FACULTAD DE CIENCIAS

La regla trapezoidal: convergencia exponencial y aceleración mediante cambios de variable

(The trapezoidal rule: exponential convergence and acceleration by means of changes of variable)

Trabajo de fin de grado para acceder al

grado en Matemáticas

Autor: Jaime Garrido Aldea

Correo: jaime.garrido@alumnos.unican.es

Director: José Javier Segura Sala

13 de noviembre de 2024

Agradecimientos (Acknowledgements)

Quiero comenzar agradeciendo a mi supervisor, Javier Segura, por haberme apoyado y ayudado en todo momento durante la realización de este trabajo. Sé que hacer un seguimiento online de un trabajo de fin de grado es un extra de complicación. Sin embargo, como alumno, solo he vivido facilidades que han hecho mucho más ligera la tarea de acabar este trabajo.

Quiero agradecer a mis amigos Carlos y Lander por esperarme y aguantarme todas aquellas veces que he llegado algo tarde por estar escribiendo ésto y por ayudarme a desconectar todas las veces que lo he necesitado.

Agradezco a Eli el apoyarme durante estos meses, por ayudarme a abarcar trabajo, estudio, deporte y vida. Por animarme siempre que tuve la sensación de no poder alcanzar todo y por recordarme constantemente que "ya queda poco".

Finalmente, quiero dar las gracias a toda mi familia por apoyarme siempre, desde el principio de la universidad hasta su final, de forma totalmente incondicional.

Resumen

Esta texto explora la regla trapezoidal en integración numérica, poniendo especial énfasis en su convergencia exponencial en determinados casos y su uso junto a cambios de variable. Se comienza presentando la regla trapezoidal como una fórmula de Newton-Cotes para la estimación de integrales en intervalos cerrados $[a, b]$. A continuación, se introducen los polinomios de Bernoulli para llegar a la fórmula de Euler-Maclaurin, un resultado importante para el análisis del error de la regla trapezoidal. También se introduce la integración Romberg. Posteriormente, se examina el error de la regla trapezoidal en funciones periódicas y para integrales en la recta real, demostrando su buen comportamiento en estos casos. Finalmente, se introducen cambios de variable para adaptar cualquier integral a otra más adecuada para la regla trapezoidal. El texto concluye con una breve comparación entre la regla trapezoidal, la integración Romberg y el uso de cambios de variable.

Palabras clave: regla trapezoidal, integración numérica, polinomios de Bernoulli, fórmula de Euler-Maclaurin, integración de Romberg, convergencia exponencial, fórmulas doblemente exponenciales.

Abstract

This text explores the trapezoidal rule in numerical integration, with a focus on its exponential convergence in some cases and its usage via changes of variable. It first introduces the trapezoidal rule as a Newton-Cotes formula for estimating integrals over closed intervals $[a, b]$. Then it introduces the Bernouilli polynomials to arrive to the Euler-Maclaurin formula, an important result for analyzing the error of the trapezoidal rule. Romberg integration is introduced as well. Afterwards, the error of the trapezoidal rule is examined for periodic functions and integrals over the real line. The text ends up demonstrating the great accuracy of the trapezoidal rule in these two cases. Changes of variable are introduced to adapt any integral to a more convenient case for the trapezoidal rule. Finally, a brief comparison of the trapezoidal rule against Romberg and the trapezoidal rule with appropriate changes of variables is discussed, demonstrating the enhanced accuracy of the trapezoidal rule when applied to periodic and analytic functions.

Keywords: trapezoidal rule, numerical integration, Bernouilli polynomials, Euler-Maclaurin formula, Romberg integration, exponential convergence, doubly exponential formulas.

Contents

Some books are to be tasted, others to be swallowed, and some to be chewed and digested.

Sir Francis Bacon

Chapter 1

Introduction

Numerical integration is often introduced through the trapezoidal rule due to its simplicity. However, this simplicity can sometimes lead to the misconception that the trapezoidal rule lacks versatility for general applications.

This degree thesis aims to demonstrate that, in fact, the trapezoidal rule is an excellent choice when using equally spaced integration nodes, either directly or by employing changes of variables. Chapter 2 introduces the foundational concepts of the trapezoidal rule in compact intervals. Chapter 3 then presents the Euler-Maclaurin formula, which shows how the trapezoidal rule achieves remarkable accuracy for periodic functions over their periods. Chapter 4 rigorously proves the exponential convergence of the trapezoidal rule in two scenarios: periodic functions over their periods and functions over the real line, with detailed discussions on the analyticity requirements of these theorems. Chapter 5 explores how variable changes can transform certain integrals into forms that are highly compatible with the trapezoidal rule. Finally, Chapter 6 presents numerical examples that illustrate the theoretical results in practice.

Throughout the text, we have prioritized readability and accessibility, using fundamental or well-known results familiar to mathematics students. The appendix provides supporting material and relevant results from related areas to aid comprehension, with clear references throughout the text for ease of use.

With these guidelines, we hope this thesis provides an approachable and comprehensive exploration of the trapezoidal rule's effectiveness in numerical integration.

A designer knows he has achieved perfection not when there is nothing left to add, but when there is nothing left to take away.

Antoine de Saint-Exupry

Chapter 2

The trapezoidal rule in $[a, b]$ [1], [2], [3]

We shall introduce in this chapter the trapezoidal rule in compact intervals. We will start with the basic definitions and properties in order to introduce the Euler-Mclaurin formula in the next chapter.

The main goal to accomplish is to compute :

$$\int_a^b f(x)dx \tag{2.1}$$

In this case, the integration interval is compact, but the tools we will describe can be applied for infinite length intervals after appropriate adaptations. We will calculate the integral as a weighted sum of function evaluations over some points of the interval. We shall define this idea now [1, page 250], [3, page 124]:

Definition 2.1. ***(Quadrature rule)***. *Let $f(x)$ be an integrable function in $[a, b]$. Consider the approximation:*

$$I(f) = \int_a^b f(x)dx \approx \sum_{i=0}^{n} w_i f(x_i) \tag{2.2}$$

where $a \leq x_0 < x_1 < \ldots < x_n \leq b$. We call:

$$Q(f) = \sum_{i=0}^{n} w_i f(x_i) \tag{2.3}$$

A quadrature rule of $n+1$ points. The coefficients w_i are called weights and the points x_i nodes.

It is important to note that the nodes in definition 2.1 the extreme points of the interval are nodes as well. One can define quadrature rules without including these points, but in this text we will stick to the definition and include the extreme points.

Definition 2.2. ***(Truncation error)***. *Let $Q(f)$ be a quadrature rule. With the notation of definition 1, we call the truncation error the quantity:*

$$E(f) = I(f) - Q(f) \tag{2.4}$$

Definition 2.3. ***(Degree of exactness)*** *A natural number n is the degree of exactness of a quadrature rule if n is the greatest natural such that the quadrature rule yields an exact result for all polynomials of degree less or equal to n.*

2.1 The trapezoidal rule as a Newton-Cotes formula

The trapezoidal rule is just a particular example of the so called Newton-Cotes formulas. These formulas consist on replacing the function $f(x)$ on the integral by the n-th degree interpolating polynomial in $n+1$ points [1, page 263]. If we approximate $f(x)$ with a two points interpolation at a and b i.e $f(x) \approx P_1(x) = f(x_0)\frac{x-x_1}{x_0-x_1} + f(x_1)\frac{x-x_0}{x_1-x_0}$, we obtain [1, page 252]:

$$I(f) = \int_a^b f(x) \approx \int_a^b P_1(x) = \frac{h}{2}(f(a) + f(b)) \tag{2.5}$$

Which is the trapezoidal rule. The value $h = b - a$ is the so called step and in a general Newton-Cotes formula it corresponds to the distance between nodes.

Definition 2.4. ***(Trapezoidal rule)*** *Let $f(x)$ be an integrable function in $[a, b]$. Consider the approximation:*

$$I(f) = \int_a^b f(x)dx \approx \frac{h}{2}\left(f(a) + f(b)\right) \tag{2.6}$$

Where $h = b - a$ is the step. We call this quadrature rule the trapezoidal rule.

Consider a first degree polynomial. Its first degree polynomial is himself. Hence, it is obvious that by construction of the trapezoidal rule that its degree of exactness is at least one.

Another famous example of a Newton-Cotes formula comes from the interpolation with a degree-two polynomial. Using Lagarange interpolation [1, page 134] in $x_0 = a$, $x_1 = {}^{(a+b)}/_2$ and $x_2 = b$ and the distance between nodes as $h = {}^{(b-a)}/_2$ we have [1, page 256]:

$$\int_a^b f(x)dx \approx \frac{h}{3}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right) \tag{2.7}$$

This is the so called Simpson's rule

Definition 2.5. ***(Simpson's rule)*** *Let $f(x)$ be an integrable function in $[a, b]$. Consider the approximation:*

$$I(f) = \int_a^b f(x)dx \approx \frac{h}{3}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right) \tag{2.8}$$

Where $h = \frac{b-a}{2}$. We call this quadrature rule the Simpson's rule.

With the same rationale as before, it is obvious by construction of the Simpon's rule that its degree of exactness is at least 2. Nevertheless, it can be proven that in fact, its degree of exactness is 3. However, for brevity, we will not show it here.

