuadratura.

▷ **Fórmula do punto medio ou de Poncelet.**

É unha fórmula de tipo i. p. de $n = 1$ nodo, tomando o punto medio do interalo (a, b) . Así:

$$\int_a^b f(x) dx \approx A_1 f\left(\frac{a+b}{2}\right). \quad (1.16)$$

E do mesmo xeito que no exemplo anterior, obtemos que $A_1 = b - a$.

▷ **Fórmula de Simpson.**

É unha fórmula de tipo i. p. de $n = 3$ nodos relativos novamente ao intervalo (a, b) , onde $x_1 = a, x_2 = \frac{a+b}{2}, x_3 = b$. Así pois, temos que

$$\int_a^b f(x) dx \approx \frac{b-a}{6} \left[f(a) + 4f\left(\frac{a+b}{2}\right) + f(b) \right]. \quad (1.17)$$

Observación 1.10. Por construción, a fórmula do punto medio ten polo menos de grao de precisión 0; máis aínda, pódese comprobar que ten grao de precisión 1. Asemade, a fórmula de Simpson ten polo menos grao de precisión 2, pero tamén se pode comprobar que ten grao de precisión 3. Son ambas, polo tanto, exemplos de fórmulas *superexactas*.

1.1.1. Propiedades dos coeficientes de cuadratura cando os nodos están simetricamente distribuídos

A partir de agora consideraremos os nodos de xeito que estes estean simetricamente distribuídos respecto ao punto medio do intervalo.

Esta escolla particular da distribución dos nodos ten unha gran valía práctica á hora de calcular os coeficientes dun xeito máis doado. Tocante a esta cuestión, enunciámos a seguinte proposición.

Proposición 1.11. *Supoñamos que os nodos dunha fórmula de cuadratura de tipo i. p., $\sum_{i=1}^n A_i f(x_i)$, están simetricamente distribuídos respecto ao punto medio do intervalo de integración, isto é,*

$$x_i - \frac{a+b}{2} = \frac{a+b}{2} - x_{n-(i-1)} \quad \text{para } i = 1, \dots, n. \quad (1.18)$$

Entón

$$A_i = A_{n-(i-1)} \quad \text{para } i = 1, \dots, n. \quad (1.19)$$

1.2. Fórmulas de Newton-Cotes

Este escrito comezou vendo que dado un número calquera de nodos, n , podemos construír unha fórmula de cuadratura de n puntos e grao de precisión polo menos $n - 1$. Vimos unha pequena pero valiosa propiedade acerca das fórmulas cuxos nodos están simetricamente distribuídos respecto ao punto medio do intervalo de integración $[a, b]$; un subconxunto destas regras de cuadratura reciben o nome de *fórmulas de Newton-Cotes*.

Definición 1.12. Chámanse **fórmulas de Newton-Cotes** ás fórmulas de cuadratura de tipo i. p. que se obteñen cando os nodos x_1, \dots, x_n están equiespaciados mediante un paso h , de xeito que $x_i = x_1 + (i - 1)h$ para $i = 1, \dots, n$ e ademais $a = x_1 - kh$ e $b = x_n + kh$ onde, ou ben $k = 0$ nas dúas igualdades, ou ben $k = 1$ nas dúas igualdades.

Unha fórmula de Newton-Cotes pode escribirse como

$$\int_a^b f(x) dx = \int_{x_1 - kh}^{x_n + kh} f(x) dx = \sum_{i=1}^n A_i f(x_1 + (i - 1)h) + E(f), \quad (1.20)$$

onde

$$A_i = \int_a^b l_i(x) dx = \int_{x_1 - kh}^{x_n + kh} l_i(x) dx. \quad (1.21)$$

Observación 1.13. Teñamos en conta que estas fórmulas están nas hipóteses da Proposición 1.11, co cal

$$A_i = A_{n-(i-1)} \quad \text{para } i = 1, \dots, n. \quad (1.22)$$

En función dos dous valores posibles de k , distinguimos dous tipos de fórmulas de Newton-Cotes.

Definición 1.14. Chamamos fórmulas de **Newton-Cotes de tipo pechado** a aquelas nas que $k = 0$, é dicir, se $x_1 = a$ e $x_n = b$. Neste caso,

$$x_i = x_1 + (i - 1)h = a + (i - 1)h \quad \text{para } i = 1, \dots, n, \quad (1.23)$$

$$b - a = x_n - x_1 = (n - 1)h, \quad (1.24)$$

$$h = \frac{b - a}{n - 1}. \quad (1.25)$$

Definición 1.15. Chamamos fórmulas de **Newton-Cotes de tipo aberto** a aquelas nas que $k = 1$, é dicir, se $x_1 = a + h$ e $x_n = b - h$. Neste caso os extremos do intervalo de integración non son nodos da fórmula. E temos que

$$x_i = x_1 + (i - 1)h = a + ih \quad \text{para } i = 1, \dots, n, \quad (1.26)$$

$$b - a = x_n + h - (x_1 - h) = x_n - x_1 + 2h = (n - 1 + 2)h = (n + 1)h, \quad (1.27)$$

$$h = \frac{b - a}{n + 1}. \quad (1.28)$$

En ambos casos as fórmulas quedan determinadas por

$$x_i = a + (i - 1 + k)h \quad \text{para } i = 1, \dots, n, \quad (1.29)$$

$$h = \frac{b - a}{n - 1 + 2k}. \quad (1.30)$$

1.2.1. Cálculo dos coeficientes

Xa vimos na sección anterior un método para o cálculo dos coeficientes de cuadratura, o *método dos coeficientes indeterminados*. Cando se traballa con este tipo de fórmulas, a simetría dos coeficientes ($A_i = A_{n-(i-1)}$) resulta especialmente práctica, pois as ecuacións do sistema que resulten de impoñer a exactitude das fórmulas para as funcións x^m con m impar resultan redundantes e polo tanto non se teñen que ter en conta. A continuación, ilustremos a obtención das fórmulas máis simples (as que empregan 1, 2 e 3 nodos).

Exemplo 1.16 (Fórmula de Newton-Cotes aberta de 1 nodo). Estamos no caso de ter

$$\int_a^b f(x) dx \approx A_1 f(x_1), \quad (1.31)$$

onde $x_1 = \frac{a+b}{2}$ é o punto medio do intervalo de integración, e o paso é obviamente $h = b - a$. Por construción sabemos que ten polo menos grao de precisión 0, impoñemos logo exactitude na integración de constantes e obtemos a ecuación

$$A_1 = b - a. \quad (1.32)$$

Polo tanto

$$\int_a^b f(x) dx \approx (b - a) f\left(\frac{a + b}{2}\right) = hf(x_1). \quad (1.33)$$

É a chamada **fórmula do punto medio** ou **fórmula de Poncelet**.

Exemplo 1.17 (Fórmula de Newton-Cotes pechada de 2 nodos). Estamos no caso de ter

$$\int_a^b f(x) dx \approx \sum_{i=1}^2 A_i f(x_i), \quad (1.34)$$

onde $x_1 = a$, $x_2 = b$ e o paso é $h = b - a$.

Como os coeficientes son invariantes por traslacións, podemos supoñer, por comodidade, que

$$a = x_1 = 0, \quad b = x_2 = h. \quad (1.35)$$

Ademais, pola simetría dos coeficientes (Proposición 1.11) sabemos que $A_1 = A_2$. Agora, sabemos que esta fórmula terá polo menos grao de precisión 1, en particular grao de precisión 0, logo cumpre a ecuación

$$A_1 + A_2 = b - a, \quad (1.36)$$

que por simetría é $2A_1 = b - a$. Co cal

$$A_1 = A_2 = \frac{b - a}{2}. \quad (1.37)$$

Obtendo

$$\int_a^b f(x) dx \approx \frac{b - a}{2}(f(a) + f(b)) = \frac{h}{2}(f(x_1) + f(x_2)). \quad (1.38)$$

É a chamada **fórmula do trapecio**.

Exemplo 1.18 (Fórmula de Newton-Cotes pechada de 3 nodos). Estamos no caso de ter

$$\int_a^b f(x) dx \approx \sum_{i=1}^3 A_i f(x_i), \quad (1.39)$$

onde $x_1 = a$, $x_2 = \frac{a+b}{2}$, $x_3 = b$ e o paso é $h = \frac{b-a}{2}$.

Novamente, por simetría e invariancia por traslacións dos coeficientes temos que

$$A_1 = A_3, \quad a = x_1 = 0, \quad x_2 = h, \quad b = x_3 = 2h. \quad (1.40)$$

Agora, sabemos que esta fórmula ten polo menos grao de precisión 2, polo tanto os coeficientes A_1, A_2 obtéñense de resolver o sistema

$$\begin{cases} 2A_1 + A_2 = b - a, \\ A_2 h + A_1 2h = \frac{b^2 - a^2}{2}, \\ A_1 4h^2 + A_2 h^2 = \frac{b^3 - a^3}{3}. \end{cases} \quad (1.41)$$

Como mencionamos ao comezo, grazas á simetría dos coeficientes a ecuación que resulta de impoñer exactitude en x^m con m impar resulta redundante e non aporta información. No noso caso é a segunda ecuación a que non imos ter en conta (resultante de impor exactitude para o monomio de grao 1). Co cal, resolvendo temos

$$A_1 = A_3 = \frac{b - a}{6}, \quad A_2 = \frac{2(b - a)}{3}. \quad (1.42)$$

obtendo

$$\int_a^b f(x) dx \approx \frac{b - a}{6}(f(a) + 4f\left(\frac{a + b}{2}\right) + f(b)) = \frac{h}{3}(f(x_1) + 4f(x_2) + f(x_3)). \quad (1.43)$$

1.2.2. O erro das fórmulas de Newton-Cotes

Anteriormente, xa calculamos o erro da fórmula do trapecio (de 2 puntos) mediante o Teorema 1.5, pois non ten ningún nodo no intervalo (a, b) . Pola contra, o resto das fórmulas de Newton-Cotes teñen todas algún dos seus nodos neste intervalo, co cal non podemos empregar con elas ese resultado. Por iso enunciaremos uns teoremas propios para este tipo de regras de cuadratura.

Teorema 1.19 (representación do erro para as fórmulas pechadas de Newton-Cotes). *Para as fórmulas de Newton-Cotes pechadas cúmprese que:*

1. *Se n é impar e $f \in C^{n+1}([a, b])$, existe $\xi \in (a, b)$ tal que*

$$E(f) = \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_0^{n-1} t^2(t-1) \cdots (t-(n-1)) dt. \quad (1.44)$$

2. *Se n é par e $f \in C^n([a, b])$, existe $\xi \in (a, b)$ tal que*

$$E(f) = \frac{h^{n+1} f^{(n)}(\xi)}{n!} \int_0^{n-1} t(t-1) \cdots (t-(n-1)) dt. \quad (1.45)$$

Teorema 1.20 (representación do erro para as fórmulas abertas de Newton-Cotes). *Para as fórmulas de Newton-Cotes abertas cúmprese que:*

1. *Se n é impar e $f \in C^{n+1}([a, b])$, existe $\xi \in (a, b)$ tal que*

$$E(f) = \frac{h^{n+2} f^{(n+1)}(\xi)}{(n+1)!} \int_{-1}^n t^2(t-1) \cdots (t-(n-1)) dt. \quad (1.46)$$

2. *Se n é par e $f \in C^n([a, b])$, existe $\xi \in (a, b)$ tal que*

$$E(f) = \frac{h^{n+1} f^{(n)}(\xi)}{n!} \int_{-1}^n t(t-1) \cdots (t-(n-1)) dt. \quad (1.47)$$

Observación 1.21. Destes teoremas dedúcese que cando o número de puntos é par o grao de precisión da fórmula é n , mentres que cando o número de puntos é impar o grao de precisión da fórmula é $n+1$ (son superexactas).

Un concepto que cobrará relevancia máis adiante é o de *orde de precisión* ou *orde de converxencia*, ese é o motivo polo cal os erros dos teoremas anteriores están expresados en función dunha potencia do paso h . Vexamos como aplicar o teorema correspondente para calcular o erro da fórmula do trapecio.

Exemplo 1.22 (Erro da fórmula do trapecio). Consideremos a fórmula do trapecio

$$\int_a^b f(x) dx = \frac{b-a}{2}(f(a) + f(b)) + E(f) \quad (1.48)$$

onde $n = 2$, $h = b - a$, $x_1 = a$ e $x_2 = b$.

É unha fórmula de Newton-Cotes pechada con n par, entón se $f \in \mathcal{C}^2([a, b])$, por (1.45), existe $\xi \in (a, b)$ de xeito que

$$E(f) = \frac{h^3 f''(\xi)}{2!} \int_0^1 t(t-1) dt. \quad (1.49)$$

Calculando a integral obtemos o erro da regra do trapecio:

$$E(f) = \frac{h^3 f''(\xi)}{2} \frac{-1}{6} = -\frac{1}{12} h^3 f''(\xi). \quad (1.50)$$

1.3. Fórmulas compostas

Sabemos por apartados anteriores que a precisión obtida mediante unha regra de cuadratura no cálculo dunha integral definida mellora sensiblemente coa diminución da lonxitude do intervalo de integración. Dado que o intervalo de integración é un dato fixado que non podemos modificar, a idea central para mellorar a precisión dunha fórmula de cuadratura (incluso se a función do integrando oscila varias veces no intervalo) é dividir o intervalo de integración en *subintervalos* (normalmente da mesma lonxitude), aplicar a regra en cada un destes novos intervalos e sumar os resultados obtidos. Este proceso dá lugar ás denominadas **fórmulas compostas**. Para comezar, veremos cal é a idea xeral para a construción destas fórmulas.

Supoñamos querer aproximar $\int_a^b f(x) dx$, onde (a, b) é un intervalo acotado. Imos dividir o intervalo $[a, b]$ en m subintervalos

$$[y_j, y_{j+1}] \quad j = 1, \dots, m \quad (1.51)$$

da mesma lonxitude $H = \frac{b-a}{m}$, onde $y_j = a + (j-1)H$ con $j = 1, \dots, m+1$ (así $y_1 = a$ e $y_{m+1} = b$). Entón temos que

$$\int_a^b f(x) dx = \sum_{j=1}^m \int_{y_j}^{y_{j+1}} f(x) dx. \quad (1.52)$$

Agora, en cada subintervalo $[y_j, y_{j+1}]$ aproximamos $\int_{y_j}^{y_{j+1}} f(x) dx$ cunha fórmula de cuadratura simple, neste caso empregaremos unha de Newton-Cotes pechada de n puntos, é dicir:

$$\int_{y_j}^{y_{j+1}} f(x) dx = \sum_{i=1}^n A_i^{(j)} f(y_j + (i-1)h) + E^{(j)}(f), \quad (1.53)$$

onde agora o paso é $h = \frac{H}{n-1}$. (Hai que ter en conta que os superíndices (j) son simplemente notación para indicar que son os termos asociados ao intervalo $[y_j, y_{j+1}]$; de feito, algunhas veces non aparecerán por comodidade.)

Unha vez chegados a este punto, en cada subintervalo empregamos as fórmulas simples vistas ata agora, e polo tanto son aplicables tódolos resultados de capítulos anteriores.

Exemplo 1.23 (Fórmula do trapecio composta). Calculemos a integral $\int_a^b f(x) dx$ mediante a fórmula do trapecio composta.

Dividimos o intervalo $[a, b]$ en m subintervalos de lonxitude $H = \frac{b-a}{m}$, obtendo entón $m+1$ nodos y_1, \dots, y_{m+1} .