More Newton-Cotes formulas can be built by increasing the degree of the interpolating polynomial. In this text we shall restrict ourselves to the trapezoidal rule.

For every numerical method it is always convenient, if possible, to have an estimation of the error with respect to the exact solution. For the trapezoidal rule, we can obtain a simple estimation by applying the mean value for integrals [4, page 189]. We include its proof in the main text because its reasoning is used several times in this text.

Theorem 2.1. ***(Mean value theorem for integrals)*** *Let $f(x)$ be a continuous function in $[a, b]$. Then, there is $c \in [a, b]$ such that:*

$$\int_a^b f(x)dx = f(c)(b - a) \tag{2.9}$$

Proof. Since f is a continuous function and $[a, b]$ is a compact set, then $f([a, b]) = [m, M]$, $m, M \in \mathbb{R}$ and $m < M$. This implies the following inequalities:

$$m(b - a) \leq \int_a^b f(x)dx \leq M(b - a) \tag{2.10}$$

$$m(b - a) \leq f(x)(b - a) \leq M(b - a)\ \forall x \in [a, b] \tag{2.11}$$

The fact that $f([a, b]) = [m, M]$ implies that $f(x)$ acquires every value in $[m, M]$ and hence, $f(x)(b - a)$ acquires all values in $[m(b - a), M(b - a)]$. The previous bound for the integral illustrates that its value must lie on the interval $[m(b - a), M(b - a)]$. As a result, there is a $c \in [a, b]$ such that:

$$\int_a^b f(x)dx = f(c)(b - a) \tag{2.12}$$

And the proof is complete. □

Now we have all the tools necessary to estimate the error of the trapezoidal rule [1, page 252], [2].

Theorem 2.2. ***(Error of the trapezoidal rule)*** *Let $f(x)$ be at least twice continuously differentiable in $[a,b]$. Then, there is $c \in [a,b]$ such that*

$$\int_a^b f(x)dx = \frac{h}{2}\left(f(a)+f(b)\right) - \frac{h^3}{12}f^{(2)}\left(c\right) \tag{2.13}$$

So the error term is:

$$E_{trap}(f) = -\frac{h^3}{12}f^{(2)}\left(c_1\right) \tag{2.14}$$

Proof. If we interpolate $f(x)$ with a degree 1 polynomial taking a and b as nodes, theorem 8.1 states that the interpolation error for is:

$$E_{interpolation}(x) = f(x) - \frac{(b-x)f(a)+(x-a)f(b)}{b-a} = (x-a)(x-b)f[a,b,x] = (x-a)(x-b)\frac{f^{(2)}(\eta_x)}{2} \tag{2.15}$$

Where $f[a,b,x]$ denotes the finite differences. We have that $f^{(2)}(\eta_x)$ is continuous as a function of x because $f[a,b,x] = \frac{f^{(2)}(\eta_x)}{2}$ and finite differences are continuous [2].

Since the trapezoidal rule is just the integral of the degree 1 interpolating polynomial, the error of the trapezoidal rule is then the integral of the interpolation error term:

$$E_{trap}(f) = \int_a^b (x-a)(x-b)\frac{f^{(2)}(\eta_x)}{2}dx \tag{2.16}$$

Now, since $\frac{f^{(2)}(\eta_x)}{2} = f[a,b,x]$ is continuous in $[a,b]$, there are $m, M \in \mathbb{R}$ such that $m \leq \frac{f^{(2)}(\eta_x)}{2} \leq M$ in $[a,b]$. And since $(x-a)(x-b)$ is a continuous function in $[a,b]$ and is lower or equal than zero in $[a,b]$, we have the two inequalities:

$$\begin{aligned} m\int_a^b (x-a)(x-b)dx \geq \int_a^b \frac{f^{(2)}(\eta_x)}{2}(x-a)(x-b)dx \geq M\int_a^b (x-a)(x-b)dx \\ m\int_a^b (x-a)(x-b)dx \geq \frac{f^{(2)}(\eta_x)}{2}\int_a^b (x-a)(x-b)dx \geq M\int_a^b (x-a)(x-b)dx \end{aligned} \tag{2.17}$$

These two inequalities and the fact that $\frac{f^{(2)}(\eta_x)}{2}$ is continuous in $[a,b]$ imply that there is $c_1 \in [a,b]$ such that:

$$\int_a^b \frac{f^{(2)}(\eta_x)}{2}(x-a)(x-b)dx = \frac{f^{(2)}(c_1)}{2}\int_a^b (x-a)(x-b)dx \tag{2.18}$$

And therefore, the error of the trapezoidal rule is:

$$E_{trap}(f) = \frac{f''\left(c_1\right)}{2}\int_a^b \left(x-a\right)\left(x-b\right)dx = -\frac{h^3}{12}f''\left(c_1\right) \tag{2.19}$$

□

Corollary 2.1. ***(Degree of exactness of the trapezoidal rule)*** *The degree of exactness of the trapezoidal rule is 1.*

Proof. By using the previous theorem. Take $f(x) = \alpha x + b$ as a degree 1 polynomial. $f(x)$ is twice continuously differentiable so that theorem 2.2 applies. Note that $f^{(2)}(x) = 0$. Plugging this into the expression of the error yields 0 for a degree 1 polynomial.

We shall check now that the trapezoidal rule is not exact for degree two polynomials. Let us take $f(x) = \alpha x^2 + \beta x + \gamma$ with nonzero α. We have that $f^{(2)} = 2\alpha$. Hence the error is $2h^3\alpha/12$ and the rule is not exact □

Quadrature rules with a higher number of nodes can give a better degree of exactness. For example, for the Simpson's rule:

Proposition 2.1. ***(Degree of exactness of the Simpon's rule)*** *The degree of exactness of the Simpon's rule is 3.*

Proof. The proof is not relevant at all for this text, so we do not write it here. The interested reader can see a proof in [1, page 257]. □

Since increasing the number of interpolation nodes leads to an improvement of the degree of exactness, one might think that the way to go in order to improve the numerical results is increasing the number of nodes of the interpolation. However, this approach does not guarantee a better approximation of the function when the nodes are equally spaced. This phenomenon receives the name *Runge phenomenon* [3, pages 54-55] (Check the appendix for an example). Hence, increasing the number of nodes, or equivalently, having a higher degree of exactness, does not guarantee a more accurate quadrature rule.

The theorem of the error of the trapezoidal rule, theorem 2.2, features an h^3 dependence of the error. Therefore, the smaller h the better. Nevertheless, in this section h has always been the length of the interval of integration, which is not always of a small length. The obvious approach to try to solve this issue is dividing the interval of integration in small subintervals where the trapezoidal rule is applied so that h is small in all of those subintervals. This approach is described in the following section. This way of improving the numerical integration is in fact the appropriate way to go instead of increasing the number of interpolation nodes. Additionally, we will see with the Romberg integration approach that, if desired, higher degree of exactness quadrature rules can be obtained automatically by a recursive usage of the quadrature rule.

2.2 Composite trapezoidal rule

In order to improve the precision of the integration with the trapezoidal rule, let us start by dividing $[a,b]$ in subintervals. To do so, we take $n+1$ points:

$$a = x_0 < x_1 < \ldots < x_n = b \quad ; \quad x_i = x_0 + ih, i = 0, \ldots, n \quad ; \quad h = (b-a)/n \tag{2.20}$$

This way we obtain the subintervals $[x_i, x_i + 1]$ where the trapezoidal rule can be applied:

$$\int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{h}{2}\left(f(x_i + f(x_{i+1})\right) \tag{2.21}$$

Doing this for the full integral:

$$\int_a^b f(x)dx = \int_{x_0}^{x_n} f(x)dx = \sum_{i=0}^{n-1} \int_{x_i}^{x_{i+1}} f(x)dx \approx \frac{h}{2}\left(f_0 + f_n\right) + h\sum_{i=1}^{n-1} f_i \tag{2.22}$$

Where we are writing $f_i = f(x_i)$. We shall write this in its own definition [1, page 253], [2], [3, page 126].

Definition 2.6. ***(Composite trapezoidal rule)*** *Let $f(x)$ be an integrable function in $[a,b]$. Consider the approximation:*

$$\int_a^b f(x)dx = \approx \frac{h}{2}\left(f_0 + f_n\right) + h\sum_{i=1}^{n-1} f_i \tag{2.23}$$

Where $x_i = x_0 + ih$; $i = 0, \ldots, n$; $x_0 = a$, $x_n = b$; $f_i = f(x_i)$ and $h = {}^{b-a}/_{n}$. We call this quadrature rule the composite trapezoidal rule and we denote:

$$T_n(f) = \frac{h}{2}\left(f_0 + f_n\right) + h\sum_{i=1}^{n-1} f_i \tag{2.24}$$

The rest of this section is devoted to the properties of the composite trapezoidal rule. We will maintain the notation $f_i = f(x_i)$ without further explanation. We can jump directly to the error of the composite trapezoidal rule [1, page 253], [2], [3, page 127].