Apliquemos a fórmula do trapecio simple a cada subintervalo, é dicir, á integral

$$\int_{y_j}^{y_{j+1}} f(x) dx \quad (1.54)$$

onde $j = 1, \dots, m$. Teriamos entón que

$$\begin{aligned} \text{Para } [y_1, y_2] = [a, y_2] &\longrightarrow \int_a^{y_2} f(x) dx \approx \frac{h}{2}(f(a) + f(y_2)), \\ \text{Para } [y_2, y_3] &\longrightarrow \int_{y_2}^{y_3} f(x) dx \approx \frac{h}{2}(f(y_2) + f(y_3)), \\ &\vdots \\ \text{Para } [y_m, y_{m+1}] = [y_m, b] &\longrightarrow \int_{y_m}^b f(x) dx \approx \frac{h}{2}(f(y_m) + f(b)), \end{aligned}$$

onde h é claramente $h = H$, posto que a regra do trapecio emprega só dous puntos. É dicir: o paso é a lonxitude do intervalo.

Agora, sumando tódalas aproximacións temos que:

$$\int_a^b f(x) dx \approx \frac{h}{2}[f(a) + 2f(y_2) + \dots + 2f(y_m) + f(b)], \quad (1.55)$$

e agrupando termos (nótese que os valores da función evaluada nos nodos intermedios do intervalo aparecen dúas veces):

$$\int_a^b f(x) dx \approx \frac{h}{2}[f(a) + f(b)] + h \sum_{j=2}^m f(y_j). \quad (1.56)$$

En canto ao erro, é evidente que este será a suma de tódolos erros cometidos en cada subintervalo ao aplicar a fórmula simple, e, como xa sabemos, o erro en cada intervalo será, se $f \in \mathcal{C}^2([y_j, y_{j+1}])$,

$$E^j(f) = -\frac{h^3}{12}f''(\xi), \quad (1.57)$$

para certo $\xi \in (y_j, y_{j+1})$. Este é o erro cometido en cada subintervalo, polo tanto en $[a, b]$ o erro será

$$\sum_{j=1}^m E^j(f) = -m \frac{h^3}{12} f''(\xi). \quad (1.58)$$

Así, tendo en conta que $h = H = \frac{b-a}{m}$ temos que $m = \frac{b-a}{h}$, o que substituíndo na expresión anterior proporciónanos o *erro da fórmula do trapecio composta*:

$$E(f) = -\frac{b-a}{12} h^2 f''(\xi). \quad (1.59)$$

O relevante da expresión anterior é a orde de converxencia ou precisión. Nótese que se se escollemos K unha constante axeitada, é claro que:

$$E(f) \leq Kh^2, \quad (1.60)$$

o cal significará que a fórmula do trapecio ten orde de converxencia 2, polo que se poderá denotar o resultado simplemente mediante

$$E(f) = \mathcal{O}(h^2), \quad (1.61)$$

dícese que a orde é un “*O grande*” de h^2 .

Disto extraemos dúas conclusións importantes:

▷ A orde de converxencia da fórmula composta é a magnitude que a efectos prácticos ten sentido manexar, xa que a lonxitude do intervalo de integración $(b-a)$ é unha magnitude fixa na práctica. Polo tanto, se reducimos á metade (factor 2) o paso h , dividindo así cada subintervalo pola metade, a cota do erro diminúe nun factor de $2^2 = 4$. É dicir, o erro queda dividido por 4 simplemente dividindo o paso á metade.

Observación 1.24. En xeral, se temos unha fórmula composta cun erro de orde $\mathcal{O}(h^p)$ e multiplicamos o número de intervalos por un número natural q , o que estamos a facer é diminuír o paso h nun factor q . É dicir, estamos considerando para o mesmo intervalo un novo paso, \hat{h} , tal que $\hat{h} = \frac{h}{q}$, e polo tanto o erro queda dividido por q^p .

Capítulo 2

Método de Romberg e casos de converxencia superalxébrica da fórmula do trapecio composta

Neste capítulo desenvolveremos diferentes métodos nos que se mellora a orde 2 de converxencia na fórmula do trapecio composta apoiándonos en [1]. Para comezar, explicaremos algúns métodos que nos serán de utilidade para entender o obxectivo do texto.

2.1. Extrapolación de Richardson

A extrapolación de Richardson é unha técnica recursiva que combina dúas aproximacións para obter unha terceira, mellorando así a orde de converxencia.

Sexa α_0 unha cantidade descoñecida que aproximamos por unha cantidade calculable $A(y)$, con $y > 0$. Supoñamos que, para un certo $k \in \mathbb{N}$, o valor $A(y)$ admite un *desenvolvemento asintótico* de A en potencias de y , isto é, admite o desenvolvemento

$$A(y) = \alpha_0 + \alpha_1 y + \alpha_2 y^2 + \dots + \alpha_k y^k + \mathcal{O}(y^{k+1}), \quad (2.1)$$

onde os coeficientes $\alpha_1, \dots, \alpha_k$ son independentes de y . Nótese que entón

$$\lim_{y \rightarrow 0} A(y) = \alpha_0. \quad (2.2)$$

Observación 2.1. Se $A(y) - \alpha_0 = \mathcal{O}(y^p)$ diremos que $A(y)$ é unha aproximación de α_0 de orde p con respecto ao parámetro y .

Ademais

$$A(y) - \alpha_0 = \alpha_1 y + \mathcal{O}(y^2). \quad (2.3)$$

Polo tanto, $A(y)$ é unha aproximación de α_0 de, polo menos, orde 1 con respecto ao parámetro y , e é exactamente de orde 1 se $\alpha_1 \neq 0$. A extrapolación de Richardson permitiranos obter aproximacións de α_0 de orde maior que 1.

Descrición da extrapolación de Richardson

Descríbimos o método e forma xenérica.

Partimos de

$$\phi_0(y) = A(y). \quad (2.4)$$

Agora construímos, para $n = 0, \dots, k-1$,

$$\phi_{n+1}(y) = \frac{\phi_n(ry) - r^{n+1}\phi_n(y)}{1 - r^{n+1}}. \quad (2.5)$$

Temos o seguinte resultado.

Proposición 2.2. *Para $n = 0, \dots, k$, $\phi_n(y)$ ten un desenvolvemento da forma*

$$\phi_n(y) = \alpha_0 + \beta_{n,n+1}y^{n+1} + \dots + \beta_{n,k}y^k + \mathcal{O}(y^{k+1}) \quad (2.6)$$

Demostración. Probaremos o resultado por indución.

Para $n = 0$, (2.6) é válida con

$$\beta_{0,l} = \alpha_l \quad l = 1, \dots, k. \quad (2.7)$$

Supoñamos certa (2.6) para un certo enteiro n , con $0 \leq n \leq k-1$. Entón:

$$\phi_n(ry) = \alpha_0 + \beta_{n,n+1}r^{n+1}y^{n+1} + \beta_{n,n+2}r^{n+2}y^{n+2} + \dots + \beta_{n,k}r^k y^k + \mathcal{O}(y^{k+1}) \quad (2.8)$$

e, por (2.5),

$$\phi_{n+1}(y) = \alpha_0 + \beta_{n,n+2} \frac{r^{n+2} - r^{n+1}}{1 - r^{n+1}} y^{n+2} + \dots + \beta_{n,k} \frac{r^k - r^{n+1}}{1 - r^{n+1}} y^k + \mathcal{O}(y^{k+1}), \quad (2.9)$$

isto é,

$$\phi_{n+1} = \alpha_0 + \beta_{n+1,n+2}y^{n+2} + \dots + \beta_{n+1,k}y^k + \mathcal{O}(y^{k+1}), \quad (2.10)$$

onde

$$\beta_{n+1,l} = \beta_{n,l} \frac{r^l - r^{n+1}}{1 - r^{n+1}} \quad l = n+2, \dots, k. \quad (2.11)$$

Con isto queda demostrado, pois a expresión obtida é un desenvolvemento do mesmo tipo que (2.6), con n substituído por $n+1$. \square

Para $n = k$ temos

$$\phi_k(y) = \alpha_0 + \mathcal{O}(y^{k+1}) \quad (2.12)$$

e xa non se podería continuar o proceso.

Observación 2.3. A ecuación (2.6) en particular implica que, para calquera n , $0 \leq n \leq k$,

$$\phi_n(y) = \alpha_0 + \mathcal{O}(y^{n+1}). \quad (2.13)$$

Implementación do método

Na práctica, tomamos un valor $y_0 > 0$ e queremos calcular a aproximación $\phi_M(y_0)$ para un certo M fixado de antemán onde $1 \leq M \leq k$. O proceso descríbese do seguinte xeito:

1. Calculamos $\phi_0(y_0) = A(y_0)$.
2. Calculamos $\phi_0(ry_0) = A(ry_0)$, e entón

$$\phi_1(y_0) = \frac{\phi_0(ry_0) - r\phi_0(y_0)}{1 - r}. \quad (2.14)$$

Se $M = 1$, o proceso remata aquí.

3. Calculamos, por orde, primeiro $\phi_0(r^2y_0) = A(r^2y_0)$ e despois

$$\phi_1(ry_0) = \frac{\phi_0(r^2y_0) - r\phi_0(ry_0)}{1 - r} \quad (2.15)$$

e

$$\phi_2(y_0) = \frac{\phi_1(ry_0) - r^2\phi_1(y_0)}{1 - r^2}. \quad (2.16)$$

Se $M = 2$, o proceso remata aquí.

4. Calculamos, por orde, primeiro $\phi_0(r^3y_0) = A(r^3y_0)$ e despois

$$\phi_1(r^2y_0) = \frac{\phi_0(r^3y_0) - r\phi_0(r^2y_0)}{1 - r}, \quad \phi_2(ry_0) = \frac{\phi_1(r^2y_0) - r^2\phi_1(ry_0)}{1 - r^2} \quad (2.17)$$

e

$$\phi_3(y_0) = \frac{\phi_2(ry_0) - r^3\phi_2(y_0)}{1 - r^3}. \quad (2.18)$$

Se $M = 3$, o proceso remata aquí.

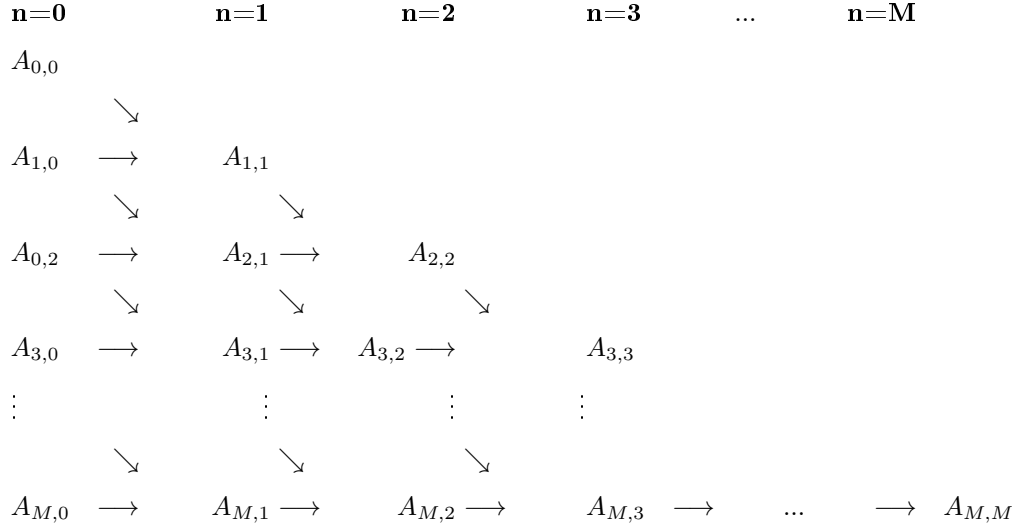
Se $M > 3$, o proceso continúa de xeito análogo ata chegar a $\phi_M(y_0)$.

Observación 2.4. Todo o proceso está baseado na aplicación recursiva da fórmula (2.5).

Definición 2.5. Para $n = 0, \dots, M$ e $m = n, n + 1, \dots, M$, definimos

$$A_{m,n} = \phi_n(r^{m-n}y_0). \quad (2.19)$$

Con esta nova notación, o proceso descrito anteriormente pódese esquematizar na seguinte táboa:



Cada elemento da táboa, agás os da primeira columna ($n = 0$), calcúlase a partir do elemento situado inmediatamente á esquerda e o situado xusto por riba deste último, empregando as fórmulas (2.19) e (2.5).

Obténse así que, para todo $n = 0, \dots, M - 1$ e para todo $m = n + 1, \dots, M$,

$$A_{m,n+1} = \phi_{n+1}(r^{m-(n+1)}y_0) = \frac{\phi_n(r^{m-n}y_0) - r^{n+1}\phi_n(r^{m-n-1}y_0)}{1 - r^{n+1}} = \frac{A_{m,n} - r^{n+1}A_{m-1,n}}{1 - r^{n+1}}. \quad (2.20)$$

Na práctica, a táboa pode xerarse por columnas (calculando primeiro os $A_{m,0}$ e implementando despois un bucle externo en n e un interno en m) ou por filas (cun bucle externo en m e un interno en n).

2.2. Fórmula de Euler-Maclaurin

A fórmula de Euler-Maclaurin relaciona integrais con series, e se pode empregar para aproximar integrais por sumas finitas, de aí a súa utilidade. Vexámola.

Sexa (a, b) un intervalo acotado e non baleiro, e sexa

$$x_i = a + (i - 1)h, \quad i = 1, \dots, n, \quad h = \frac{b - a}{n - 1}, \quad (2.21)$$

unha partición uniforme do intervalo $[a, b]$. Chamemos $T(h)$ á aproximación de $\int_a^b f(x) dx$ obtida ao aplicar a regra do trapecio composta con paso h , isto é:

$$T(h) = \frac{h}{2}[f(a) + f(b)] + h \sum_{i=2}^{n-1} f(x_i). \quad (2.22)$$

Xa podemos escribir a fórmula de Euler-Maclaurin.

Proposición 2.6. *Se $f \in \mathcal{C}^{2k+2}([a, b])$, $k \in \mathbb{N} \cup \{0\}$, cúmprese a seguinte igualdade, coñecida como **fórmula de Euler-Maclaurin**:*

$$T(h) = \int_a^b f(x) dx + \sum_{l=1}^k \frac{B_{2l}}{(2l)!} h^{2l} [f^{2l-1}(b) - f^{2l-1}(a)] + \frac{B_{2k+2}}{(2k+2)!} h^{2k+2} (b-a) f^{2k+2}(\xi) \quad (2.23)$$

para certo $\xi \in (a, b)$, onde os B_i son os números de Bernoulli.

Os números de Bernoulli identifícanse co termo independente dos polinomios de Bernoulli $\hat{B}_k(x)$, e estes últimos quedan totalmente definidos por:

$$\begin{aligned} \hat{B}_0(x) &= 1, \\ \hat{B}_1(x) &= x - \frac{1}{2}, \\ \hat{B}_{i+1}(x) &= (i+1)\hat{B}_i(x), \quad i \geq 1, \\ \hat{B}_{2l+1}(0) &= \hat{B}_{2l+1}(1) = 0 \quad l \geq 1. \end{aligned} \quad (2.24)$$

Nótese que a condición $\hat{B}_{2l+1}(0) = 0$ para $l \geq 1$ implica que os números de Bernoulli de índice impar $i \geq 3$ son nulos.

Observación 2.7. O interese da fórmula de Euler-Maclaurin vese cando $k \geq 1$, pois se $k = 0$ a fórmula redúcese ao que xa sabiamos da fórmula do trapecio composta.