Theorem 2.3. ***(Error of the composite trapezoidal rule)*** *Let $f(x)$ be at least twice continuously differentiable in $[a, b]$. For the composite trapezoidal rule with n subintervals, $T_n(f)$, we have:*

$$\int_a^b f(x)dx = T_n(f) - \frac{(b-a)h^2}{12} f^{(2)}(\tau) \tag{2.25}$$

With $\tau \in [a, b]$. We note the error of $T_n(f)$ as:

$$E_n^T(f) = -\frac{(b-a)h^2}{12} f^{(2)}(\tau) \tag{2.26}$$

Proof. By invoking the theorem 2.2 for the error of the trapezoidal rule, we have:

$$\int_{x_i}^{x_{i+1}} f(x)dx = \frac{h}{2}\left(f_{i-1} + f_{i+1}\right) - \frac{h^3}{12} f^{(2)}(c_i) \tag{2.27}$$

With $c_i \in [x_i, x_{i+1}]$. Hence, the total error of the integral over $[a, b]$ is just the sum of all the errors of the integrals over the subintervals:

$$E_n^T(f) = -\frac{h^3}{12} \sum_{i=0}^{n} f^{(2)}(c_i) \tag{2.28}$$

Now, note that:

$$m = \min_{x\in[a,b]} nf^{(2)}(x) \leq \sum_{i=0}^{n} f^{(2)}(c_i) \leq \max_{x\in[a,b]} nf^{(2)}(x) = M \tag{2.29}$$

Since $nf^{(2)}(x)$ is a continuous function in $[a, b]$, it reaches all values between m and M. Hence, there is $\tau \in [a, b]$ such that $nf^{(2)}(\tau) = \sum_{i=0}^{n} f^{(2)}(c_i)$. If we plug this last result into the last expression for the error $E_n^T(f)$, we get:

$$E_n^T(f) = -\frac{h^3}{12} nf^{(2)}(\tau) = -\frac{h^2(b-a)}{12} f^{(2)}(\tau) \tag{2.30}$$

And we get the desired result. □

The estimation of the error in the limit $h \longrightarrow 0$ will have important consequences and it will provide a first hint on a more general result we will prove later (The Euler-Maclaurin formula) . This limit can be analyzed as a corollary of the previous theorem [1, page 254], [3, page 128]:

Corollary 2.2. ***(Asymptotic estimation of the error of the composite trapezoidal rule)*** *Following the notation of the previous theorem, we have:*

$$E_n^T(f) = -\frac{h^3}{12} \sum_{i=0}^{n} f^{(2)}(c_i) \tag{2.31}$$

Let us write

$$\widehat{E}_n^T(f) = -\frac{(b-a)^2}{12n^2}\left(f'(b) - f'(a)\right) = -\frac{h^2}{12}\left(f'(b) - f'(a)\right) \tag{2.32}$$

Then, we have:

$$\lim_{n\to\infty} \frac{\widehat{E}_n^T(f)}{E_n^T(f)} = 1 \tag{2.33}$$

Which means that $\widehat{E}_n^T(f)$ is the asymptotic estimation of the error of the composite trapezoidal rule.

Proof. We proceed by direct computation:

$$\lim_{n\to\infty}\frac{\widehat{E}_n^T(f)}{E_n^T(f)}=\lim_{n\to\infty}\frac{-\frac{(b-a)^2}{12n^2}\left(f'(b)-f'(a)\right)}{-\frac{(b-a)^3}{12n^3}\sum_{i=0}^{n}f^{(2)}(c_i)}=\lim_{n\to\infty}\frac{f'(b)-f'(a)}{\frac{(b-a)}{n}\sum_{i=0}^{n}f^{(2)}(c_i)}=\\ =\frac{f'(b)-f'(a)}{\int_a^b f^{(2)}(x)dx}=1 \tag{2.34}$$

And we have used the fact that $\lim_{n\to\infty}\frac{b-a}{n}\sum_{i=0}^{n}f^{(2)}(c_i)$ is the Riemann integral $\int_a^b f^{(2)}(x)dx$. □

The asymptotic behaviour of the error in the composite trapezoidal rule shows how the error will decrease more rapidly when $f'(b)=f'(a)$. We will see in chapter 3 that the convergence of the integral is even better when $f^{(n)}(b)=f^{(n)}(a)$, the higher the n, the better. This makes the trapezoidal rule especially suitable to integrate periodic functions over their period.

We finish this section with a definition that we write for the sake of completeness.

Definition 2.7. ***(Order of a cuadrature rule)*** *The exponent of h in the value of the error of a composite Newton-Cotes quadrature rule receives the name of order of the quadrature rule.*

With the previous definition, the order of the composite trapezoidal rule is 2 according to theorem 2.3. The order of the composite Simpson's rule is 4. The interested reader can find a proof in [1, page 258].

2.3 Sequential evaluation of the integral

The trapezoidal rule has an interesting property: if one wants to reduce the step to $h/2$, the new trapezoidal calculation does not need to double the number of function evaluations since some of the evaluations of the quadrature with step h can be reused [3, page 129]. Let us write this explicitly. The trapezoidal rule with step h yields:

$$\int_a^b f(x)dx\approx T(f,h)=\frac{h}{2}\sum_{i=0}^{n-1}(f_i+f_{i+1}) \tag{2.35}$$

The trapezoidal rule with step $2h$ yields:

$$\int_a^b f(x)dx\approx T(f,2h)=h\sum_{i=0}^{2n-1}(f_{2i}+f_{2i+2}) \tag{2.36}$$

As a result, we can relate both quadratures:

$$T(f,h)=\frac{T(f,2h)}{2}+h\sum_{i=1}^{2n-1}f_{2i-1} \tag{2.37}$$

With this rationale, in order to reduce the step to a half, we only need to evaluate the function in new nodes while we divide by two the value of the previous quadrature.

This philosophy is generalized with the Romberg integration approach that we will introduce in the next chapter.

To achieve great things, two things are needed: a plan, and not quite enough time.

Leonard Bernstein

Chapter 3

The Euler-Maclaurin formula and Romberg integration [2], [6]

In this chapter we will introduce the Euler-Maclaurin formula, a key result for the trapezoidal rule that extends corollary 2.2. As a consequence of this formula, we will develop the Romberg integration framework.

3.1 The Euler-Maclaurin formula

The Euler-Maclaurin formula is obtained exploiting the properties of the Bernouilli polynomials [6, page 136],[8, page 127],[1, page 284]. We shall start by introducing them and some useful properties.

Definition 3.1. ***(Definition of the Bernouilli polynomials)***. *Consider the following expansion:*

$$\frac{t(e^{xt}-1)}{e^t-1} = \sum_{k=0}^{\infty} B_k(x)\frac{t^k}{k!} \tag{3.1}$$

The polynomials $B_k(x)$ obtained with the previous expansion are called Bernoulli polynomials of degree k.

The explicit expressions for the Bernouilli polynomials can be obtained by expanding the left hand side of the equality (3.1). The function on the left side of (3.1) is analytic as a function of t in all points of the complex plane except in those in which $e^t = 1$. Hence, the function admits a Laurent series (theorem 8.4). Since the pole at $t = 0$ is a removable singularity, all the terms of the Laurent series a_{-k} with $k > 0$ are zero and we can focus on those a_k , $k \geq 0$ (Using corollary 8.2). Using theorem 8.4 with $\gamma = \mathcal{C}(0,1)$:

$$a_0 = \frac{1}{2\pi i}\oint_\gamma \frac{t(e^{xt}-1)}{(e^t-1)t}dt = \frac{1}{2\pi i}\oint_\gamma \frac{e^{xt}-1}{e^t-1}dt \overset{\text{Theorem 8.5}}{=} 0 \tag{3.2}$$

$$a_1 = \frac{1}{2\pi i}\oint_\gamma \frac{t(e^{xt}-1)}{(e^t-1)t^2}dt = \frac{1}{2\pi i}\oint_\gamma \frac{e^{xt}-1}{(e^t-1)t}dt \overset{\text{Theorem 8.5}}{=} x \tag{3.3}$$

$$a_2 = \frac{1}{2\pi i}\oint_\gamma \frac{t(e^{xt}-1)}{(e^t-1)t^3}dt = \frac{1}{2\pi i}\oint_\gamma \frac{e^{xt}-1}{(e^t-1)t^2}dt \tag{3.4}$$

The function $f(t) = (e^{xt}-1)/(e^t-1)t^2$ has an order 2 pole at $t = 0$. Using corollary 8.2, the last nonzero term a_{-k} with $k > 0$ is a_{-2}, which goes with $1/t^2$. Therefore, if we multiply the function by t^2, the residue a_{-1} goes with t and hence, we can calculate the residue of the function at $t = 0$ as:

$$Res(f;0) = \lim_{t\to 0}\frac{d}{dt}\left(\frac{e^{xt}-1}{e^t-1}\right) = x^2 - x \tag{3.5}$$

As a result, the first three Bernouilli polynomials are:

$$\begin{aligned} B_0(x) &= 0 \\ B_1(x) &= x \\ B_2(x) &= x^2 - x \end{aligned} \tag{3.6}$$

Definition 3.2. ***(Definition of the Bernouilli numbers)*** *Consider the expansion*

$$\frac{t}{e^t - 1} = \sum_{k=0}^{\infty} B_k \frac{t^k}{k!} \tag{3.7}$$

The constants B_k obtained by the previous expansion are called Bernouilli numbers.

Remark 3.1. *Some other sources such as [3, page 332] or [9] define the Bernouilli polynomials with a slightly different generating function:*

$$\frac{te^{tx}}{e^t - 1} = \sum_{k=0}^{\infty} \mathcal{B}_k(x) \frac{t^k}{k!} \tag{3.8}$$

Where the $\mathcal{B}_k$ are the Bernouilli polynomials. With this alternative definition, the Bernouilli numbers $\mathcal{B}_k$ are simply $\mathcal{B}_k(1)$. Both definitions of the Bernouilli polynomials are related as [9] $B_k(x) = \mathcal{B}_k(x) - \mathcal{B}_k$. In this text we will stick to definition 3.1 since it is convenient to proof the Euler-Mclaurin formula.

Proposition 3.1. ***(Properties of the Bernouilli polynomials)*** *The Bernouilli polynomials satisfy the following properties:*

1. $B_k(0) = 0 \ \forall k \geq 0$.
2. $B_k(1) = 0 \ \forall k \neq 1$.
3. $B_{2k+1} = 0 \ \forall k > 0$
4. $B'_{2k}(x) = 2kB_{2k-1}(x) \ \forall k > 1$.
5. $B'_{2k+1}(x) = (2k+1)(B_{2k}(x) + B_{2k}) \ \forall k \geq 0$.
6. $B_k = -\int_0^1 B_j(x)dx \ , j \geq 1$

Proof. We shall proceed property by property:

1. For $x = 0$, using (3.1) we have:

$$0 = \sum_{k=0}^{\infty} B_k(0) \frac{t^k}{k!} \ \forall t \in \mathbb{R} \tag{3.9}$$

   As a result: $B_k(0) = 0 \ \forall k \geq 0$

2. For $x = 1$, using (3.1) we have:

$$\begin{aligned} t = \sum_{k=0}^{\infty} B_k(1) \frac{t^k}{k!} \ \forall t \in \mathbb{R} &\Rightarrow B_0(1) + B_1(1)t \sum_{k=2}^{\infty} B_k(1) \frac{t^k}{k!} - t = 0 \\ &\Rightarrow B_0(1) + t(B_1(1) - 1) + \sum_{k=2}^{\infty} B_k(1) \frac{t^k}{k!} = 0 \ \forall t \in \mathbb{R} \end{aligned} \tag{3.10}$$

   Since the equality holds $\forall t \in \mathbb{R}$, all t powers must vanish, yielding:

$$B_1(1) = 1 \quad B_k(1) = 0 \ \forall k \neq 1 \tag{3.11}$$

3. Let us name $f(t)$ the generating function in (3.7). Then, we have:

$$f(-t) - f(t) = \frac{-t}{e^{-t}-1} - \frac{t}{e^t-1} = \frac{te^t}{e^t-1} - \frac{t}{e^t-1} = \frac{t(e^t-1)}{e^t-1} \tag{3.12}$$

Using the definition of the Bernouilli numbers (3.7) and the Bernouilli polynomials (3.1), the previous equality translates into:

$$\sum_{k=0}^{\infty} B_k \frac{(-t)^k}{k!} - \sum_{k=0}^{\infty} B_k \frac{t^k}{k!} = \sum_{k=0}^{\infty} B_k(1) \frac{t^k}{k!} \tag{3.13}$$

Now, we can consider those terms with $k = 2m+1,\ m > 0$:

$$\sum_{m=0}^{\infty} B_{2m+1} \frac{(-t)^{2m+1}}{(2m+1)!} - \sum_{m=0}^{\infty} B_{2m+1} \frac{t^{2m+1}}{(2m+1)!} = \sum_{m=0}^{\infty} B_{2m+1}(1) \frac{t^{2m+1}}{(2m+1)!} \tag{3.14}$$

$$-\sum_{m=0}^{\infty} B_{2m+1} \frac{t^{2m+1}}{(2m+1)!} - \sum_{m=0}^{\infty} B_{2m+1} \frac{t^{2m+1}}{(2m+1)!} = \sum_{m=0}^{\infty} B_{2m+1}(1) \frac{t^{2m+1}}{(2m+1)!} \tag{3.15}$$

The last equality implies that both sides must be equal term by term for each power:

$$-B_{2m+1} - B_{2m+1} = B_{2m+1}(1) = (\text{Using property 2}) = 0 \Rightarrow B_{2m+1} = 0\ \forall m > 0 \tag{3.16}$$

4. We will prove properties 4 and 5 at the same time. Firstly, let us write the generating function of the Bernouilli polynomials (3.1) as $f(x,t) = {}^{t(e^{xt}-1)}/_{e^t-1}$. Then, we have:

$$\partial_x f(x,t) = \frac{t^2 e^{xt}}{e^t-1} = \sum_{k=0} B_k'(x) \frac{t^k}{k!} \tag{3.17}$$

Since $B_k(x) = 0$ (3.6), we can conveniently write the previous equality as:

$$\partial_x f(x,t) = \frac{t^2 e^{xt}}{e^t-1} = \sum_{k=1} B_k'(x) \frac{t^k}{k!} \tag{3.18}$$

Using (3.1) and (3.7), we have the following equalities:

$$\partial_x f(x,t) - \frac{t^2}{e^t-1} = \frac{t^2(e^{xt}-1)}{e^t-1} = \sum_{k=0}^{\infty} B_k(x) \frac{t^{k+1}}{k!} = \sum_{k=1}^{\infty} B_k'(x) \frac{t^k}{k!} - \sum_{k=0}^{\infty} B_k \frac{t^{k+1}}{k!} \tag{3.19}$$

$$\sum_{k=0}^{\infty} B_k(x) \frac{t^{k+1}}{k!} = \sum_{k=0}^{\infty} B_{k+1}'(x) \frac{t^{k+1}}{(k+1)!} - \sum_{k=0}^{\infty} B_k \frac{t^{k+1}}{k!} \tag{3.20}$$

This last equality implies that both sides must be equal for each power of t. Thus:

$$\frac{B_k(x)}{k!} = \frac{B_{k+1}'(x)}{(k+1)!} - \frac{B_k}{k!} \xrightarrow{\text{Solving for } B_{k+1}'(x)} B_{k+1}'(x) = (k+1)(B_k(x) + B_k) \tag{3.21}$$

If we consider the case of odd k i.e $k = 2m-1\ , m > 1$:

$$B_{2m+1}'(x) = 2m(B_{2m-1(x)} + B_{2m-1}) \xrightarrow{\text{Using property 3}} B_{2m+1}'(x) = 2mB_{2m-1(x)}\ , \forall m > 1 \tag{3.22}$$

If we consider even k i.e $k = 2m\ , m \geq 0$:

$$B_{2m+1}'(x) = (2m+1)(B_{2m}(x) + B_{2m})\ , \forall m \geq 0 \tag{3.23}$$

5. Proven in the previous step.

6. By integrating equation (3.1) on x:

$$\int_0^1 \frac{t(e^{xt}-1)}{e^t-1}dx = 1 - \frac{t}{e^t-1} \overset{(3.7)}{=} 1 - \sum_{k=0}^{\infty} B_k \frac{t^k}{k!} = \int_0^1 \sum_{k=0}^{\infty} \frac{t^k}{k!} B_k(x)dx = \sum_{k=0}^{\infty} \frac{t^k}{k!} \int_0^1 B_k(x)dx \quad (3.24)$$

Where the last equality follows from using theorems 8.12 and 8.13. The function $g(x) = \sum_{k=0}^{\infty} |B_k(x)| \frac{|t|^k}{k!}$ dominates the partial sums $f_m(x) = \sum_{k=0}^{m} B_k(x) \frac{t^k}{k!}$. The function $g(x)$ is integrable as a consequence of the monotone convergence theorem for integrals 8.12 applied to the non-decreasing sequence of non-negative measurable functions $g_m(x) = \sum_{k=0}^{m} |B_k(x)| \frac{|t|^k}{k!}$. $g_m(x)$ is measurable since it is a finite sum of polynomials, which are all measurable. Thus, the previous equality is fully justified. Moving terms in the previous equality, we get:

$$1 = \sum_{k=0}^{\infty} \frac{t^k}{k!} \left(\int_0^1 B_k(x)dx + B_k \right) \quad (3.25)$$

For the equality to be valid, all coefficients on the right hand side with $k \geq 1$ must vanish, yielding:

$$B_k = -\int_0^1 B_k(x)dx \ , \forall k \geq 1 \quad (3.26)$$

And no we have proved all the properties. □

With all these properties we have the necessary tools to prove one of the most important results about the trapezoidal rule: the Euler-Mclaurin formula [6, pages 137-140], [1, page 285].