2.2.1. Método de Romberg

O método de Romberg xera unha matriz triangular cuxos elementos son estimacións da integral definida $\int_a^b f(x) dx$, empregando a extrapolación de Richardson reiteradamente sobre a regra do trapecio, e reducindo o paso á metade en cada iteración. A intención deste método é mellorar a orde 2 de converxencia da regra do trapecio composta, que como vimos, ten un erro $\mathcal{O}(h^2)$.

Supoñamos que $f \in \mathcal{C}^{2k+2}([a, b])$. En virtude da fórmula de Euler-Maclaurin temos que:

$$T(h) = \int_a^b f(x) dx + \sum_{l=1}^k \alpha_l h^{2l} + \mathcal{O}(h^{2k+2}), \quad (2.25)$$

onde

$$\alpha_l = \frac{B_{2l}}{(2l)!} h^{2l} [f^{2l-1}(b) - f^{2l-1}(a)], \quad l = 1, \dots, k. \quad (2.26)$$

Agora, chamando $\alpha_0 = \int_a^b f(x) dx$ e declarando $y = h^2$ e $A(y) = T(h)$, obtemos o desenvolvemento asintótico:

$$T(h) = \alpha_0 + \alpha_1 y + \alpha_2 y^2 + \dots + \alpha_k y^k + \mathcal{O}(y^{k+1}), \quad (2.27)$$

podendo así aplicar a técnica da extrapolación de Richardson para obter unha aproximación de α_0 , isto é: de $\int_a^b f(x) dx$.

Tomamos un paso inicial $h_0 = \frac{b-a}{n_0}$, $n_0 \in \mathbb{N}$ (xeralmente será $n_0 = 1$) e $r = \frac{1}{4}$ (é dicir, como $y = h^2$, isto equivale a dividir o paso inicial sucesivamente entre 2).

Lembremos agora o algoritmo da extrapolación de Richardson:

$$A_{m,0} = A(r^m y_0), \quad m = 0, \dots, M, \quad (2.28)$$

$$A_{m,n+1} = \frac{A_{m,n} - r^{n+1} A_{m-1,n}}{1 - r^{n+1}}, \quad n = 0, \dots, M-1, \quad m = n+1, \dots, M. \quad (2.29)$$

Nótese que, como $r = 1/4$,

$$A_{m,0} = A\left(\left(\frac{1}{4}\right)^m y_0\right) = T\left(\frac{h_0}{2^m}\right) = T\left(\frac{b-a}{2^m n_0}\right), \quad (2.30)$$

e o algoritmo convértese en:

$$A_{m,0} = T\left(\frac{b-a}{2^m n_0}\right), \quad m = 0, \dots, M, \quad (2.31)$$

$$A_{m,n+1} = \frac{4^{n+1} A_{m,n} - A_{m-1,n}}{4^{n+1} - 1}, \quad n = 0, \dots, M-1, \quad m = n+1, \dots, M. \quad (2.32)$$

Finalmente, cando remata o proceso dase como aproximación final de $\int_a^b f(x) dx$ o valor $A_{M,M}$.

Para comprender o método, é útil pensar que se traballa en niveis de aproximación. Nun primer nivel aplicamos a regra do trapecio, e para poder empregar a fórmula debemos duplicar cada vez o número de subintervalos (é o que conseguimos cando tomamos $r = 1/4$ para dividir o paso á metade).

No segundo nivel de aproximación é onde empregamos a fórmula (2.32) tomando as parellas contiguas do nivel de aproximación anterior.

Despois pasamos ao nivel de aproximación 3 e facemos o proceso análogo. Continuaremos así ata o último nivel M , fixado de antemán, onde xa só teremos unha parella no nivel anterior. En xeral, se iniciamos o primer nivel con M aproximacións, entón debemos acadar o nivel de aproximación M .

Ilustrémolo cun exemplo.

Exemplo 2.8. Aplicamos o método de Romberg para calcular a integral $\int_{-1}^1 e^x dx$ con $M = 5$ niveis de extrapolación. Obtemos así a seguinte táboa:

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
3.086161269630488				
2.543080634815244	2.362053756543496			
2.399166282614003	2.351194831880255	2.350470903569373		
2.362631333585210	2.350453017242279	2.350403562933081	2.350402494034092	
2.353462010374782	2.350405569304639	2.350402406108797	2.350402387746506	2.350402387329692

Os díxitos en negrita son os díxitos de coincidencia co valor exacto da integral. Vemos como vai aumentando a orde de converxencia, proporcionándonos o método a aproximación

$$\int_{-1}^1 e^x dx \approx \mathbf{2.350402387329692}.$$

2.3. Superconverxencia e converxencia superalxébrica da fórmula do trapecio composta

Tendo en conta a fórmula de Euler-Maclaurin (2.23), é claro que a fórmula do trapecio composta ten un erro $\mathcal{O}(h^{2k+2})$, é dicir, ten orde de converxencia $2k + 2$, se

$$f \in \mathcal{C}^{2k+2}([a, b]), \quad (2.33)$$

e

$$f^{2l-1}(a) = f^{2l-1}(b) \quad l = 1, \dots, k. \quad (2.34)$$

Polo tanto, cando $k \geq 1$, a fórmula é como mínimo de orde 4. Este feito evidencia a **superconverxencia** da fórmula cando f satisfai (2.33) e (2.34).

Observación 2.9. As condicións (2.33) e (2.34) danse cando $f \in \mathcal{C}$, sendo

$$\mathcal{C} = \left\{ g|_{[a,b]} : g \in \mathcal{C}^{2k+2}(\mathbb{R}) \text{ é periódica de período } b - a \right\}.$$

Ademais, isto non só o cumpren as funcións trigonométricas, senón tamén aquelas funcións de clase \mathcal{C}^{2k+2} con soporte compacto contido no intervalo (a, b) .

Tamén remarcar que, se $f \in \mathcal{C}^\infty([a, b])$ satisfai a condición (2.34) para todo $l \in \mathbb{N}$, a orde de converxencia é superior a s , para todo $s \in \mathbb{N}$, que é o que se chama **converxencia superalxébrica** da regra do trapecio composta.

Exemplo 2.10. Aplicamos o método de Romberg para calcular $\int_0^{2\pi} e^{\sin(x)} dx$ con $M = 5$ niveis de aproximación. Temos entón a táboa:

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
6.283185307179585				
6.283185307179586	6.283185307179586			
7.989323439822037	8.558036150702854	8.709692873604405		
7.954927772701778	7.943462550328358	7.902490976970059	7.889678248452054	
7.9549265210128477	7.954926103783204	7.955690340680193	7.956534775024799	7.956796957481946

Destácanse en negrita os díxitos que coinciden coa solución exacta. Obsérvase que a primeira columna do método de Romberg (onde se aplica repetidamente a fórmula do trapecio composta) aproxímase ao valor exacto da solución moito máis rapidamente que o valor da última fila (aproximación que daría o método). Isto débese á converxencia superalxébrica da fórmula, pois vemos que estamos nas hipóteses de $f \in C^\infty([a, b])$ e (2.34).

En conclusión, con estas hipóteses de converxencia superalxébrica sobre a fórmula do trapecio composta, esta é moito máis eficaz e sinxela de aplicar que o método de Romberg, polo que empregar este último método non sería o mellor.

Exemplo 2.11. Tendo en mente a observación 2.9, vexamos que ocorre ca converxencia cando consideramos unha función periódica pero con período distinto de $b - a$, como por exemplo a integral $\int_{-1}^1 e^{\sin(x)} dx$. Aplicando o método de Romberg con $M = 5$ niveis de aproximación obtemos:

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
2.750852775361445				
2.375426387680723	2.250284258453815			
2.304855822610269	2.281332300920118	2.283402170417871		
2.288533938888505	2.283093310981250	2.283210711651992	2.283207672623962	
2.284524777463879	2.283188390322337	2.283194728945076	2.283194475251315	2.283194423496913

Vemos que, neste caso, o valor dado polo método de Romberg aproxímase moito máis que o que nos porporcionaría a fórmula do trapeio composta, isto é porque o integrando non ten período ($b - a$) e xa non estamos polo tanto nas condicións de superconverxencia.

Exemplo 2.12. Vexamos que ocorre se tomamos unha función $f \in C^\infty([a, b])$ que satisfai a condición (2.34) pero que non é periódica. Calculemos por exemplo a integral

$$\int_0^{2\pi} e^{\sin(x)} + 1 + x - \pi x^2 + \frac{x^3}{3} dx$$

con $M = 5$ niveis de aproximación. Obtemos a táboa:

Nivel 1	Nivel 2	Nivel 3	Nivel 4	Nivel 5
-97.573208628798696				
-97.573208628798682	-97.573208628798682			
-95.867070496156245	-95.298357785275428	-95.146701062373879		
-95.901466163276496	-95.912931385649912	-95.953902959008218	-95.966715687526232	
-95.901467414965410	-95.901467832195053	-95.900703595298054	-95.899859160953454	-95.899596978496305

Neste caso, como apreciamos, a fórmula do trapecio composta é máis eficaz que o método de Romberg. Isto é porque, se nos fixamos, atopámonos en condicións de converxencia superalxébrica. En efecto:

- i) $f \in C^\infty([a, b])$.
- ii) Cúmrese (2.34) para todo $l \in \mathbb{N}$, pois:

$$f'(0) = f'(2\pi), f'''(0) = f'''(2\pi),$$

e tódalas derivadas a partir da cuarta coinciden en 0 e 2π .

Superconverxencia

Fagamos un breve resumo dos casos nos que temos a superconverxencia da fórmula do trapecio composta.

A fórmula do trapecio composta é superconverxente para:

1. Funcións que cumbran as condicións (2.33) e (2.34).
2. Particularizando o apartado anterior, temos as funcións regulares periódicas e con derivadas periódicas, sempre e cando se integre nun intervalo de lonxitude o período. Estas funcións poden ser tanto trigonométricas como funcións regulares cuxo soporte é un compacto contido no interior do intervalo de integración.
3. Combinacións lineais de funcións que ofrezan superconverxencia (pola linealidade da integral).

En xeral, se tomamos unha f regular que non cumpra algún dos supostos anteriores, a fórmula do trapecio composta non superará a orde 2 de converxencia.

Capítulo 3

Aplicación ao cálculo de coeficientes de Fourier

As series de Fourier teñen diversas aplicacións en campos como a física ou a enxeñaría. A idea central destas series é que case toda función periódica de período T pode expresarse coma unha suma trigonométrica de senos e cosenos de período T . Facendo uso de [2] veremos unha serie de requisitos para garantir que, efectivamente, exista unha igualdade entre a función considerada e esta serie trigonométrica que recibe o nome de serie de Fourier. Neste marco, dende o punto de vista numérico xurdirá a idea de aproximar a función truncando a súa serie de Fourier.

3.1. Funcións periódicas

Lembremos que unha función $f : \mathbb{R} \rightarrow \mathbb{R}$ dise **periódica** de período T ou *T-periódica* se existe unha constante $T > 0$ tal que

$$f(x + T) = f(x), \quad (3.1)$$

para todo x do dominio de definición de $f(x)$. Nótese ademais que se T é período da función, entón kT con $k \in \mathbb{N}$ tamén é período.

Diremos que $f : [a, b] \rightarrow \mathbb{R}$ é periódica de período T ou *T-periódica* se

$$b - a = kT \quad \text{para certo } k \in \mathbb{N} \quad e \quad f(x + T) = f(x) \quad \forall x \in [a, b] : x + T \in [a, b]. \quad (3.2)$$

Observación 3.1. Entenderemos que o período T é o menor número real positivo que cumpre a condición (3.1) ou (3.2). Non trataremos casos nos que ese período mínimo non exista.

Enunciamos unha propiedade importante das funcións periódicas en canto á integrabilidade das mesmas:

Proposición 3.2. *Sexa f unha función periódica de período T . Se f é integrable nun intervalo de lonxitude T , entón é integrable en calquera intervalo da mesma lonxitude. Ademais, o valor da integral é o mesmo, é dicir:*

$$\int_a^{a+T} f(x) \, dx = \int_b^{b+T} f(x) \, dx, \quad \forall a, b \in \mathbb{R}. \quad (3.3)$$

As funcións trigonométricas como *seno* ou *coseno* son casos básicos de funcións periódicas, nas que o período son 2π radiáns. Imos ver como podemos representar funcións de período $T = 2\pi$ en termos das funcións simples de período 2π ,

$$\{1, \cos(x), \sin(x), \cos(2x), \sin(2x), \dots, \cos(nx), \sin(nx), \dots\} \quad (3.4)$$

que é o demonimado *sistema trigonométrico*. Debemos ter en conta que este sistema é ortogonal no intervalo $[-\pi, \pi]$ (e, consecuentemente, en calquera intervalo de lonxitude 2π , debido á periodicidade).

3.2. Series de Fourier

As funcións periódicas que nos atopamos nos casos prácticos poden chegar a ser bastante complicadas, o noso obxectivo será representalas en termos de funcións periódicas simples para poder traballar mellor con elas.

3.2.1. Series de Fourier para funcións de período 2π

Supoñamos que $f(x)$ é unha función periódica de período 2π que admite unha descomposición nunha serie trigonométrica infinita, de xeito que

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)). \quad (3.5)$$

O noso obxectivo é determinar unha expresión para os coeficientes a_0 , a_n e b_n con $n \in \mathbb{N}$.

Cálculo dos coeficientes

Observación 3.3. Vexamos unha serie de resultados que nos serán de utilidade no cálculo dos coeficientes de Fourier.

Lembramos que, dado $n \in \mathbb{N}$, $n \neq 0$, temos

$$\int_{-\pi}^{\pi} \cos(nx) \, dx = \left[\frac{\sin(nx)}{n} \right]_{x=-\pi}^{x=\pi} = 0, \quad (3.6)$$

$$\int_{-\pi}^{\pi} \sin(nx) \, dx = \left[\frac{-\cos(nx)}{n} \right]_{x=-\pi}^{x=\pi} = 0, \quad (3.7)$$

e, ademais,

$$\int_{-\pi}^{\pi} (\cos(nx))^2 \, dx = \int_{-\pi}^{\pi} \frac{1 + \cos(2nx)}{2} \, dx = \pi, \quad (3.8)$$

$$\int_{-\pi}^{\pi} (\sin(nx))^2 \, dx = \int_{-\pi}^{\pi} \frac{1 - \cos(2nx)}{2} \, dx = \pi. \quad (3.9)$$

Agora, tendo en conta as fórmulas trigonométricas xa coñecidas

$$\begin{aligned} \cos(\alpha) \cos(\beta) &= \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)], \\ \sin(\alpha) \sin(\beta) &= \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)], \\ \sin(\alpha) \cos(\beta) &= \frac{1}{2} [\sin(\alpha + \beta) + \sin(\alpha - \beta)], \end{aligned} \quad (3.10)$$

é doado chegar ao resultado seguinte: dado $m \in \mathbb{N}$, $m \neq n$, temos

$$\int_{-\pi}^{\pi} \cos(nx) \cos(mx) \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\cos(n+m)x + \cos(n-m)x] \, dx = 0, \quad (3.11)$$

$$\int_{-\pi}^{\pi} \sin(nx) \sin(mx) \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\cos(n-m)x - \cos(n+m)x] \, dx = 0, \quad (3.12)$$

$$\int_{-\pi}^{\pi} \sin(nx) \cos(mx) \, dx = \frac{1}{2} \int_{-\pi}^{\pi} [\sin(n+m)x + \sin(n-m)x] \, dx = 0. \quad (3.13)$$

Observese que isto demostra a ortogonalidade do sistema trigonométrico no intervalo $[-\pi, \pi]$.