Theorem 3.1. ***(The Euler-Maclaurin formula)*** *Let $f \in \mathcal{C}^{2m+2}[a,b]$, $m \in \mathbb{N}$. Let us write $h = {}^{(b-a)}/_{n}$ and $x_j = a + jh$, $j = 0, 1, \ldots, n$. With the notation of definition 3.7, we have the following equality:*

$$\int_a^b f(x)dx = h\sum_{j=0}^{n} f(a+jh) - \frac{h}{2}\left[f(b)+f(a)\right] - \sum_{k=1}^{m} \frac{B_{2k}}{(2k)!} h^{2k} \left[f^{(2k-1)}(b) - f^{(2k-1)}(a) \right] + E_m \quad (3.27)$$

Where the error term is given by:

$$E_m = \frac{h^{2m+2}}{(2m+2)!} \int_a^b \overline{B}_{2m+2}\left(\frac{x-a}{h}\right) f^{(2m+2)}(x)dx \quad (3.28)$$

And $\overline{B}_j(x)$ is the periodic extension of the Bernoulli polynomials:

$$\overline{B}_j(x) = \begin{cases} B_j(x) & \text{if } 0 \leq x < 1 \\ \overline{B}_j(x-1) & \text{if } x \geq 1 \end{cases} \quad (3.29)$$

Proof. The proof of this results is long and it relies a lot in the properties of proposition 3.1. We also aim to discuss the rationale of the proof so that the reader can understand what ideas motivated the big picture of the proof. The proof written in this text is a completed version of that given in [6, pages 137-140].

Firstly, let us begin with the main idea to start the proof. Properties 4 and 5 of proposition 3.1 relate the derivatives of the Bernoulli polynomials to other Bernouilli polynomials and the Bernouilli numbers. These properties can be exploited when integrating by parts a function g times a Bernouilli polynomial. The integration by parts will be straightforward if $g(x) = f^{(k)}(x)$ for a certain f and $k \in \mathbb{N}$. Moreover, the limits of the integral are such that we evaluate $B_k(0)$ and $B_k(1)$, we will also be able to exploit properties 1 and 2 of 3.1. Using these ideas, it feels "natural" to try to compute the following integral:

$$\begin{aligned}
X_1 &:= \frac{1}{2}\int_0^h B_2\left(\frac{y}{h}\right) f''(x_0+y)dy \overset{(3.6)}{=} \frac{1}{2}\int_0^h \left(\frac{y^2}{h^2}-\frac{y}{h}\right) f''(x_0+y)dy \overset{\text{By parts}}{=} -\int_0^h \left(\frac{y}{h^2}-\frac{1}{2h}\right) f'(x_0+y)dy \\
&\overset{\text{By parts}}{=} -\frac{1}{2h}[f(x_0+h)+f(x_0)] + \frac{1}{h^2}\int_0^h f(x_0+y)dy \\
&\overset{x_0+y=z}{=} -\frac{1}{2h}[f(x_0+h)+f(x_0)] + \frac{1}{h^2}\int_{x_0}^{x_0+h} f(z)dz
\end{aligned} \tag{3.30}$$

The first term on the right-hand side of the last equality is just the trapezoidal rule applied to f on the interval $[x_0, x_0+h]$. Hence, the previous equalities relate the trapezoidal rule, the actual integral and an integral of the second derivative of f times a Bernouilli polynomial. The next question to ask ourselves is: may a similar relation appear for higher derivatives of f?. To check that, we shall define for $k > 1$:

$$X_k := \frac{1}{(2k)!}\int_0^h B_{2k}\left(\frac{y}{h}\right) f^{(2k)}(x_0+y)dy \tag{3.31}$$

We take only even derivatives and even degree Bernouilli polynomials for the sake of conveniency as we will see later. Now, we integrate by parts several times (3.31):

$$\begin{aligned}
X_k &:= \frac{1}{(2k)!}\int_0^h B_{2k}\left(\frac{y}{h}\right) f^{(2k)}(x_0+y)dy \overset{\text{By parts and property 4}}{=} \\
&= \frac{1}{(2k)!}\left[\underbrace{B_{2k}\left(\frac{y}{h}\right) f^{(2k-1)}(x_0+y)\Big|_0^h}_{\text{Vanishes due to properties 1 and 2}} - \int_0^h f^{(2k-1)}(x_0+y)\frac{2k}{h}B_{2k-1}\left(\frac{y}{h}\right)dy\right] = \\
&= -\frac{1}{h(2k-1)!}\int_0^h f^{(2k-1)}(x_0+y)B_{2k-1}\left(\frac{y}{h}\right)dy \overset{\text{By parts and property 5}}{=} \\
&= -\frac{1}{h(2k-1)!}\left[\underbrace{B_{2k-1}\left(\frac{y}{h}\right) f^{(2k-2)}(x_0+y)\Big|_0^h}_{\text{Vanishes due to properties 1 and 2}} - \int_0^h f^{(2k-2)}(x_0+y)\frac{2k-1}{h}\left(B_{2k-2}\left(\frac{y}{h}\right)+B_{2k-2}\right)dy\right] = \\
&= \frac{1}{h^2(2k-2)!}\int_0^h f^{(2k-2)}(x_0+y)B_{2k-2}\left(\frac{y}{h}\right)dy + \frac{B_{2k-2}}{h^2(2k-2)!}\left[f^{(2k-3)}(x_0+h)-f^{(2k-3)}(x_0)\right] \overset{X_k \text{ definition (3.31)}}{=} \\
&= \frac{X_{k-1}}{h^2} + \frac{B_{2k-2}}{h^2(2k-2)!}\left[f^{(2k-3)}(x_0+h)-f^{(2k-3)}(x_0)\right]
\end{aligned} \tag{3.32}$$

The last equality yields the following relation:

$$h^2X_k - X_{k-1} = \frac{B_{2k-2}}{(2k-2)!}\left[f^{(2k-3)}(x_0+h)-f^{(2k-3)}(x_0)\right] \tag{3.33}$$

So in the end we have not obtained an expression relating the trapezoidal rule, the actual integral and the integral of a high derivative of f times a Bernouilli polynomial. Nevertheless, the last equality is very useful since we can obtain all X_k starting from X_1. Moreover, the expression relates X_k and X_{k-1} via an easy to compute expression that does not involve integrals. Then, by using (3.30) and (3.33) and the hypothesis that $f \in \mathcal{C}^{2m+2}[a,b]$, $m \in \mathbb{N}$:

$$\begin{aligned}\frac{1}{2}\Big[f(x_0+h)+f(x_0)\Big] =& \frac{1}{h}\int_{x_0}^{x_0+h} f(y)dy - hX_1 = \frac{1}{h}\int_{x_0}^{x_0+h} f(y)dy + \frac{hB_2}{2!}\left[f'(x_0+h)-f'(x_0)\right] - h^3X_2 \\ =& \frac{1}{h}\int_{x_0}^{x_0+h} f(y)dy + \frac{hB_2}{2!}\left[f'(x_0+h)-f'(x_0)\right] + \frac{h^3B_4}{4!}\left[f'''(x_0+h)-f'''(x_0)\right] - h^5X_3 \\ =& \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ =& \frac{1}{h}\int_{x_0}^{x_0+h} f(y)dy + \sum_{k=1}^{m}\frac{B_{2k}}{(2k)!}h^{2k-1}\left[f^{(2k-1)}(x_0+h)-f^{(2k-1)}(x_0)\right] - h^{2m+1}X_{m+1}\end{aligned} \tag{3.34}$$

Note that in the right hand side of (3.34) we have some terms that resemble those in the Euler-Maclaurin formula (3.27). However, we want to compute the integral of f over $[a,b]$ and so far we only have contributions from x_0 to x_0+h. We can obtain the rest of the contributions by repeating the previous process but for integrals from jh to $(j+1)h$ with $1 \leq j \leq n-1$ and using the periodic extension of the Bernouilli polynomials (3.29). Analogously to (3.30), we have:

$$\begin{aligned}Y_1^j :=& \frac{1}{2}\int_{jh}^{(j+1)h} \overline{B}_2\left(\frac{y}{h}\right) f''(x_0+y)dy \overset{\frac{y}{h}-j=z}{=} \frac{h}{2}\int_0^1 \overline{B}_2(z+j)f''(x_0+h(z+j))dz \overset{(3.29)}{=} \\ =& \frac{h}{2}\int_0^1 B_2(z)f''(x_0+h(z+j))dz \overset{\text{By parts}}{=} \\ =& \frac{h}{2}\left[\underbrace{\frac{B_2(z)f'(x_0+h(z+j))}{h}\Big|_0^1}_{\text{Vanishes due to properties 1 and 2}} - \frac{1}{h}\int_0^1 f'(x_0+h(z+j))B_2'(z)dz\right] \overset{\text{By parts}}{=} \\ =& -\frac{1}{2}\left[\frac{B_2'(z)f(x_0+h(z+j))}{h}\Big|_0^1 - \frac{1}{h}\int_0^1 B_2''(z)f(x_0+h(z+j))dz\right] \overset{(3.6)}{=} \\ =& -\frac{1}{2h}\left[f(x_0+(j+1)h)+f(x_0+jh) - 2\int_0^1 f(x_0+jh+hz)dz\right] \overset{jh+hz=y}{=} \\ =& -\frac{1}{2h}\left[f(x_0+(j+1)h)+f(x_0+jh) - \frac{2}{h}\int_{jh}^{(j+1)h} f(x_0+y)dy\right] = \\ =& -\frac{1}{2h}\left[f(x_0+(j+1)h)+f(x_0+jh)\right] + \frac{1}{h^2}\int_{jh}^{(j+1)h} f(x_0+y)dy \overset{x_0+y=z}{=} \\ =& -\frac{1}{2h}\left[f(x_0+(j+1)h)+f(x_0+jh)\right] + \frac{1}{h^2}\int_{x_0+jh}^{x_0+(j+1)h} f(z)dz\end{aligned} \tag{3.35}$$