Agora xa podemos proceder ao cálculo dos coeficientes. Supoñamos que f é integrable no intervalo $[-\pi, \pi]$ e que se dan as condicións necesarias para intercambiar as operacións de suma e integración. Integrando a ambos lados da igualdade (3.5) temos que

$$\int_{-\pi}^{\pi} f(x) \, dx = \int_{-\pi}^{\pi} \frac{a_0}{2} \, dx + \sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos(nx) \, dx + b_n \int_{-\pi}^{\pi} \sin(nx) \, dx \right). \quad (3.14)$$

Por (3.6) e (3.7), a expresión queda reducida a

$$\int_{-\pi}^{\pi} f(x) \, dx = \int_{-\pi}^{\pi} \frac{a_0}{2} \, dx = \pi a_0. \quad (3.15)$$

Fixamos agora un $m \in \mathbb{N}$ e multiplicamos (3.5) por $\cos(mx)$ a ambos lados supoñendo novamente que podemos integrar termo a termo o resultado obtido, é dicir

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \cos(mx) \, dx &= \frac{a_0}{2} \int_{-\pi}^{\pi} \cos(mx) \, dx + \\ &\sum_{n=1}^{\infty} \left(a_n \int_{-\pi}^{\pi} \cos(nx) \cos(mx) \, dx + b_n \int_{-\pi}^{\pi} \sin(nx) \cos(mx) \, dx \right). \end{aligned} \quad (3.16)$$

Novamente, polo resultado (3.6), o primeiro termo á dereita da igualdade é nulo. Ademais, en virtude de (3.11) e (3.13), a expresión anterior queda reducida a un único elemento á dereita da igualdade, o referente a $m = n$, pois no caso de ser distinto tódolos sumandos son nulos. Deste xeito obtemos

$$\int_{-\pi}^{\pi} f(x) \cos(mx) \, dx = a_m \int_{-\pi}^{\pi} (\cos(mx))^2 \, dx = a_m \pi. \quad (3.17)$$

aplicando neste último paso a igualdade (3.8).

Finalmente obtemos que

$$a_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(mx) \, dx. \quad (3.18)$$

Facendo un proceso totalmente análogo (multiplicando por $\sin(mx)$ en ambos lados de (3.5) e integrando), obtemos

$$b_m = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(mx) \, dx. \quad (3.19)$$

Finalmente, de (3.15), (3.18) e (3.19) concluimos que

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx \quad \text{para } n \in \mathbb{N} \cup \{0\}, \quad (3.20)$$

$$b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx \quad \text{para } n \in \mathbb{N}. \quad (3.21)$$

Os números dados por (3.20) e (3.21) son os denominados **coeficientes de Fourier** da función f , e a serie trigonométrica

$$Sf(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)) \quad (3.22)$$

cos coeficientes $\{a_n\}_{n=0}^{\infty}$, $\{b_n\}_{n=1}^{\infty}$ dados por (3.20) e (3.21) denomínase **serie de Fourier** de f .

Observación 3.4. Estamos supoñendo que a serie de Fourier de f converge e ten á función f como suma. Isto pode non ser así, a serie pode converxer a unha función que non sexa f e ata diverxer. Na seguinte sección atenderemos a esta cuestión.

Observación 3.5. Sexa f unha función de período $2L$ definida no intervalo $[-L, L]$. Para desenvolvela en serie de Fourier basta considerar o cambio de variable $x = \frac{Lt}{\pi}$, e entón a serie de Fourier de f quedaría definida por

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right), \quad (3.23)$$

onde

$$a_n = \frac{1}{L} \int_{-L}^L f(x) \cos\left(\frac{n\pi x}{L}\right) dx, \quad n \in \mathbb{N} \cup \{0\}, \quad e \quad b_n = \frac{1}{L} \int_{-L}^L f(x) \sin\left(\frac{n\pi x}{L}\right) dx, \quad n \in \mathbb{N}. \quad (3.24)$$

3.2.2. Serie de Fourier en notación compleja

Supoñamos f unha función definida no intervalo $[-\pi, \pi]$, supoñamos que é posible desenvolve-la como unha serie do tipo (3.22), así

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nt) + b_n \sin(nt)) \quad (3.25)$$

con coeficientes dados por (3.20) e (3.21)

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx \quad n \in \mathbb{N} \cup \{0\}, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx \quad n \in \mathbb{N}. \end{aligned} \quad (3.26)$$

Consideramos as *fórmulas de Euler*

$$\begin{aligned} e^{i\theta} &= \cos(\theta) + i\sin(\theta) \\ e^{-i\theta} &= \cos(\theta) - i\sin(\theta), \end{aligned} \quad (3.27)$$

e tomando $\theta = nx$ inmediatamente vemos que

$$\cos(nx) = \frac{e^{inx} + e^{-inx}}{2} \quad e \quad \sin(nx) = i \frac{e^{-inx} - e^{inx}}{2}. \quad (3.28)$$

Substituíndo agora as expresións (3.28) na serie de Fourier temos que

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(\frac{a_n - ib_n}{2} e^{inx} + \frac{a_n + ib_n}{2} e^{-inx} \right), \quad (3.29)$$

e se denotamos

$$\frac{a_0}{2} = c_0, \quad \frac{a_n - ib_n}{2} = c_n, \quad \frac{a_n + ib_n}{2} = c_{-n} \quad n \in \mathbb{N}, \quad (3.30)$$

entón

$$f(x) = c_0 + \sum_{n=1}^{\infty} (c_n e^{inx} + c_{-n} e^{-inx}), \quad (3.31)$$

ou equivalentemente,

$$\begin{aligned}
 f(x) &= c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=1}^{\infty} c_{-n} e^{-inx} \\
 &= c_0 + \sum_{n=1}^{\infty} c_n e^{inx} + \sum_{n=-\infty}^{-1} c_n e^{inx} \\
 &= \sum_{n=0}^{\infty} c_n e^{inx} + \sum_{n=-\infty}^{-1} c_n e^{inx} \\
 &= \sum_{n=-\infty}^{\infty} c_n e^{inx}.
 \end{aligned} \tag{3.32}$$

Esta é a denominada **forma complexa da serie de Fourier** de $f(x)$. Os coeficientes (3.30) son os denominados **coeficientes complexos de Fourier**. Imos ver unha expresión máis rigurosa destes últimos. Tendo en conta (3.30) podemos ver sen máis que desenvolver as fórmulas que, para todo $n \in \mathbb{N}$, temos que

$$\begin{aligned}
 c_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx - i \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos(nx) - i \sin(nx)) \, dx \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx,
 \end{aligned} \tag{3.33}$$

e

$$\begin{aligned}
 c_{-n} &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx - i \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) (\cos(nx) + i \sin(nx)) \, dx \\
 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{inx} \, dx.
 \end{aligned} \tag{3.34}$$

Polo tanto, podemos escribir

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} \, dx \quad \forall n \in \mathbb{Z}. \tag{3.35}$$

3.3. Convergencia das series de Fourier

Ata este momento sempre establecemos unha igualdade entre a función f e a súa serie de Fourier, mais isto non ter por que ser así. Para expresar dun xeito correcto a serie de

Fourier dunha función f sen estudar previamente se esta converge a f debemos escribir

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)). \quad (3.36)$$

Esta notación o que indica é que a serie trigonométrica do segundo membro *corresponde* á función f , é dicir, ten os coeficientes de Fourier de f como coeficientes, co cal trátase da serie de Fourier de f . Este signo \sim pode substituírse por $=$ cando teñamos asegurado que a serie converge e que a súa suma coincide con f .

Tocante a esta cuestión, plantexamos o problema da converxencia da serie de Fourier: ata que punto a serie de Fourier dunha función é unha representación válida da mesma?

Observación 3.6. É preciso dicir que cabe a posibilidade de que a serie de Fourier sexa *diverxente*. Non é o máis habitual, pero podémosnos atopar con funcións continuas cuxa serie de Fourier é diverxente, ou ben nun punto ou incluso nun conxunto denso.

É ben sabido que unha condición necesaria (que non suficiente) para que se dea a converxencia dunha serie é que o termo xeral desta tenda a cero. No noso contexto, imos ver que debe cumprir a función f para que os coeficientes de Fourier $\{a_n\}_{n=0}^{\infty}$ e $\{b_n\}_{n=1}^{\infty}$ cumpran esta condición. Como resultados preliminares enunciaremos dous lemas:

Lema 3.7. *Sexa f unha función continua nun intervalo $[a, b]$. Entón, para todo $\epsilon > 0$, existe unha función continua en $[a, b]$ e regular a cachos, g , tal que*

$$|f(x) - g(x)| \leq \epsilon, \quad \forall x \in [a, b]. \quad (3.37)$$

Notación 3.8. Diremos que f é unha función continua e regular a cachos no intervalo $[a, b]$ se, dada unha partición $\{a = x_0, \dots, x_n = b\}$ do intervalo $[a, b]$, a función f é continua e posúe unha derivada continua en cada intervalo (x_{k-1}, x_k) , $k = 0, \dots, n$, da partición.

Definición 3.9. Diremos que f é unha **función absolutamente integrable** en $[a, b]$ se é unha función cuxo valor absoluto é integrable, isto é,

$$\int |f(x)| \, dx < \infty. \quad (3.38)$$

Lema 3.10. *Sexa f unha función absolutamente integrable no intervalo $[a, b]$. Entón, para algún $\epsilon > 0$, existe unha función continua en $[a, b]$ e regular a cachos, g , tal que*

$$\int_a^b |f(x) - g(x)| \, dx \leq \frac{\epsilon}{2}. \quad (3.39)$$

Nótese ademais que, se f é absolutamente integrable e periódica, podemos tomar g dentro da clase das funcións periódicas.

Con todo isto, vemos o seguinte teorema.

Teorema 3.11. *Para toda función f absolutamente integrable cúmprese que*

$$\lim_{m \rightarrow \infty} \int_a^b f(x) \cos(mx) \, dx = \lim_{m \rightarrow \infty} \int_a^b f(x) \sin(mx) \, dx = 0. \quad (3.40)$$

Demostración. Pode consultarse a referencia [2]. □

Observación 3.12. Con este teorema acabamos de probar que os coeficientes de Fourier dunha función absolutamente integrable tenden a 0 cando $n \rightarrow \infty$. O interese deste resultado reside en que xa temos unha condición suficiente para poder asegurar a converxencia a 0 dos coeficientes de Fourier.

Condições suficientes para a converxencia

Imos ver unha serie de condicións suficientes para garantir que unha función f poida ser representada pola súa serie de Fourier. Coñecémolas como *condicións de Dirichlet*, e poden ser resumidas nos seguintes enunciados:

1. A función f debe ser periódica.
2. A función f debe estar ben definida, ser continua (salvo posiblemente nun número finito de puntos) e debe ter un número finito de máximos e mínimos.
3. A función f debe ser absolutamente integrable, isto é,

$$\int_{-\frac{T}{2}}^{\frac{T}{2}} |f(x)| \, dx < \infty, \quad (3.41)$$

onde $[-\frac{T}{2}, \frac{T}{2}]$ é o intervalo de integración da función periódica de período T .

Ademais,

- i) Se f é continua no punto x_0 , entón a serie de Fourier de f en x_0 converxe ao valor $f(x_0)$.
- ii) Se f é discontinua no punto x_0 , entón a serie de Fourier de f en x_0 converxe ao valor

$$\frac{f(x_0^+) + f(x_0^-)}{2}, \quad (3.42)$$

sendo $f(x_0^+)$ e $f(x_0^-)$ o límite de f cando x tende a x_0 pola dereita e pola esquerda, respectivamente.

3.3.1. Velocidade de converxencia das series de Fourier

Veremos que canto maior sexa a regularidade da nosa función, maior será a velocidade de converxencia da serie de Fourier á función.

Primeiramente, un resultado previo que nos será de utilidade:

Proposición 3.13. *Sexa f unha función absolutamente continua e periódica no intervalo $[-\pi, \pi]$. Entón a serie de Fourier de f' coincide ca serie de Fourier que obtemos derivando termo a termo a serie de Fourier de f , é dicir, se a serie de Fourier de f é*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)), \quad (3.43)$$

entón a serie de Fourier de f' é

$$\sum_{n=1}^{\infty} (nb_n \cos(nx) - na_n \sin(nx)). \quad (3.44)$$

A continuación vemos un resultado acerca da velocidade de converxencia dos coeficientes.

Teorema 3.14. *Sexa $f \in \mathcal{C}_P^k([-\pi, \pi])$ tal que $f^{(k)}$ é unha función absolutamente continua e 2π -periódica no intervalo $[-\pi, \pi]$ para algún $k \in \mathbb{N}$. Entón*

$$\lim_{n \rightarrow \infty} n^{k+1} a_n = \lim_{n \rightarrow \infty} n^{k+1} b_n = 0, \quad (3.45)$$

onde a_n e b_n son os coeficientes de Fourier de f .

Notación 3.15. $f \in \mathcal{C}_P^k([-\pi, \pi])$ significa que f é unha función 2π -periódica, k veces continuamente derivable en $[-\pi, \pi]$ e que $f^{(i)}(-\pi) = f^{(i)}(\pi)$ para $i = 0, 1, \dots, k$.

Demostración. Véxase [2]. □

Como consecuencia, obtemos un resultado acerca do ritmo de converxencia das series de Fourier.

Corolario 3.16. *Sexa $f \in \mathcal{C}_P^k([-\pi, \pi])$ tal que $f^{(k)}$ é unha función absolutamente continua e 2π -periódica no intervalo $[-\pi, \pi]$ para algún $k \in \mathbb{N}$. Entón, existe unha constante $c > 0$ tal que*

$$\max_{x \in [-\pi, \pi]} |s_m(x) - f(x)| \leq \sum_{n=m+1}^{\infty} \frac{c}{n^{k+1}}. \quad (3.46)$$

Notación 3.17. Denotamos por s_m á m -ésima suma parcial da serie de Fourier de f , isto é

$$s_m(x) = \frac{a_0}{2} + \sum_{k=1}^m (a_k \cos(kx) + b_k \sin(kx)). \quad (3.47)$$

3.4. Cálculo dos coeficientes de Fourier mediante integración numérica

Consideramos unha función f definida en $[-\pi, \pi]$ periódica e de período 2π que cumpre as condicións necesarias para poder establecer a igualdade entre ela e a súa serie de Fourier correspondente. Isto é, f é unha función que pode representarse mediante unha serie trigonométrica:

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)), \quad (3.48)$$

onde os coeficientes da serie veñen dados polas fórmulas xa coñecidas

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) \, dx \quad \forall n \geq 0, \\ b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx \quad \forall n \geq 1. \end{aligned} \quad (3.49)$$

No caso de que a expresión analítica da función periódica non sexa coñecida, non se poden calcular as integrais que permiten obter os coeficientes da serie de Fourier. Do mesmo xeito, a función pode ser coñecida pero as integrais non ter unha expresión ben definida. Neste caso adoitamos coñecer algunha mostra de valores da función ao longo dun período, en particular do intervalo. Vexamos como afrontar esta situación mediante a *fórmula do trapecio composta*.

Temos unha función f de período 2π e coñecemos $k + 1$ valores da función

$$f(y_1), \dots, f(y_{k+1}), \quad -\pi = y_1 < \dots < y_{k+1} = \pi, \quad (3.50)$$

con $f(y_1) = f(y_{k+1})$ por periodicidade.

Aplicamos a fórmula do trapecio composta para calcular os coeficientes de Fourier:

Cálculo do coeficiente a_0

Sabemos por (3.15) que

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \, dy \approx \frac{1}{\pi} \frac{h}{2} (f(y_1) + 2f(y_2) + \dots + 2f(y_k) + f(y_{k+1})). \quad (3.51)$$

Agora como $h = \frac{2\pi}{k+1}$ e $f(y_1) = f(y_{k+1})$, obtemos

$$a_0 \approx \frac{2}{k+1} ((f(y_1) + f(y_2) + \dots + f(y_k))). \quad (3.52)$$

Cálculo dos coeficientes a_n

Sabemos por (3.20) que

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ny) \, dy \\ &\approx \frac{1}{\pi} \frac{h}{2} (f(y_1) \cos(ny_1) + 2f(y_2) \cos(ny_2) + \cdots + 2f(y_k) \cos(ny_k) + f(y_{k+1}) \cos(ny_{k+1})). \end{aligned} \quad (3.53)$$

Agora como $h = \frac{2\pi}{k-1}$ e $f(y_1) = f(y_{k+1})$, obtemos

$$\begin{aligned} a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \cos(ny) \, dy \approx \frac{1}{\pi} \frac{2}{k+1} (f(y_1) \cos(ny_1) + f(y_2) \cos(ny_2) + \cdots \\ &\quad + f(y_k) \cos(ny_k)). \end{aligned} \quad (3.54)$$

Cálculo dos coeficientes b_n

Sabemos por (3.21) que

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ny) \, dy \\ &\approx \frac{1}{\pi} \frac{h}{2} (f(y_1) \sin(ny_1) + 2f(y_2) \sin(ny_2) + \cdots + 2f(y_k) \sin(ny_k) + f(y_{k+1}) \sin(ny_{k+1})). \end{aligned} \quad (3.55)$$

Agora como $h = \frac{2\pi}{k-1}$ e $f(y_1) = f(y_{k+1})$, obtemos

$$\begin{aligned} b_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \sin(ny) \, dy \approx \frac{1}{\pi} \frac{2}{k+1} (f(y_1) \sin(ny_1) + f(y_2) \sin(ny_2) + \cdots \\ &\quad + f(y_k) \sin(ny_k)). \end{aligned} \quad (3.56)$$

É lóxico pensar que a aproximación que nos proporciona a fórmula de integración numérica será mellor canto maior sexa o número de nodos (valores da mostra), mais unha vantaxe que nos proporcionan as series de Fourier é que, como vimos na sección anterior, a converxencia da serie de Fourier á función que representa é moi rápida cando a función é o suficientemente regular. Este tipo de funcións, se ademais non son de carácter moi oscilante, resultan ser a combinación perfecta para empregar a fórmula do trapecio composta, pois obteríamos aproximacións moi boas para un número de valores, k , relativamente pequeno.

Capítulo 4

Transformada discreta de Fourier vía FFT

Aínda que o obxectivo deste capítulo é tratar a transformada rápida de Fourier (FFT), antes debemos ter unha serie de coñecementos previos acerca da transformada de Fourier e da transformada discreta de Fourier (DFT), pois a transformada rápida de Fourier non é máis que un algoritmo para o cálculo eficaz da transformada discreta de Fourier. Para a realización deste capítulo consultamos as referencias [3] e [4].

4.1. Introducción á transformada de Fourier

Podemos pensar a transformada de Fourier como a forma continua da serie de Fourier. É unha ferramenta básica no estudo de funcións oscilantes. Ven sendo unha operación matemática que transforma unha función no dominio do tempo (t) ao dominio da frecuencia (λ), todo isto sen alterar en ningún momento a información que esta contén. As súas extensas aplicacións abarcan ámbitos como a medicina, as comunicacións, o campo electromagnético, o procesamento de sinais e imaxes...

Definición 4.1. Sexa $f : \mathbb{R} \rightarrow \mathbb{C}$ unha función continuamente diferenciable e absolutamente integrable. Definimos a **transformada de Fourier de f** como a función $\hat{f} : \mathbb{R} \rightarrow \mathbb{C}$ definida por

$$\mathcal{F}[f](\lambda) = \hat{f}(\lambda) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t)e^{i\lambda t} dt. \quad (4.1)$$

Definición 4.2. Se as funcións f e \hat{f} son integrables, entón definimos a **transformada**

inversa de Fourier de f como

$$\mathcal{F}^{-1}[f](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\lambda) e^{i\lambda x} d\lambda. \quad (4.2)$$

4.2. Transformada discreta de Fourier (DFT)

Cando traballamos no estudo das funcións oscilantes, é común atoparnos con sinais que non son continuas, polo que non podemos facer uso da transformada de Fourier nin da propia serie de Fourier. Nalgunhas aplicacións, especialmente as relacionadas coa análise de información almacenada en soportes dixitais, adoitamos ter un conxunto de mostras discreto, polo que precisamos dunha adaptación da transformada de Fourier para estes casos. Desta necesidade xorde a transformada discreta de Fourier.

Definición 4.3. Sexa \mathcal{S}_n un conxunto de **sinais discretas** (ou mostras) de tamaño n . Cada elemento de \mathcal{S}_n é un vector de números complexos $y = (y_0, \dots, y_{n-1})^T$ que pode ser pensado como unha sinal discreta onde cada elemento y_j , $j \in \{0, \dots, n-1\}$, fai referencia ao valor da sinal no nodo temporal $t = t_j$.

A idea subxacente á DFT é un simple cambio de coordenadas en \mathbb{C}^n entre dúas bases cun significado un tanto especial. Estamos a falar de:

- En primeiro lugar, a base canónica dada polos vectores da forma $e_j[i] = 0$ se $i \neq j$ e $e_j[j] = 1$. Nótese que, cando escribimos $y = (y_0, \dots, y_{n-1})^T$, en realidade estamos escribindo $y = y_0 e_0 + \dots + y_{n-1} e_{n-1} = \sum_{j=0}^{n-1} y_j e_j$. Con isto estamos considerando unha representación no dominio do *tempo*.

- Por outra banda, tomamos a base que nos describen estas sinais en termos das *frecuencias*, $\{\tilde{w}_k\}_{k=0}^{n-1}$. Esta é unha base ortogonal formada polos vectores $\tilde{w}_k = \sum_{j=0}^{n-1} e^{\frac{2i\pi k j}{n}} e_j$.

Xa podemos definir a transformada discreta de Fourier:

Definición 4.4. Sexa $y = \{y_j\}_{j=0}^{n-1} \in \mathcal{S}_n$. A **transformada discreta de Fourier** de y ven dada por unha secuencia $(\mathcal{F}_n \{y\})_k = \hat{y}_k$, onde

$$\hat{y}_k = \sum_{j=0}^{n-1} y_j e^{\frac{-2i\pi k j}{n}}. \quad (4.3)$$

Polo que se refire ao cálculo da transformada discreta de Fourier, podemos observar que é equivalente ao seguinte produto de matrices:

$$\mathcal{F}_n \{y\} = \hat{y} = (\bar{F}_n)(y), \quad (4.4)$$

onde $\hat{y} = (\hat{y}_0, \dots, \hat{y}_{n-1})^T$, $y = (y_0, \dots, y_{n-1})^T$ e

$$F_n = \begin{pmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^{n-1} \\ 1 & w^2 & w^4 & \dots & w^{2(n-1)} \\ \vdots & & \ddots & \ddots & \\ 1 & w^{n-1} & w^{2(n-1)} & \dots & w^{(n-1)^2} \end{pmatrix}, \quad (4.5)$$

tomando $w = e^{\left(\frac{2\pi i}{n}\right)}$.

Observación 4.5. A matriz F_n é unha matriz simétrica de orde $n \times n$.

Inmediatamente, podemos deducir que a **transformada discreta de Fourier inversa** virá dada por

$$\mathcal{F}_n^{-1}(\hat{y}) = y, \quad (4.6)$$

sendo

$$y_j = \frac{1}{n} \sum_{k=0}^{n-1} \hat{y}_k w^{jk}. \quad (4.7)$$

Unha propiedade importante a ter en conta, e que tomará relevancia na seguinte sección, é que, se temos unha mostra $y \in \mathcal{S}_n$, entón

$$(\mathcal{F}\{y\})_{n-k} = (\bar{\mathcal{F}}\{y\})_k. \quad (4.8)$$

Equivalentemente,

$$\hat{y}_{n-k} = \bar{\hat{y}}_k, \quad k = 0, \dots, n-1. \quad (4.9)$$

Esta igualdade permítenos reducir o noso cálculo a simplemente os valores $\hat{y}_0, \dots, \hat{y}_{\frac{n}{2}-1}$ se n é par, e $\hat{y}_0, \dots, \hat{y}_{\lfloor \frac{n}{2} \rfloor - 1}$ cando n sexa impar.

Notación 4.6. Denotamos por $\lfloor x \rfloor$ ao maior número enteiro igual ou menor que x .

4.2.1. Relación cos coeficientes de Fourier

Como xa mencionamos, a idea baixo a transformada discreta de Fourier é axudar á aproximación dos coeficientes de Fourier dunha función continua $f(t)$.

Comecemos aplicando a fórmula de cuadratura do trapecio composta para realizar unha aproximación da integral

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(t) dt \quad (4.10)$$

cun paso $h = \frac{2\pi}{n}$.

En base ao visto no capítulo 1, podemos concluír que

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(t) dt \approx \frac{1}{2\pi} \frac{2\pi}{n} \left[\frac{Y_0}{2} + Y_1 + \dots + Y_{n-1} + \frac{Y_n}{2} \right], \quad (4.11)$$

onde $Y_j = F(-\pi + hj) = F(-\pi + \frac{2\pi j}{n})$, $j = 0, \dots, n$. Agora, se $F(t)$ é 2π -periódica, entón $Y_0 = Y_n$, e polo tanto a fórmula anterior convértese en

$$\frac{1}{2\pi} \int_{-\pi}^{\pi} F(t) dt \approx \frac{1}{n} \sum_{j=0}^{n-1} Y_j. \quad (4.12)$$

Aplicando esta fórmula ao cálculo dos coeficientes de Fourier (3.35), temos que

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} dt \approx \frac{1}{n} \sum_{j=0}^{n-1} f\left(-\pi + \frac{2\pi j}{n}\right) e^{(-\frac{2\pi ijk}{n} + ik\pi)}, \quad (4.13)$$

co cal

$$c_k \approx \frac{e^{ik\pi}}{n} \sum_{j=0}^{n-1} y_j \bar{w}^{jk} = \frac{e^{ik\pi}}{n} \hat{y}_k, \quad (4.14)$$

onde

$$y_j = f\left(-\pi + \frac{2\pi j}{n}\right) \quad \text{e} \quad w = e^{(\frac{2\pi i}{n})}. \quad (4.15)$$

Notación 4.7. Neste caso, por \bar{w} entendemos o conxugado de w .

Nótese que esta expresión non nos aproxima os c_k se $k \geq n$, pois os c_k non son funcións n -periódicas. De feito, esta expresión só aproxima os c_k cando k é un valor relativamente pequeno en comparanza con n , porque a regra do trapecio proporciónanos aproximacións precisas só se o tamaño do paso, $h = \frac{2\pi}{n}$, é relativamente pequeno comparado coa frecuencia k .

Deste xeito, acabamos de atopar un método bixectivo para calcular aproximacións dos coeficientes de Fourier a partir dunha serie de mostras ou sinais.

4.3. Transformada rápida de Fourier (FFT)

Calcular a transformada discreta de Fourier mediante a fórmula (4.3) pode resultarnos moi custoso. Obsérvase que ao multiplicar o vector $\hat{y} = (\hat{y}_0, \dots, \hat{y}_{n-1})^T$ (de dimensión $n \times 1$) pola matriz \bar{F}_n (de dimensión $n \times n$) fanse un total de $n(n-1) + n^2$ operacións (sumas e multiplicacións). Por este motivo, para un n relativamente común como pode ser $n = 1000$, teríamos que facer máis dun millón de operacións. De aí a importancia de buscar un algoritmo que calcule a DFT cun custo computacional moito menor, a *transformada rápida de Fourier*, FFT.

O algoritmo FFT baséase en 3 pasos principais:

1. Dividimos o problema en varios subproblemas análogos menores.
2. Resolvemos recursivamente cada un dos subproblemas co mesmo algoritmo.

3. Obtemos a solución do problema orixinal combinando a solución de cada subproblema.

Algoritmo FFT

Este algoritmo está baseado en dividir a expresión (4.3) en dúas sumas, por un lado os termos *pares* e por outro os *impares*.

Por comodidade, tomaremos como número de nodos $n = 2N$. Agora consideremos a mostra $y = (y_0, \dots, y_{2N-1})$ (con período $n = 2N$). Entón os \hat{y}_k calcúlanse mediante

$$\hat{y}_k = \sum_{j=0}^{2N-1} y_j \bar{w}^j k. \quad (4.16)$$

Separando a ecuación anterior en dous sumandos segundo estes sexan pares ou impares, obtemos

$$\hat{y}_k = \sum_{j=0}^{N-1} y_{2j} \bar{w}^{2jk} + \sum_{j=0}^{N-1} y_{2j+1} \bar{w}^{(2j+1)k}. \quad (4.17)$$

Lembramos que $w = e^{\left(\frac{2\pi i}{n}\right)}$ e $n = 2N$. Agora consideremos $W = e^{\left(\frac{2\pi i}{N}\right)} = w^2$, e así

$$\hat{y}_k = \sum_{j=0}^{N-1} y_{2j} \bar{W}^{jk} + \bar{w}^k \left(\sum_{j=0}^{N-1} y_{2j+1} \bar{W}^{jk} \right). \quad (4.18)$$

Nótese que a expresión anterior non é máis que expresar \hat{y}_k como suma de dúas transformadas de Fourier (onde $n = N$). En efecto,

$$\hat{y}_k = \mathcal{F}_N [\{y_0, y_2, \dots, y_{2N-2}\}]_k + \bar{w}^k \mathcal{F}_N [\{y_1, y_3, \dots, y_{2N-1}\}]_k \quad (4.19)$$

para $0 \leq k \leq 2N - 1$.

Agora substituímos k por $k + N$ na ecuación anterior e, empregando as seguintes propiedades:

- *Periodicidade* : $\mathcal{F}_N[y_{2j}]$ e $\mathcal{F}_N[y_{2j+1}]$ teñen período N ,
- *Simetría* : $\bar{w}^{k+N} = \bar{w}^k e^{-\pi i} = -\bar{w}^k$,

temos que, se $0 \leq k \leq N - 1$, entón

$$\begin{aligned} \hat{y}_k &= \mathcal{F}_N [\{y_0, y_2, \dots, y_{2N-2}\}]_k + \bar{w}^k \mathcal{F}_N [\{y_1, y_3, \dots, y_{2N-1}\}]_k, \\ \hat{y}_{k+N} &= \mathcal{F}_N [\{y_0, y_2, \dots, y_{2N-2}\}]_k - \bar{w}^k \mathcal{F}_N [\{y_1, y_3, \dots, y_{2N-1}\}]_k. \end{aligned} \quad (4.20)$$

En conclusión, acabamos de escribir $(\mathcal{F}_{2N}y)_k = \hat{y}_k$ para $0 \leq k \leq 2N - 1$ en termos de $\mathcal{F}_N [\{y_0, y_2, \dots, y_{2N-2}\}]_k$ e $\mathcal{F}_N [\{y_1, y_3, \dots, y_{2N-1}\}]_k$ para $0 \leq k \leq N - 1$.

Dun xeito análogo, podemos describir a transformada inversa de Fourier como

$$\begin{aligned} y_k &= \frac{1}{2} \left\{ \mathcal{F}_N^{-1} [\{\hat{y}_0, \hat{y}_2, \dots, \hat{y}_{2N-2}\}]_k + w^k \mathcal{F}_N^{-1} [\{\hat{y}_1, \hat{y}_3, \dots, \hat{y}_{2N-1}\}]_k \right\}, \\ y_{k+N} &= \frac{1}{2} \left\{ \mathcal{F}_N^{-1} [\{\hat{y}_0, \hat{y}_2, \dots, \hat{y}_{2N-2}\}]_k - w^k \mathcal{F}_N^{-1} [\{\hat{y}_1, \hat{y}_3, \dots, \hat{y}_{2N-1}\}]_k \right\}, \end{aligned} \quad (4.21)$$

para $0 \leq k \leq N-1$.