We can define as we did with (3.31) for $k>1$:

$$Y_k^j := \frac{1}{(2k)!}\int_{jh}^{(j+1)h} \overline{B}_{2k}\left(\frac{y}{h}\right) f^{(2k)}(x_0+y)dy \tag{3.36}$$

Which also leads to a recursion relation:

$$
\begin{aligned}
Y_k^j &:= \frac{1}{(2k)!}\int_{jh}^{(j+1)h} \overline{B}_{2k}\left(\frac{y}{h}\right) f^{(2k)}(x_0+y)dy \overset{\text{By parts and property 4}}{=} \\
&= \frac{1}{(2k)!}\left[\underbrace{\overline{B}_{2k}\left(\frac{y}{h}\right) f^{(2k-1)}(x_0+y)\Big|_{jh}^{(j+1)h}}_{\text{Vanishes due to property 1 (3.29)}} - \int_{jh}^{(j+1)h} f^{(2k-1)}(x_0+y)\frac{2k}{h}\overline{B}_{2k-1}\left(\frac{y}{h}\right)dy\right] = \\
&= -\frac{1}{h(2k-1)!}\int_{jh}^{(j+1)h} f^{(2k-1)}(x_0+y)\overline{B}_{2k-1}\left(\frac{y}{h}\right)dy \overset{\text{By parts and property 5}}{=} \\
&= -\frac{1}{h(2k-1)!}\left[\underbrace{\overline{B}_{2k-1}\left(\frac{y}{h}\right) f^{(2k-2)}(x_0+y)\Big|_{jh}^{(j+1)h}}_{\text{Vanishes due to property and (3.29)}} - \int_{jh}^{(j+1)h} f^{(2k-2)}(x_0+y)\frac{2k-1}{h}\left(\overline{B}_{2k-2}\left(\frac{y}{h}\right)+B_{2k-2}\right)dy\right] = \\
&= \frac{1}{h^2(2k-2)!}\int_{jh}^{(j+1)h} f^{(2k-2)}(x_0+y)\overline{B}_{2k-2}\left(\frac{y}{h}\right)dy \\
&+ \frac{\overline{B}_{2k-2}}{h^2(2k-2)!}\left[f^{(2k-3)}(x_0+(j+1)h) - f^{(2k-3)}(x_0+jh)\right] \overset{Y_k \text{ definition (3.36)}}{=} \\
&= \frac{Y_{k-1}}{h^2} + \frac{B_{2k-2}}{h^2(2k-2)!}\left[f^{(2k-3)}(x_0+(j+1)h) - f^{(2k-3)}(x_0+jh)\right]
\end{aligned}
\tag{3.37}
$$

So then:

$$
h^2 Y_k^j - Y_{k-1}^j = \frac{B_{2k-2}}{(2k-2)!}\left[f^{(2k-3)}(x_0+(j+1)h) - f^{(2k-3)}(x_0+jh)\right] \tag{3.38}
$$

With this recursion relation, we get a similar result to (3.34) using the fact that $f \in C^{2m+2}[a,b]$:

$$
\begin{aligned}
&\frac{1}{2}\left[f(x_0+(j+1)h)+f(x_0+jh)\right] = \frac{1}{h}\int_{x_0+jh}^{x_0+(j+1)h} f(y)dy - hY_1^j = \\
=&\frac{1}{h}\int_{x_0+jh}^{x_0+(j+1)h} f(y)dy + \frac{hB_2}{2!}\left[f'(x_0+(j+1)h) - f'(x_0+jh)\right] - h^3Y_2^j = \\
=&\frac{1}{h}\int_{x_0+jh}^{x_0+(j+1)h} f(y)dy + \frac{hB_2}{2!}\left[f'(x_0+(j+1)h) - f'(x_0+jh)\right] + \frac{h^3B_4}{4!}\left[f'''(x_0+(j+1)h) - f'''(x_0+jh)\right] - h^5Y_3^j = \\
=&\cdots\cdots\cdots\cdots\cdots\cdots \\
=&\frac{1}{h}\int_{x_0+jh}^{x_0+(j+1)h} f(y)dy + \sum_{k=1}^{m}\frac{B_{2k}}{(2k)!}h^{2k-1}\left[f^{(2k-1)}(x_0+(j+1)h) - f^{(2k-1)}(x_0+jh)\right] - h^{2m+1}Y_{m+1}
\end{aligned}
\tag{3.39}
$$

Now, if we sum (3.34) and (3.39) for $j = 1, \ldots, n-1$, we obtain for the left hand side:

$$
\begin{aligned}
&\frac{1}{2}\left[f(x_0+h)+f(x_0)\right] + \frac{1}{2}\sum_{j=1}^{n-1}\left[f(x_0+(j+1)h)+f(x_0+jh)\right] = \\
=&\frac{1}{2}\left[f(x_0+h)+f(x_0)\right] + \frac{1}{2}\sum_{k=2}^{n} f(x_0+kh) + \frac{1}{2}\sum_{j=1}^{n-1} f(x_0+jh) \overset{\text{Take the term with } k=1 \text{ out of the sum}}{=} \\
=&\frac{1}{2}\left[f(x_0+nh)+f(x_0)\right] + \frac{1}{2}\sum_{k=1}^{n-1} f(x_0+kh) + \frac{1}{2}\sum_{j=1}^{n-1} f(x_0+jh) = \\
=&\frac{1}{2}\left[f(x_0+nh)+f(x_0)\right] + \sum_{j=1}^{n-1} f(x_0+jh) \overset{\text{Summing 0}}{=} \\
=&-\frac{1}{2}\left[f(x_0+nh)+f(x_0)\right] + \sum_{j=0}^{n} f(x_0+jh)
\end{aligned}
\tag{3.40}
$$

And for the right hand side:

$$\begin{aligned}
&\frac{1}{h}\int_{x_0}^{x_0+nh} f(y)dy + \sum_{k=1}^{m}\frac{B_{2k}h^{2k-1}}{(2k)!}\left[f^{(2k-1)}(x_0+h) - f^{(2k-1)}(x_0) + \sum_{j=1}^{n-1}\left[f^{(2k-1)}(x_0+(j+1)h) - f^{(2k-1)}(x_0+jh)\right]\right] \\
&- h^{2m+1}(X_{m+1} + \sum_{j=1}^{n-1} Y_{m+1}^{j}) = \text{(Using definition (3.36) and noting the telescopic behaviour of the series)} = \\
&= \frac{1}{h}\int_{x_0}^{x_0+nh} f(y)dy + \sum_{k=1}^{m}\frac{B_{2k}h^{2k-1}}{(2k)!}\left[f^{(2k-1)}(x_0+nh) - f^{(2k-1)}(x_0)\right] \\
&- \frac{h^{2m+1}}{(2k+2)!}\int_{x_0}^{x_0+nh} \overline{B}_{2m+2}\left(\frac{y}{h}\right) f^{(2k)}(x_0+y)dy \overset{x_0+y=z}{=} \\
&= \frac{1}{h}\int_{x_0}^{x_0+nh} f(y)dy + \sum_{k=1}^{m}\frac{B_{2k}h^{2k-1}}{(2k)!}\left[f^{(2k-1)}(x_0+nh) - f^{(2k-1)}(x_0)\right] \\
&- \frac{h^{2m+1}}{(2k+2)!}\int_{x_0}^{x_0+nh} \overline{B}_{2m+2}\left(\frac{z-x_0}{h}\right) f^{(2k)}(z)dz
\end{aligned} \tag{3.41}$$

Finally, combining (3.40) and (3.41) and writing $x_0 = a$ and $x_0 + nh = b$, we obtain:

$$\begin{aligned}
\int_a^b f(y)dy = h\sum_{j=0}^{n} f(a+jh) - \frac{h}{2}\left[f(b)+f(a)\right] - \sum_{k=1}^{m}\frac{B_{2k}h^{2k}}{(2k)!}\left[f^{(2k-1)}(b) - f^{(2k-1)}(a)\right] \\
+ \frac{h^{2m+2}}{(2m+2)!}\int_a^b \overline{B}_{2m+2}\left(\frac{z-a}{h}\right) f^{(2k)}(z)dz
\end{aligned} \tag{3.42}$$

Which coincides with (3.27) and the proof is now complete. □

For an easier analysis of this important result, we shall write the following corollary [1, page 288], [3, page 131]:

Corollary 3.1. *The error term in (3.28) can be expressed as:*

$$E_m = -\frac{h^{2m+2}(b-a)B_{2m+2}}{(2m+2)!}f^{(2m+2)}(\zeta) \tag{3.43}$$

Proof. For this proof, we will need property 4. in proposition 7.1, which states that $(-1)^k B_{2k}(x) > 0, 0 < x < 1$ for all $k \geq 1$. The corresponding proof can be found in the proposition.