Matricialmente, podemos expresar a transformada discreta de Fourier como

$$\mathcal{F}_{2N} y = (\bar{F}_{2N}) (y) = \begin{pmatrix} I_N & \bar{D}_N \\ I_N & -\bar{D}_N \end{pmatrix} \begin{pmatrix} \bar{F}_N & 0 \\ 0 & \bar{F}_N \end{pmatrix} \begin{pmatrix} y_{2j} \\ y_{2j+1} \end{pmatrix}, \quad (4.22)$$

sendo F_N a matriz definida en (4.5), I_N a matriz identidade de orde $N \times N$ e D_N unha matriz diagonal con elementos $1, w, w^2, \dots, w^{N-1}$.

Deste xeito transformamos unha DFT de $2N$ mostras en dúas de tamaño N . Considerando a complexidade de cálculo da DFT, este método fai o proceso substancialmente máis doado. Quédanos entón un algoritmo recursivo que ao chegar a certo tamaño (o equivalente ao número de mostras) deixará de executarse e empregará o algoritmo usual da DFT (o descrito na sección anterior).

Así, en lugar dun custo de $n^2 = 4N^2$ que tería o método usual da DFT, a FFT ofrécenos un custo de $2N^2 + N + 1$ multiplicacións, o cal esta practicamente reducindo á metade o usual. Esta diferenza é moi notable na práctica, tendo en conta a gran cantidade de mostras que normalmente tomamos.

Estes algoritmos aplícanse cando n é par, pero para aproveitar mellor a recursividade é preferible que n sexa unha potencia de 2, isto é $n = 2^L$. Deste xeito, podemos iterar o algoritmo un total de L veces. Así, unha vez chegamos á iteración L atopámonos cunha DFT de tamaño 1 (unha mostra), derivando nun cálculo moi sinxelo.

Outra cuestión é: **cal é o custo real deste algoritmo?** É dicir, cantas operacións (multiplicacións) son necesarias?

Supoñamos K_L o número de multiplicacións necesarias para calcular $\mathcal{F}[y]$ mediante o método anterior cun número de mostras $n = 2^L$. Fixándonos en (4.20), vemos que é necesario calcular $\mathcal{F}[y_{2j}]$ e $\mathcal{F}[y_{2j+1}]$, con $j = 0, \dots, 2^L - 1$. Ademais, nótese que $n = 2^L \implies N = 2^{L-1}$. Así mesmo, o número de multiplicacións necesarias é proporcional a

$$2K_{L-1} + N = 2K_{L-1} + 2^{L-1}, \quad (4.23)$$

e ademais

$$K_L \approx 2K_{L-1} + 2^{L-1}. \quad (4.24)$$

Con isto quérese dicir que:

$$L = 0 \implies K_0 = 0,$$

$$\begin{aligned}
L = 1 &\implies K_1 \approx 1, \\
L = 2 &\implies K_2 \approx 2 \cdot 2^1, \\
L = 3 &\implies K_3 \approx 3 \cdot 2^2, \\
L = 4 &\implies K_4 \approx 4 \cdot 2^3, \\
&\vdots
\end{aligned}$$

e así sucesivamente.

Chégase así a unha fórmula xeral para o número de multiplicacións, dada por

$$K_L \approx L \cdot 2^{L-1}. \quad (4.25)$$

De aí que, cando escollemos como número de mostras $n = 2^L$, é dicir $L = \log_2 n$, podemos ver (na fórmula 4.25) que o número de multiplicacións necesarias vai ser proporcional a $n \log_2 n$.

4.4. Aproximación dos coeficientes de Fourier mediante a DFT vía FFT

Consideremos unha función f definida no intervalo $[-\pi, \pi]$ con serie de Fourier asociada

$$\sum_{k=-\infty}^{\infty} c_k e^{ikx}, \quad (4.26)$$

onde, como vimos no capítulo anterior, os coeficientes de Fourier veñen dados pola fórmula (3.35), é dicir,

$$c_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-ikx} dx. \quad (4.27)$$

Observamos que, por ser f unha función de valores reais, entón os valores c_k e c_{-k} son complexos conxugados, é dicir

$$c_k = \bar{c}_{-k}. \quad (4.28)$$

Ademais, como sabemos que tanto c_k como c_{-k} son valores complexos, podemos expresalos do seguinte xeito

$$\begin{aligned}
c_k &= \operatorname{Re}(c_k) + i\operatorname{Im}(c_k), \\
\bar{c}_{-k} &= \operatorname{Re}(c_{-k}) - i\operatorname{Im}(c_{-k}).
\end{aligned} \quad (4.29)$$

Daquela, volvendo a (4.28), temos que

$$\operatorname{Re}(c_k) + i\operatorname{Im}(c_k) = \operatorname{Re}(c_{-k}) - i\operatorname{Im}(c_{-k}) \implies \begin{cases} \operatorname{Re}(c_k) = \operatorname{Re}(c_{-k}) \\ \operatorname{Im}(c_k) = -\operatorname{Im}(c_{-k}) \end{cases}. \quad (4.30)$$

Por unha banda, consideramos a expresión da serie de Fourier de f como

$$c_0 + \sum_{k=1}^{\infty} c_k e^{ikx} + \sum_{k=1}^{\infty} c_{-k} e^{-ikx} = c_0 + \sum_{k=1}^{\infty} (c_k e^{ikx} + c_{-k} e^{-ikx}). \quad (4.31)$$

Por outra banda, tendo en mente a fórmula de Euler e a paridade das funcións cos e sin, temos

$$\begin{aligned} c_k e^{ikx} + c_{-k} e^{-ikx} &= c_k (\cos(kx) + i \sin(kx)) + c_{-k} (\cos(-kx) + i \sin(-kx)) \\ &= c_k \cos(kx) + i c_k \sin(kx) + c_{-k} \cos(kx) - i c_{-k} \sin(kx) \\ &= (c_k + c_{-k}) \cos(kx) + i(c_k - c_{-k}) \sin(kx). \end{aligned} \quad (4.32)$$

Separando agora os números complexos en parte real e parte imaxinaria e usando a relación obtida en (4.30), vemos que

$$c_k e^{ikx} + c_{-k} e^{-ikx} = 2\operatorname{Re}(c_k) \cos(kx) - 2\operatorname{Im}(c_k) \sin(kx). \quad (4.33)$$

Tendo en conta isto, e volvendo a (4.31), a serie de Fourier de f adopta a expresión

$$c_0 + \sum_{k=1}^{\infty} [2\operatorname{Re}(c_k) \cos(kx) - 2\operatorname{Im}(c_k) \sin(kx)]. \quad (4.34)$$

Antes de máis nada, tomamos una mostra de valores reais $y = \{y_j\}_{j=1}^N$. Consecuentemente, como xa vimos en (4.3), a transformada discreta de Fourier (DFT) de y virá dada polo vector $\hat{y} = \{\hat{y}_k\}_{k=1}^N$, sendo

$$\hat{y}_k = \sum_{j=1}^N y_j e^{-\frac{2\pi i}{N}(k-1)(j-1)}. \quad (4.35)$$

Hai que ter en conta que, en base ás propiedades de periodicidade e conxugación da DFT, sábese que

$$\hat{y}_{N-k+2} = \bar{\hat{y}}_k, \quad k = 2, \dots, N. \quad (4.36)$$

Polo tanto, o único conxunto que nos interesa calcular debido á información que contén é o dado por $\{\hat{y}_1, \dots, \hat{y}_{\lfloor \frac{N}{2} \rfloor + 1}\}$, pois tódolos demais elementos serán repeticións destes. Nótese que, no caso de que N sexa par, o elemento $\hat{y}_{\frac{N}{2}+1}$ será un valor real.

Aproximación dos coeficientes de Fourier

Facendo uso de toda a información recollida ata este momento, imos implementar un método que empregará o algoritmo da transformada rápida de Fourier (FFT) para calcular os coeficientes de Fourier en función da transformada discreta de Fourier (DFT) e mediante a fórmula de cuadratura do trapecio composta.

Primeiramente, consideremos o intervalo de definición da nosa función f , $[-\pi, \pi]$. Dividimos o intervalo en N subintervalos de lonxitude $h = \frac{2\pi}{N}$, obtendo así $N + 1$ nodos equiespaciados

$$-\pi = y_1 < \dots < y_{N+1} = \pi. \quad (4.37)$$

Obsérvese que entón estes puntos quedan totalmente definidos por

$$y_j = -\pi + (j - 1)h = -\pi + (j - 1)\frac{2\pi}{N}. \quad (4.38)$$

Empregando esta partición para o cálculo dos coeficientes de Fourier mediante a fórmula do trapecio composta temos que

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)e^{-ikx} dx \approx \frac{1}{2\pi} \left[\frac{h}{2} [f(y_1) + f(y_{N+1})] + h \sum_{j=2}^N f(y_j)e^{-iky_j} \right] \\ &= \frac{1}{2\pi} \frac{h}{2} \left[f(y_1) + 2 \sum_{j=2}^N f(y_j)e^{-iky_j} + f(y_{N+1}) \right] = \frac{1}{2N} \left[f(-\pi) + 2 \sum_{j=2}^N f(y_j)e^{-iky_j} + f(\pi) \right] \\ &= \frac{1}{2N} \left[f(-\pi) + 2 \sum_{j=2}^N f(y_j)e^{-ik(-\pi+(j-1)\frac{2\pi}{N})} + f(\pi) \right] \\ &= \frac{1}{2N} \left[f(-\pi) + 2 \sum_{j=2}^N f(y_j)e^{ik\pi} e^{-ik(j-1)\frac{2\pi}{N}} + f(\pi) \right] \\ &= \frac{1}{2N} \left[f(-\pi) + 2e^{ik\pi} \sum_{j=2}^N f(y_j)e^{-ik(j-1)\frac{2\pi}{N}} + f(\pi) \right]. \end{aligned} \quad (4.39)$$

Observación 4.8. Tocante á elección do número de mostrás, cabe mencionar que cando $N + 1 < 4k$ a aproximación anterior non ten por que ser fiable. Isto débese ao feito de ter que integrar as funcións $\cos(kx)$ e $\sin(kx)$, pois estas teñen $4k$ extremos no intervalo $[-\pi, \pi]$.

Simultaneamente, o máis importante é decatarse de que o sumatorio involucrado na expresión anterior, $\sum_{j=2}^N f(y_j)e^{-ik(j-1)\frac{2\pi}{N}}$, correspóndese ca DFT dunha mostra de valores $\{f(y_j)\}_{j=1}^N$, pois nótese que xa non temos en conta o valor asociado ao nodo $N + 1$, que é $f(\pi)$.

É máis, a transformada discreta de Fourier para o conxunto de mostrás $\{f(y_j)\}_{j=1}^N$ segundo a súa definición sabemos que é

$$\{\hat{y}_k\}_{k=1}^N, \quad \text{onde } \hat{y}_k = \sum_{j=1}^N f(y_j)e^{-\frac{2\pi i}{N}(j-1)(k-1)}. \quad (4.40)$$

Pero se nos fixamos podemos observar que

$$\begin{aligned}
\frac{1}{N}\hat{y}_k &= \frac{1}{N} \sum_{j=1}^N f(y_j) e^{-\frac{2\pi i}{N}(j-1)(k-1)} = \frac{1}{N} \left[f(y_1) + \sum_{j=2}^N f(y_j) e^{-\frac{2\pi i}{N}(j-1)(k-1)} \right] \\
&= \frac{1}{2N e^{ik\pi}} \left[f(-\pi) + 2e^{ik\pi} \sum_{j=2}^N f(y_j) e^{-\frac{2\pi i}{N}(j-1)(k-1)} + f(\pi) \right] \\
&\quad + \frac{1}{2N e^{ik\pi}} \left((2e^{ik\pi} - 1)f(-\pi) - f(\pi) \right).
\end{aligned} \tag{4.41}$$

Logo podemos identificar a relación entre o primeiro sumando da ecuación anterior e a aproximación dada por (4.39), chegando a que

$$\frac{1}{N}\hat{y}_k \approx c_{k-1} \frac{1}{e^{ik\pi}} + \frac{1}{2N e^{ik\pi}} \left((2e^{ik\pi} - 1)f(-\pi) - f(\pi) \right). \tag{4.42}$$

Finalmente, despxando temos que

$$c_{k-1} \approx \frac{1}{N} \left(e^{ik\pi} \hat{y}_k - \frac{1}{2} \left[(2e^{ik\pi} - 1)f(-\pi) - f(\pi) \right] \right), \quad k = 1, \dots, \left\lfloor \frac{N}{2} \right\rfloor + 1. \tag{4.43}$$

Deste xeito, somos capaces de obter o valor dos coeficientes de Fourier $c_0, \dots, c_{\lfloor \frac{N}{2} \rfloor}$ coa precisión que nos proporciona a regra do trapezio composta de $N+1$ nodos, e a comodidade en canto a número de operacións da DFT mediante o algoritmo FFT.

Observación 4.9. No caso de tratar cunha función 2π -periódica obtemos trivialmente a relación $c_{k-1} \approx \frac{e^{ik\pi}}{N} \hat{y}_k$, pois neste caso $f(-\pi) = f(\pi)$.

Cómpre dicir que esta idea ten especial interese cando tratamos con funcións regulares periódicas e con derivadas periódicas, sempre e cando a integral estea definida nun intervalo de lonxitude o período, pois así estaríamos en condicións de garantir a superconverxencia da fórmula do trapezio composta (como vimos no capítulo 2, grazas á fórmula de Euler Maclaurin), e polo tanto acadaríamos unha aproximación especialmente boa cun custo notablemente menor do habitual dos coeficientes de Fourier.

Capítulo 5

Resultados numéricos

Neste capítulo imos calcular os coeficientes de Fourier dunha función mediante a DFT vía FFT empregando un código de Matlab, observando tamén a converxencia da serie de Fourier á función orixinal e os erros cometidos.

5.1. Aproximación dos coeficientes de Fourier dunha función discontinua

Consideramos a función pulso

$$f(t) = \begin{cases} -\frac{t}{2} - \frac{\pi}{2} & \text{se } -\pi \leq t < 0, \\ -\frac{t}{2} + \frac{\pi}{2} & \text{se } 0 \leq t \leq \pi, \end{cases} \quad (5.1)$$

con $t \in [-\pi, \pi]$. Trátase dunha función discontinua no intervalo $[-\pi, \pi]$. Calculamos unha aproximación dos coeficientes de Fourier empregando a DFT vía FFT, tal e como explicamos no capítulo anterior.