We shall focus now on the integral in (3.28). We will use the same rationale we used to proof the mean value theorem for integrals 2.1. Since $f^{(2m+2)}$ is a continuous function in $[a,b]$ by hypothesis, it reaches a minimum and maximum value in $[a,b]$. Let $m = \min_{x\in[a,b]} f^{(2m+2)}(x)$ and $M = \max_{x\in[a,b]} f^{(2m+2)}(x)$. As a result, in $[a,b]$:

$$m \leq f^{(2m+2)}(x) \leq M \tag{3.44}$$

Since $(-1)^k B_{2k}(x) > 0, 0 < x < 1$ for all $k \geq 1$, we can prove that the integral $\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx$ is positive number:

$$\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \overset{y=\frac{x-a}{h}}{=} \frac{1}{h}\int_0^1 (-1)^{m+1}B_{2m+2}(y)dy \overset{\text{4. in prop (3.28)}}{>} 0 \tag{3.45}$$

and therefore, from (3.44):

$$\begin{aligned}
m\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \leq f^{(2m+2)}(y)\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \leq \\
\leq M\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx
\end{aligned} \tag{3.46}$$

And using that $(-1)^k B_{2k}(x) > 0, 0 < x < 1$ for all $k \geq 1$ we also have in $[a,b]$:

$$\begin{aligned} m\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \leq \int_a^b f^{(2m+2)}(x)(-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \leq \\ \leq M\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \end{aligned} \tag{3.47}$$

Inequalities (3.46), (3.47) and the fact that $f^{(2m+2)}$ is a continuous function imply that there is $\zeta \in [a,b]$ such that:

$$f^{(2m+2)}(\zeta)\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx = \int_a^b f^{(2m+2)}(x)(-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \tag{3.48}$$

With this equality, we can jump to equality (3.28) to obtain:

$$\begin{aligned} E_m &= \frac{h^{2m+2}}{(2m+2)!}\int_a^b \overline{B}_{2m+2}\left(\frac{x-a}{h}\right)f^{(2m+2)}(x)dx = \\ &= \frac{h^{2m+2}}{(2m+2)!(-1)^{m+1}}\int_a^b f^{(2m+2)}(x)(-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \overset{(3.48)}{=} \\ &= \frac{h^{2m+2}}{(2m+2)!(-1)^{m+1}}f^{(2m+2)}(\zeta)\int_a^b (-1)^{m+1}\overline{B}_{2m+2}\left(\frac{x-a}{h}\right)dx \overset{(3.29)}{=} \\ &= \frac{h^{2m+2}}{(2m+2)!}f^{(2m+2)}(\zeta)n\int_a^{a+h} B_{2m+2}\left(\frac{x-a}{h}\right)dx \overset{y=(x-a)/h}{=} \\ &= \frac{h^{2m+2}}{(2m+2)!}f^{(2m+2)}(\zeta)nh\int_0^1 B_{2m+2}(y)\,dy \end{aligned} \tag{3.49}$$

Finally, using property 6. of proposition 3.1, we can solve the integral to obtain:

$$E_m = -\frac{h^{2m+2}(b-a)B_{2m+2}}{(2m+2)!}f^{(2m+2)}(\zeta) \tag{3.50}$$

And the proof is now complete.

□

We can give now an explanation of why this result is so remarkable. Firstly, note that the first two terms of the right hand side of equality (3.27) are the result of the composite trapezoidal rule (definition 2.6). The third term of the right hand side of (3.27) can be seen as a correction term. Note that as it happened for the asymptotic error of the composite trapezoidal rule (corollary 2.2), the Euler-Mclaurin formula predicts a better accuracy of the trapezoidal rule when the integrand is periodic over the interval of integration. In fact, the Euler-Mclaurin formula goes beyond the asympotic error result since it shows that the trapezoidal rule result gets better when more and more derivatives of the function are equal in a and b.

3.2 Romberg integration

In chapter 2, we discussed how in order to improve the numerical result using a quadrature rule, a naive approach is taking more equally spaced interpolation nodes. We stated that increasing the number of nodes (or equivalently, having a higher degree of exactness) does not guarantee a better result with a quadrature rule (The so called *Runge phenomenon* [3, pages 54-55] exemplified in the appendix). Hence, going for quadrature rules beyond the trapezoidal rule is unnecessary. As a result, we aimed to improve the quadrature rule by increasing the number of subintervals where the trapezoidal rule is applied.

With the Romberg integration procedure that we are about to discuss in this section, we will have an algorithm to iteratively increase the number of subintervals (as described in section 2.3) and to

obtain quadrature rules of higher degree of exactness (if desired) by a recursive application of the trapezoidal rule. This approach supports even more the fact that working with quadrature rules other than the trapezoidal rule is not necessary since these can be obtained by Romberg integration.

To introduce Romberg integration [2], [3, pages 294-296], we shall recall the result of the Euler-Maclaurin formula in theorem 3.1. Let $T(f,h)$ denote the trapezoidal rule applied to f with step h. This result can be written as:

$$\int_a^b f(x)dx = T(f,h) - \sum_{k=1}^{m} \tau_{2k} h^{2k} + E_m \tag{3.51}$$

Where we have:

$$\tau_{2k} = \frac{B_{2k}}{(2k)!}(f^{(2k-1)}(b) - f^{(2k-1)}(a)) \tag{3.52}$$

In order to develop the Romberg integration algorithm, we need to write:

$$\begin{aligned} \int_a^b f(x)dx &= T(f,h) - (\tau_2 h^2 + \tau_4 h^4 + \tau_6 h^6 + \dots) \\ \int_a^b f(x)dx &= T(f,h/2) - \left(\tau_2 \left(\frac{h}{2}\right)^2 + \tau_4 \left(\frac{h}{2}\right)^4 + \tau_6 \left(\frac{h}{2}\right)^6 + \dots\right) \end{aligned} \tag{3.53}$$

Multiplying by 4 the second equation and substracting the first one, we obtain:

$$\int_a^b f(x)dx = \frac{4T(f,h/2) - T(f,h)}{3} + \hat{\tau}_4 \left(\frac{h}{2}\right)^4 + \hat{\tau}_6 \left(\frac{h}{2}\right)^6 \dots \tag{3.54}$$

Where we have rewritten the constants multiplying the powers of h as $\hat{\tau}_{2k}$ to simplify the notation. Hence, just by using twice the trapezoidal rule, we have obtained a quadrature rule with order 4 since the first power of h is h^4. In chapter 2 we stated that the Simpson's rule has order 4. In fact, the ratio in (3.54) is the composite Simpson's rule. The expression can also be obtained by using the explicit expression of the composite Simpson's rule. Let $S(f,h)$ be the Simpson's rule applied to f with step h. Then:

$$S(f,h) = \frac{h}{3}(f_0 + f_{2m}) + \frac{2h}{3}\sum_{j=1}^{m-1} f_{2i} + \frac{4h}{3}\sum_{j=1}^{m} f_{2i-1} \tag{3.55}$$

And by noting that:

$$\begin{aligned} T(f,h) &= \frac{h}{2}(f_0 + f_{2m}) + h\sum_{i=1}^{2m-1} f_i \\ T(f,2h) &= h(f_0 + f_{2m}) + 2h\sum_{j=1}^{m-1} f_{2j} \end{aligned} \tag{3.56}$$

We can obtain:

$$S(f,h) = \frac{4T(f,h) - T(f,2h)}{3} \tag{3.57}$$

So summarizing, with (3.53) and (3.54) we have obtained an order 4 quadrature rule by using an order 2 quadrature rule (the trapezoidal rule). This process can be generalized to obtain higher order quadrature rules with an arbitrary step. Let us denote:

$$h_i = \frac{b-a}{2^{i-1}} \quad i \in \mathbb{N} \tag{3.58}$$

And let $R(i,k)$ be the quadrature rule of order $2k$ with step h_i. Then:

$$\begin{aligned}\int_a^b f(x)dx &= R(i,k) + a_{2k}h_i^{2k} + a_{2k+2}h_i^{2k+2} + \dots \\ \int_a^b f(x)dx &= R(i+1,k) + a_{2k}\left(\frac{h_i}{2}\right)^{2k} + a_{2k+2}\left(\frac{h_i}{2}\right)^{2k+2} + \dots\end{aligned} \tag{3.59}$$

Multiplying the second equation by 4^k, substracting the first one and noting that $h_{i+1} = {}^{h_i}/_2$:

$$\int_a^b f(x)dx = \frac{4^k R(i+1,k) - R(i,k)}{4^k - 1} + \hat{a}_{2k+2}h_{i+1}^{2k+2} + \dots \tag{3.60}$$

Which is an order 2k+2 quadrature rule and then:

$$R(i+1,k+1) = \frac{4^k R(i+1,k) - R(i,k)}{4^k - 1} \quad i \in \mathbb{N} \quad , k \leq i \quad , k \in \mathbb{N} \tag{3.61}$$

And then we have a formula to obtain higher order quadrature rules from lower order quadrature rules. With this formula, we can design an algorithm to obtain high order quadrature rules, but firstly, it is convenient to get an intuition on how to proceed. To do so, we can make use of table 3.62, which orders the quadrature rules with decreasing step in rows and increasing order in columns. By inspection of (3.61), we note that an element $(i+1,k+1)$ of the table is calculated with the elements $(i+1,k)$ and (i,k) of the table i.e the elements at the left and left + above. Hence, in order to calculate $R(i+1,k+1)$, we will need to store in memory $k+1$ quadrature rules: $R(i+1,1), R(i,1), R(i,2), R(i,3), \dots, R(i,k)$ i.e the row above and the first element of the current row.