Dividimos o intervalo $[-\pi, \pi]$ en $N = 2^{20}$ subintervalos da mesma lonxitude, e empregando a expresión

$$c_{k-1} \approx \frac{1}{N} \left(e^{ik\pi} \hat{y}_k - \frac{1}{2} \left[(2e^{ik\pi} - 1)f(-\pi) - f(\pi) \right] \right), \quad k = 1, \dots, \left\lfloor \frac{N}{2} \right\rfloor + 1, \quad (5.2)$$

calculamos os coeficientes de Fourier da función f . En realidade, grazas á periodicidade da función, sabemos que $f(-\pi) = f(\pi)$, polo tanto

$$c_{k-1} \approx \frac{e^{ik\pi}}{N} \hat{y}_k, \quad k = 1, \dots, \left\lfloor \frac{N}{2} \right\rfloor + 1. \quad (5.3)$$

Observación 5.1. Cabe mencionar que, se pensamos $N = 2^p$ con $p \in \mathbb{N}$, a nosa elección do número de mostrás baséase en establecer como criterio que: tomando un parámetro de tolerancia de 10^{-6} e denotando por $c_{\lfloor \frac{N}{2} \rfloor}^i$ ao coeficiente de Fourier $\lfloor \frac{N}{2} \rfloor$ -ésimo calculado con $N = 2^i$ ($i \in \mathbb{N}$), verifícase que

$$\frac{\left| c_{\lfloor \frac{N}{2} \rfloor}^i - c_{\lfloor \frac{N}{2} \rfloor}^{i+1} \right|}{1 + \left| c_{\lfloor \frac{N}{2} \rfloor}^{i+1} \right|} < 10^{-6}. \quad (5.4)$$

Con respecto á serie de Fourier da función, lembremos que mediante a expresión

$$c_0 + \sum_{k=1}^m [2\operatorname{Re}(c_k) \cos(kx) - 2\operatorname{Im}(c_k) \operatorname{sen}(kx)] \quad (5.5)$$

podemos calcular a m -ésima suma parcial da serie de Fourier de f , denotada por s_m .

Notación 5.2. A partir de agora, referirémonos á serie de Fourier de $m + 1$ sumandos, é dicir á m -ésima suma parcial da serie de Fourier, s_m , como *suma parcial de orde m* .

Así, facendo uso dos coeficientes calculados usando a DFT vía FFT, obtemos unha aproximación da serie de Fourier de f dada polas sumas parciais m -ésimas, representadas na Figura 5.1.

Como xa sabemos, se f e f' son funcións continuas, salvo nun número finito de puntos de discontinuidade de tipo salto, como é o noso caso, a serie de Fourier converxe puntualmente a $f(t)$ nos puntos de continuidade e á media dos límites laterais nos puntos de discontinuidade. Neste caso, a función coa que estamos a traballar presenta unha singularidade (discontinuidade de salto) no punto $t = 0$. Na figura 5.1 pódese apreciar que, efectivamente, cando $t = 0$ as sumas parciais converxen á media dos límites laterais $\frac{f(0^-) + f(0^+)}{2} = \frac{-\frac{\pi}{2} + \frac{\pi}{2}}{2} = 0$. Pola contra, cando $t \neq 0$ as sumas parciais aproximan o valor da función f nes puntos t .

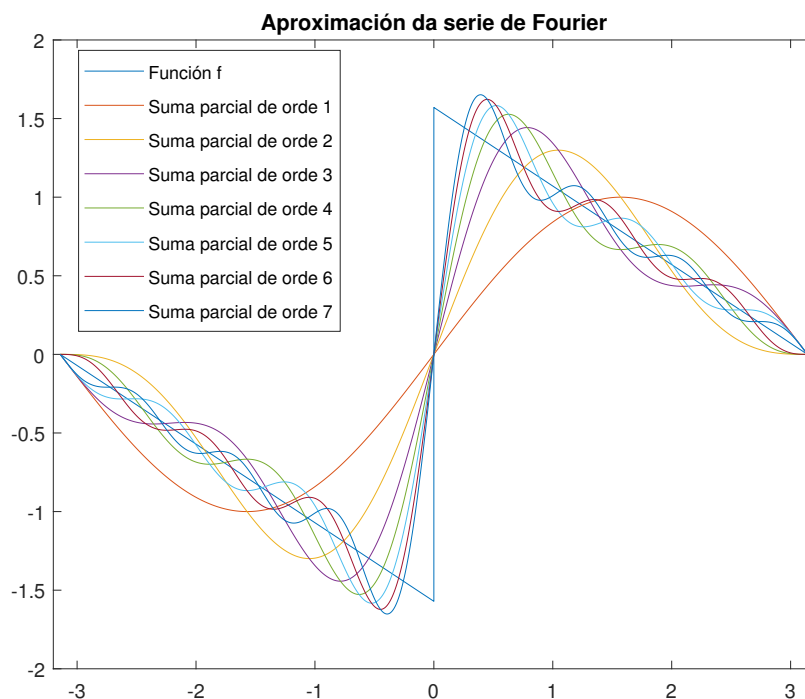


Figura 5.1: Sumas parciais.

Aproximamos agora as sumas de orde 50 e 100. Cabe esperar que a suma parcial converxa á función cando incrementamos o número de termos, pero veremos que isto ocorre en tódolos puntos salvo nas discontinuidades, cando a función pasa de $-\frac{\pi}{2}$ a $\frac{\pi}{2}$ en $t = 0$. Para visualizalo mellor, facemos nunha especie de “zoom” en torno ao punto $t = (0, \frac{\pi}{2})$ para ver o fenómeno do que estamos a falar. Queda ilustrado nas Figuras 5.2 e 5.3.

Observación 5.3. Neste exemplo atopamos o denominado *fenómeno de Gibbs*, un comportamento característico das series de Fourier asociadas a funcións definidas a cachos, con discontinuidades de salto finito e 2π -periódicas. Fixémonos en que no punto de discontinuidade representado, $t = 0$, a gráfica da suma parcial excede claramente á nosa función salto nese punto; este é o fenómeno de Gibbs.

A continuación, pasamos a estudar a velocidade de decrecemento dos coeficientes de Fourier. Para iso estudaremos as relacións que teñen os coeficientes entre si.

Podemos observar que:

$$\begin{aligned}
 c_1 &= -0.5000i, \\
 c_2 &= -0.2500i \cong \frac{c_1}{2}, \\
 c_3 &= -0.1667i \cong \frac{c_1}{3}, \\
 c_4 &= -0.1250i \cong \frac{c_1}{4}, \\
 c_5 &= -0.1000i \cong \frac{c_1}{5}, \\
 c_6 &= -0.0833i \cong \frac{c_1}{6}, \\
 c_7 &= -0.0714i \cong \frac{c_1}{7}, \\
 &\vdots
 \end{aligned} \tag{5.6}$$

concluíndo, por recursividade, que

$$c_k = \mathcal{O}\left(\frac{1}{k}\right). \tag{5.7}$$

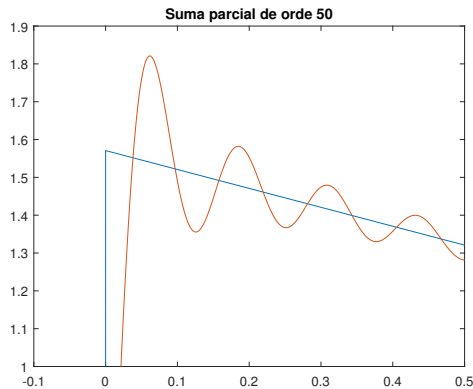


Figura 5.2: Soma parcial 50-ésima.

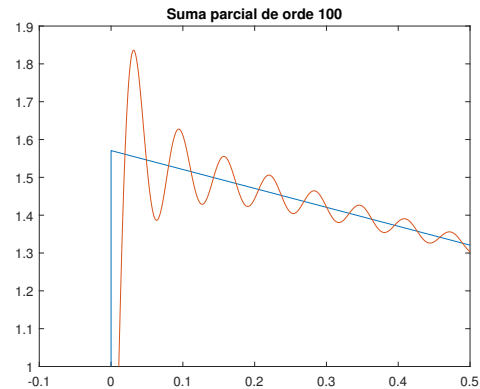


Figura 5.3: Soma parcial 100-ésima.

5.2. Aproximación dos coeficientes de Fourier dunha función continua pero non diferenciable

Consideramos unha función f tal que $f \in \mathcal{C}^0([-\pi, \pi])$ e $f \notin \mathcal{C}^1([-\pi, \pi])$. Por exemplo

$$f(t) = \begin{cases} \pi + t & \text{se } -\pi \leq t < 0, \\ \pi - t & \text{se } 0 \leq t \leq \pi, \end{cases} \tag{5.8}$$

e aproximamos os coeficientes de Fourier mediante a DFT vía FFT, tal e como fixemos no exemplo anterior.

Neste caso, empregamos un criterio de selección para o número de mostras do intervalo análogo ao do anterior exemplo, pero esixindo unha maior proximidade entre os coeficientes. Tomaremos como parámetro de tolerancia 10^{-10} e, coa mesma notación que a Observación 5.1, deberá verificarse que

$$\frac{\left| c_{\lfloor \frac{N}{2} \rfloor}^i - c_{\lfloor \frac{N}{2} \rfloor}^{i+1} \right|}{1 + \left| c_{\lfloor \frac{N}{2} \rfloor}^{i+1} \right|} < 10^{-10}. \quad (5.9)$$

Dividimos o intervalo $[-\pi, \pi]$ en $N = 2^{17}$ subintervalos da mesma lonxitude. Agora, tendo en conta que estamos a traballar cunha función que acada o mesmo valor en ambos extremos do intervalo, podemos calcular os coeficientes de Fourier mediante a aproximación

$$c_{k-1} \approx \frac{e^{ik\pi}}{N} \hat{y}_k, \quad k = 1, \dots, \left\lfloor \frac{N}{2} \right\rfloor + 1. \quad (5.10)$$

Usamos estes coeficientes para calcular as sumas parciais da serie de Fourier de f mediante a fórmula xa coñecida

$$c_0 + \sum_{k=1}^m [2\operatorname{Re}(c_k) \cos(kx) - 2\operatorname{Im}(c_k) \sin(kx)], \quad (5.11)$$

que quedan representadas na Figura 5.4.

Notamos que, para esta función, tódolos coeficientes c_k con índice par son nulos. Polo tanto, centraremos nos c_k con k impar para o estudo da velocidade de decrecemento dos mesmos. Así vemos que:

$$\begin{aligned} c_1 &= 0.6366, \\ c_3 &= 0.0707 \cong \frac{c_1}{3^2}, \\ c_5 &= 0.0255 \cong \frac{c_1}{5^2}, \\ c_7 &= 0.0130 \cong \frac{c_1}{7^2}, \\ c_9 &= 0.0079 \cong \frac{c_1}{9^2}, \\ c_{11} &= 0.0053 \cong \frac{c_1}{11^2}, \\ &\vdots \end{aligned} \quad (5.12)$$

polo tanto podemos deducir que

$$c_k = \mathcal{O}\left(\frac{1}{k^2}\right). \quad (5.13)$$

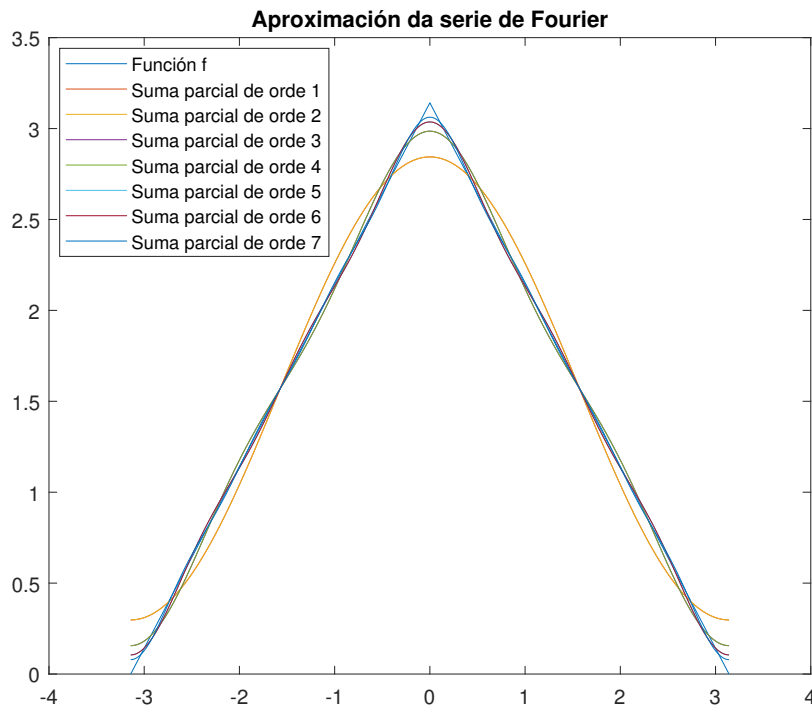


Figura 5.4: Sumas parciais.

Por outra banda, para a converxencia das sumas parciais impoñemos como criterio unha desigualdade análoga ao exemplo anterior, pero esta vez cunha menor diferenza entre os termos das sumas parciais involucradas. Así, diremos que a suma parcial de orde m pódese considerar converxente cando

$$\|s_m - s_{m-1}\| < 10^{-5}. \quad (5.14)$$

Baseándonos neste criterio, a nosa suma parcial acadou a converxencia para a suma parcial de orde 357. Representamos s_{357} na Figura 5.5, onde se pode observar unha boa aproximación da función f . Fixémonos en que esta aproximación parece moito mellor que a obtida para a función discontinua, incluso se comparamos os puntos de continuidade da función $f \notin \mathcal{C}^0([-\pi, \pi])$ co punto onde f non é diferenciable.

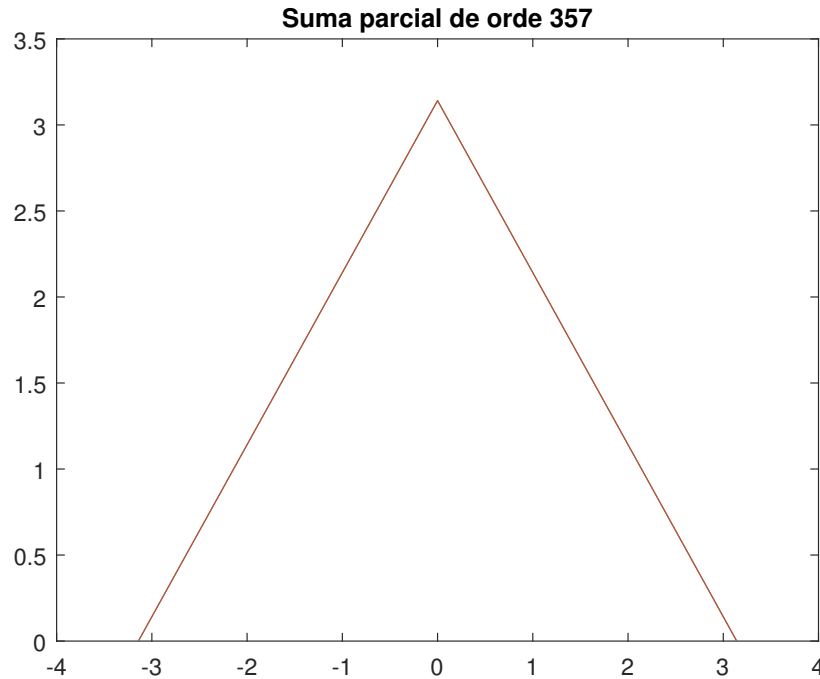


Figura 5.5: Aproximación da función f .

5.3. Aproximación dos coeficientes de Fourier dunha función de clase \mathcal{C}^1 pero non \mathcal{C}^2

Damos un paso máis na regularidade da función, e realizamos un proceso análogo ao da sección anterior pero esta vez considerando

$$f(t) = \begin{cases} t(t + \pi) & \text{se } -\pi \leq t < 0, \\ t(t - \pi) & \text{se } 0 \leq t \leq \pi. \end{cases} \quad (5.15)$$

Esta é unha función de clase $\mathcal{C}^1([-\pi, \pi])$, pero non o é de clase $\mathcal{C}^2([-\pi, \pi])$.

Empregando a DFT vía FFT calculamos os seus coeficientes de Fourier. Neste caso dividimos o intervalo $[-\pi, \pi]$ en $N = 2^{11}$ subintervalos da mesma lonxitude, elixindo esta cantidade do mesmo xeito que na sección anterior, baseándonos na condición (5.9).

Facendo uso da fórmula (4.43) xunto ca periodicidade da función no seu intervalo de definición, obtemos as aproximacións dos coeficientes de Fourier. Calculamos despois as sumas parciais ca expresión (4.30), que quedan representadas na Figura 5.6. Volvémonos atopar no caso de ter os coeficientes c_k , k par, con valor nulo. Estudemos entón aqueles coeficientes

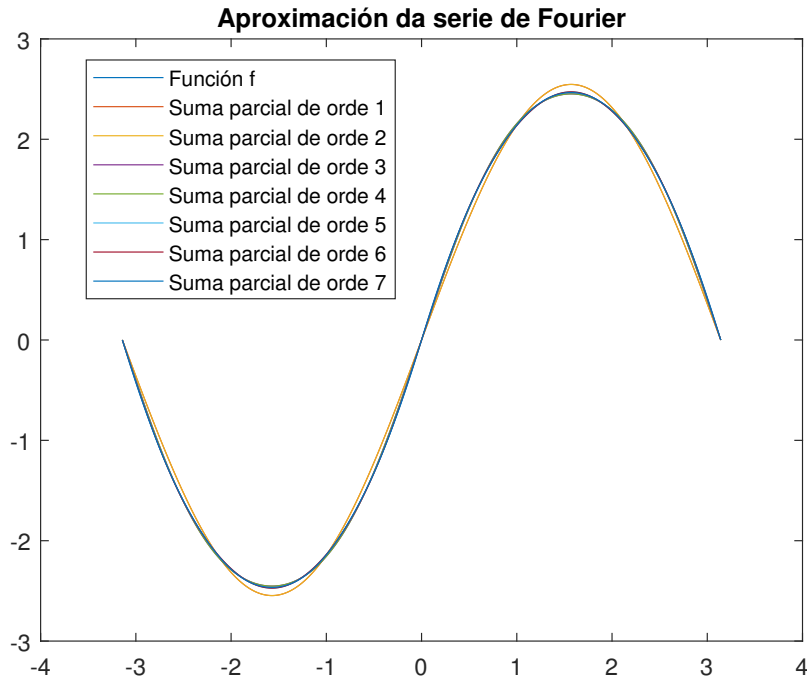


Figura 5.6: Somas parciais.

con índice impar. Observamos que:

$$\begin{aligned}
 c_1 &= -1.2732i, \\
 c_3 &= -0.0472i \cong \frac{c_1}{3^3}, \\
 c_5 &= -0.0102i \cong \frac{c_1}{5^3}, \\
 c_7 &= -0.0037i \cong \frac{c_1}{7^3}, \\
 c_9 &= -0.0017i \cong \frac{c_1}{9^3}, \\
 c_{11} &= -0.0010i \cong \frac{c_1}{11^3}. \\
 &\vdots
 \end{aligned} \tag{5.16}$$

Xa podemos ver como vai aumentando a velocidade de decrecemento dos coeficientes, concluindo entón que

$$c_k = \mathcal{O}\left(\frac{1}{k^3}\right). \tag{5.17}$$

Agora consideramos como criterio de converxencia a desigualdade

$$\|s_m - s_{m-1}\| < 10^{-7} \tag{5.18}$$

e vemos que as sumas parciais desta función converxen na suma parcial de orde 295, representada na Figura 5.7.

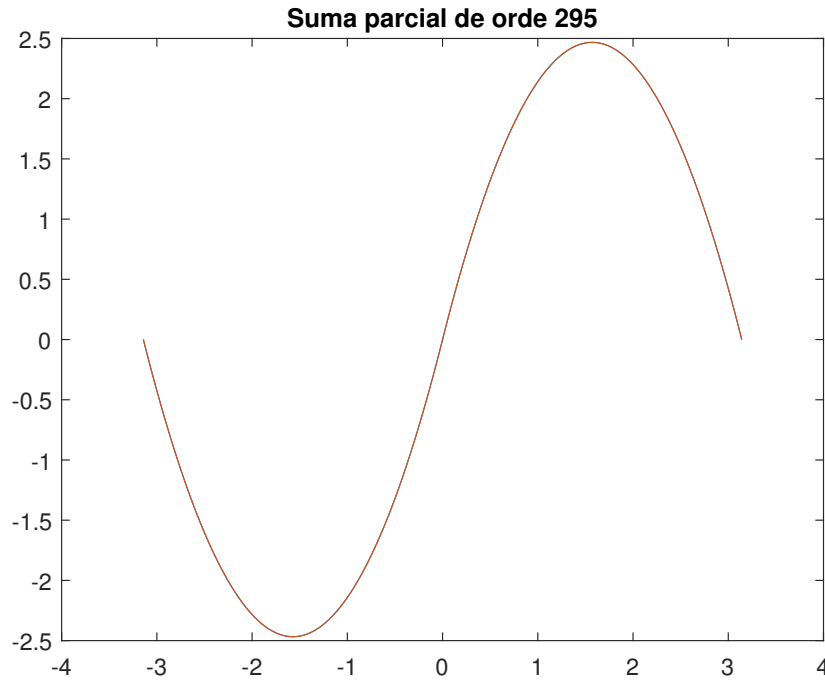


Figura 5.7: Aproximación da función f .

Queda evidenciado na gráfica a excelente aproximación que obtemos da función f en tódolos puntos. De feito, se calculamos o erro cometido ao aproximar a función f pola suma 295-ésima, temos que

$$\|f(t) - s_{295}\|_\infty = 5.5760 \times 10^{-6}, \quad \forall t \in [-\pi, \pi]. \quad (5.19)$$

5.4. Aproximación dos coeficientes de Fourier dunha función de clase \mathcal{C}^∞

Rematamos este estudo considerando unha función de clase $\mathcal{C}^\infty([-\pi, \pi])$, como pode ser $f(t) = e^{\cos(t)}$, no intervalo $[-\pi, \pi]$.

Calculamos unha aproximación dos coeficientes de Fourier co método empregado ata agora, e dividimos o intervalo de definición en $N = 2^{10}$ subintervalos de igual lonxitude, empregando a condición (5.9).

Dun xeito totalmente análogo ao da sección anterior, calculamos as sumas parciais de orde

m , que representamos na Figura 5.8. De novo, facemos un estudo sobre os coeficientes c_k para observar a súa velocidade de decrecemento. Atopámonos con que:

$$\begin{aligned}
 c_1 &= 0.5652, \\
 c_2 &= 0.1357 && \cong \frac{c_1}{4}, \\
 c_3 &= 0.0222 && \cong \frac{c_1}{25}, \\
 c_4 &= 0.0027 && \cong \frac{c_1}{209}, \\
 c_5 &= 0.0003 && \cong \frac{c_1}{1884}, \\
 c_6 &= 2.2489 \times 10^{-5} && \cong 0, \\
 &\vdots \\
 c_{10} &= 2.7529 \times 10^{-10} && \cong 0, \\
 &\vdots \\
 c_{15} &= 2.1125 \times 10^{-17} && \cong 0, \\
 &\vdots
 \end{aligned} \tag{5.20}$$

Nótese a rápida velocidade coa que decrecen os coeficientes.

En canto á converxencia, empregamos o mesmo criterio que os casos anteriores pero esta vez sendo un pouco máis esixentes, imponendo a desigualdade

$$\|s_m - s_{m-1}\| < 10^{-12}. \tag{5.21}$$

Con isto, as sumas parciais converxen xa na suma parcial de orde 14, a cal podemos ver na Figura 5.9.

É clara a precisión coa que queda representada a función de partida mediante a suma parcial calculada, exactamente

$$\|f(t) - s_{14}\| = 1.7764 \times 10^{-15}. \tag{5.22}$$

Observación 5.4. Resaltar que a rápida e boa aproximación que obtemos con este método é, en parte, grazas a que a función f coa que estamos a traballar está en condicións de garantir a converxencia superalxébrica da fórmula do trapecio composta, pois $f \in \mathcal{C}^\infty([-\pi, \pi])$ e cumpre a condición (2.34) (ver capítulo 2). Isto produce unha moi boa aproximación dos coeficientes, e polo tanto da serie de Fourier de f . Ademáis a regularidade da función provoca (como vimos no capítulo 3) que a converxencia da serie de Fourier, neste caso da súa aproximación, a f sexa moi rápida.

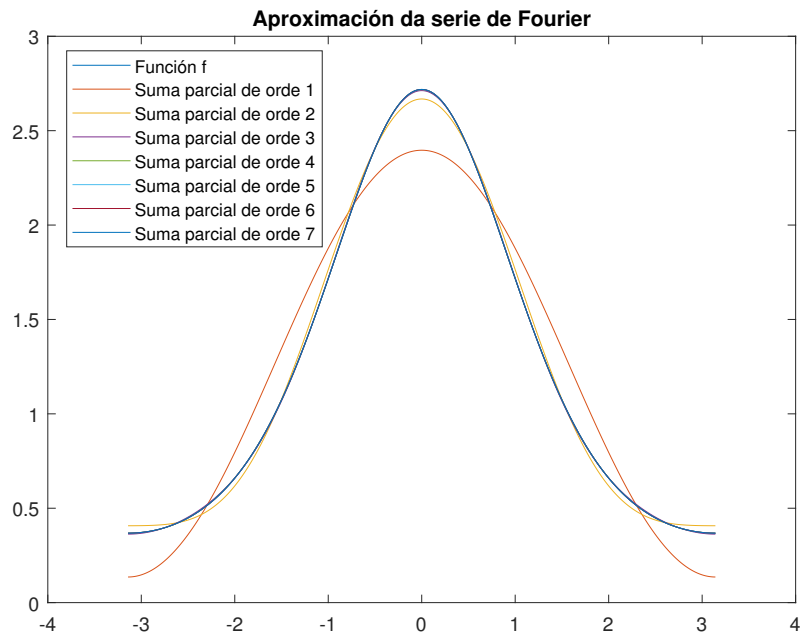
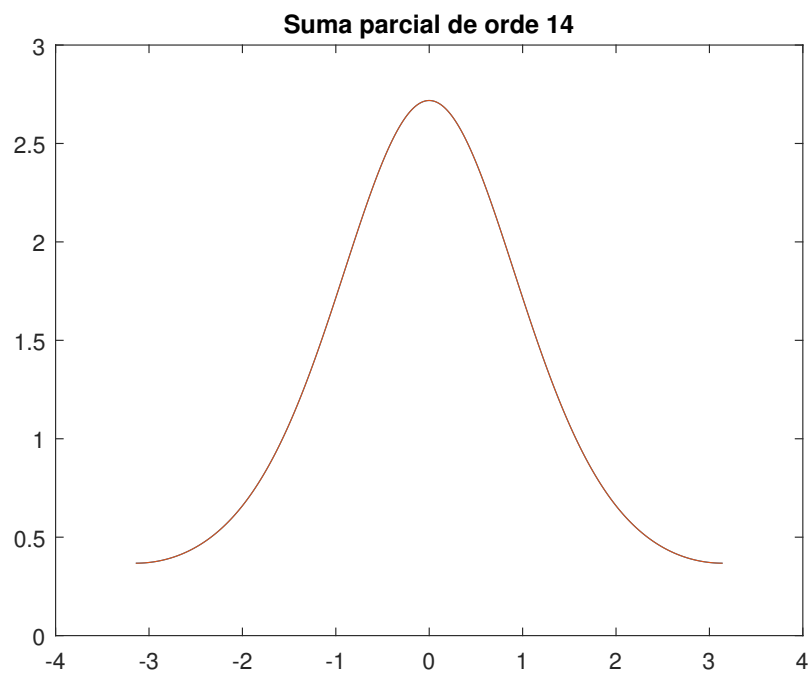


Figura 5.8: Sumas parciais.

Figura 5.9: Aproximación da función f .

5.5. Conclusións

En base aos exemplos das seccións anteriores, analizamos como o feito de aumentar a regularidade da función f inflúe no cálculo dos coeficientes de Fourier.

En primeiro lugar, sendo $N = 2^p$ o número de mostras do intervalo de definición das funcións, vemos que, a medida que aumenta a regularidade da función, é preciso un menor número de mostras para acadar unha mesma distancia (“converxencia”) entre os coeficientes calculados. En efecto, vimos que:

Se $f \notin \mathcal{C}^0([-\pi, \pi])$ temos que considerar $N = 2^{20} = 1048576$ subintervalos para obter unha diferenza entre as aproximacións dos coeficientes de 10^{-6} , mentres que se $f \in \mathcal{C}^0([-\pi, \pi])$ este número xa se reduce a $N = 2^{17} = 131072$, tendo agora unha diferenza entre coeficientes de 10^{-10} . Ca mesma tolerancia conseguimos para unha función $f \in \mathcal{C}^1([-\pi, \pi])$ reducir o valor ata $N = 2^{11} = 2048$, e finalmente coa máxima regularidade posible, $f \in \mathcal{C}^\infty([-\pi, \pi])$, bastaría tomar $N = 2^{10} = 1024$.

Polo tanto, ao aumentar a regularidade da función acadamos unha mellor aproximación cun número moito menor de mostras, e consecuentemente un menor custo.

En segundo lugar, o noso algoritmo realizou unha aproximación da serie de Fourier de f mediante as sumas parciais s_m , e intrinsecamente unha aproximación da propia función f . Pois ben, comprobamos que, a medida que aumenta a regularidade da función, a converxencia entre as sumas parciais dáse nunha orde cada vez menor, é dicir:

Para unha función $f \in \mathcal{C}^0([-\pi, \pi])$ teremos que chegar á suma parcial de orde 357 (isto é s_{357}) para obter unha aproximación da serie de Fourier cun erro de orde 10^{-5} , mentres que para unha función $f \in \mathcal{C}^1([-\pi, \pi])$ bastaría coa serie de Fourier de 295 sumandos, s_{295} , para acadar unha aproximación cun erro de orde 10^{-7} . Seguimos diminuindo a tolerancia no criterio de conxencia, e cunha función totalmente regular, $f \in \mathcal{C}^\infty([-\pi, \pi])$, bastaríanos con construír a serie de Fourier de tan só 14 sumandos, s_{14} para obter unha aproximación cun erro de orde 10^{-12} .

Polo tanto, ao aumentar a regularidade da función acadamos unha converxencia entre as sumas parciais moito máis rápida e eficaz, podendo así aproximar tanto a serie de Fourier de f coma a propia función por unha serie cun número de sumandos moi pequeno.

No mellor dos casos, se traballamos con funcións que cumpren a condición (2.34) e son de clase \mathcal{C}^∞ nun intervalo de lonxitude o seu período, podemos garantir a converxencia superalxébrica da fórmula do trapecio composta (base do algoritmo para o cálculo dos coeficientes). Isto concederíanos tódolos beneficios posibles, e seríamos capaces de: acadar

unha aproximación da serie de Fourier empregando un número moi pequeno de sumandos, cunha velocidade de converxencia dos coeficientes moi elevada. Derivando así nun resultado moi efectivo cun custo bastante reducido.

Bibliografía

- [1] Departamento de Matemática Aplicada, *Cálculo Numérico*. Universidade de Santiago de Compostela.
- [2] Tolstov, G. P. *Fourier Series*, Dover, New York, 1976.
- [3] Bogges, A. e Narcowich, F. J. (2009, segunda edición), *A First Course in Wavelets with Fourier Analysis*, Wiley, New Jersey.
- [4] Olver, S. S., *Numerical approximation of highly oscillatory integrals*, Doctoral thesis, University of Cambridge, 2008.