step \ order	2	4	6	8	
h	$R(1,1)$				
$h/2$	$R(2,1)$	$R(2,2)$			
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$		
$h/8$	$R(4,1)$	$R(4,2)$	$R(4,3)$	$R(4,4)$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$
	Trapezoidal	Simpson	$n=4$	$n=6$	$\dots$

(3.62)

Nevertheless, memory usage can be optimized by getting rid on the go of the entries that will not be used later on and using that memory for the new calculated entries. This way, only $k+1$ at most are stored simultaneously. We write explicitly this process in table 3.1. We use $I(i)$ to denote an entry of a vector following MATLAB's notation.
Once it is clear how to apply (3.61) to obtain higher order quadrature rules, we only need a stopping criteria to finish the algorithm. To obtain one, we can modify (3.61) by setting i instead of $i+1$:

$$R(i,k+1) = \frac{4^k R(i,k) \overbrace{-R(i,k) + R(i,k)}^{\text{We are adding 0 to (3.61)}} - R(i-1,k)}{4^k - 1} = R(i,k) + \frac{R(i,k) - R(i-1,k)}{4^k - 1} \tag{3.63}$$
$$i = 2,3,\dots \quad , k = 1,2,\dots,i-1$$

This expression relates a $2(k+1)$ order quadrature rule with its $2k$ order counterpart. We find convenient to define:

$$\nabla_i R(i,k) := R(i,k) - R(i-1,k) \quad i = 2,3,\dots \quad k = 1,2,\dots,i-1 \tag{3.64}$$

Consider the case in which i gives a step such that $R(i,k+1) \approx I(f)$. In such a case , we can consider ${}^{\nabla_i R(i,k)}/_{(4^k-1)}$ as the error after approximating the integral by the quadrature rule $R(i,k)$. Under these considerations, ${}^{\nabla_i R(i,k)}/_{(4^k-1)}$ can be treated as the error to consider for the stopping criteria.

Algorithm 3.1. ***(Calculating $\int_a^b f(x)dx$ with Romberg integration and an absolute error less than ϵ).*** *Inputs:*

1. $\epsilon > 0$
2. *The function* $f(x)$ *to be integrated.*
3. $a < b$ *real numbers defining the beginning and end of the interval of integration.*
4. $n \in \mathbb{N}$ *to define the initial number of subintervals for the initial calculation of the trapezoidal rule.*
5. $n_f \in \mathbb{N}$ *that defines the maximum number of subintervals to consider.*

Outputs:

1. *An approximation of* $I(f) = \int_a^b f(x)dx$

Algorithm

1. *Set* $h = \frac{b-a}{n}$.
2. *Set* $i = 1$.
3. *Set* $\Delta = 1 + \epsilon$.
4. *Calculate* $I(i) = \frac{h}{2}(f(a) + f(b))$.
5. `IF` $(n > 1)$:[1]
$$I(i) = I(i) + h\sum_{j=1}^{n-1} f(a + jh)$$
6. `DO WHILE` $(\Delta > \epsilon \;\&\; i < n_f)$: *(In this loop we follow the process of table 3.1 until the error is lower than* ϵ *or we reach the maximum number of subintervals* n_f*)*
   (a) $i = i + 1$,
   (b) $h = {}^{h}/_{2}$
   (c) $I(i) = \frac{I(i-1)}{2} + h\sum_{j=1}^{n} f(a + (2j-1)h)$ *(The trapezoidal rule with the new step is calculated using (2.37)).*
   (d) $n = 2n$ *(Because reducing the step to a half duplicates the number of subintervals for the trapezoidal rule)*
   (e) $k = 0$
   (f) `DO WHILE` $(\Delta > \epsilon \;\&\; k < i - 1)$
      i. $k = k + 1$
      ii. $\Delta_k = \frac{I(i-k+1) - I(i-k)}{4^k - 1}$ *(Here we calculate the ratio in (3.63))*
      iii. $\Delta = |\Delta_k|$
      iv. $I(i-k) = I(i-k+1) + \Delta_k$ *(Here we are calculating (3.63))*
7. `RETURN:` $I = I(i-k)$

Remark 3.2. *Note that the previous algorithm follows the process exemplified in table 3.1 but starting from* $R(n,1)$ $, n > 1$ *instead of* $R(1,1)$*. Starting from a higher number of intervals guarantees a better initial approximation of the integral. However, more steps will be needed to reach higher order quadrature rules than if started from* $R(1,1)$*. To make such realization, note how table 3.1 gets far in columns with increasing steps. If we start with n subintervals, we are basically replicating the process in table 3.1 but with a shifting in rows so that we start from* $R(n,1)$ *but we still need steps to add columns and get to high order quadrature rules.*

[1] In this step we are simply calculating the trapezoidal rule with n subintervals i.e step ${}^{(b-a)}/_{2^{n-1}}$

Remark 3.3. *The algorithm can be used with input $n = 1$. By doing so, the process in table 3.1 is fully replicated until getting $R(n_f, n_f)$.*

Remark 3.4. *We discussed above that the stopping criteria relies on the assumption that $R(i,k+1) \approx I(f)$. However, we did not clarify under which conditions the assumption is satisfied. We can introduce a way of checking this. Firstly, note that if $R(i,k+1) \approx I(f)$, then, writing $h = b - a$ we have:*

$$\begin{aligned}
I(f) &\underset{\text{Euler-Maclaurin}}{\approx} R(i,k+1) + a_{2k+2}h_i^{2k+2} \underset{(3.58)}{\approx} R(i,k+1) + a_{2k+2}h^{2k+2}2^{-(2k+2)(i-1)} \\
I(f) &\underset{\text{Euler-Maclaurin}}{\approx} R(i-1,k+1) + a_{2k+2}h_{i-1}^{2k+2} \underset{(3.58)}{\approx} R(i-1,k+1) + a_{2k+2}h^{2k+2}2^{-(2k+2)(i-2)} \\
I(f) &\underset{\text{Euler-Maclaurin}}{\approx} R(i+1,k+1) + a_{2k+2}h_{i+1}^{2k+2} \underset{(3.58)}{\approx} R(i+1,k+1) + a_{2k+2}h^{2k+2}2^{-(2k+2)i}
\end{aligned} \tag{3.65}$$

And with the previous approximations:

$$\begin{aligned}
\Delta_{rel}(k) &= \frac{R(i,k+1) - R(i-1,k+1)}{R(i+1,k+1) - R(i,k+1)} = \frac{2^{-2ki+2k-2i+2} - 2^{-2ki+4k-2i+4}}{2^{-2ki-2i} - 2^{-2ki+2k-2i+2}} \\
&= \frac{2^{-2ki-2i}2^{2k+2}\left(1 - 2^{2k+2}\right)}{2^{-2ki-2i}\left(1 - 2^{2k+2}\right)} = 2^{2k+2} = 2^{2(k+1)} = 4^{k+1}
\end{aligned} \tag{3.66}$$

So by checking the evolution of the relative error, we can verify whether or not we are in the regime in which the discussed stopping criteria makes sense. When the relative error changes as in (3.66) when reducing the step to a half, then we can argue that the stopping criteria applies.

We started this session discussing that it was not worth the effort to use quadrature rules beyond the trapezoidal rule. With this algorithm, we firmly reinforce that statement. There is no need to obtain the explicit expressions for the higher order quadrature rules because a lot of them can be obtained by means of the trapezoidal rule and Romberg integration.

With Romberg integration explained, we can stop extending the trapezoidal rule in compact intervals. The next matter will be the extension and behaviour of the trapezoidal rule when integrating over the whole real line.

step \ order	Order			
	2	4	6	8
h	$R(1,1)$			
	$I(1)$			
$h/2$	$R(2,1)$			
	$I(2)$			
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
	$I(2)$	$I(1)$		
$h/4$	$R(3,1)$			
	$I(3)$			
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
		$I(1)$		
$h/4$	$R(3,1)$	$R(3,2)$		
	$I(3)$	$I(2)$		
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
	$I(3)$	$I(2)$	$I(1)$	
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
	$I(3)$	$I(2)$	$I(1)$	
$h/8$	$R(4,1)$			
	$I(4)$			
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
		$I(2)$	$I(1)$	
$h/8$	$R(4,1)$	$R(4,2)$		
	$I(4)$	$I(3)$		
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
			$I(1)$	
$h/8$	$R(4,1)$	$R(4,2)$	$R(4,3)$	
	$I(4)$	$I(3)$	$I(2)$	
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
$h/8$	$R(4,1)$	$R(4,2)$	$R(4,3)$	$R(4,4)$
	$I(4)$	$I(3)$	$I(2)$	$I(1)$
h	$R(1,1)$			
$h/2$	$R(2,1)$	$R(2,2)$		
$h/4$	$R(3,1)$	$R(3,2)$	$R(3,3)$	
$h/8$	$R(4,1)$	$R(4,2)$	$R(4,3)$	$R(4,4)$
	$I(4)$	$I(3)$	$I(2)$	$I(1)$
$h/16$	$R(5,1)$			
	$I(5)$			

Table 3.1: Step by step application of the Romberg integration given by (3.61) keeping in mind saving memory.

Give a man a fish and you feed him for a day. Teach a man how to fish and you feed him for a lifetime.

Lao Tzu

Chapter 4

Exponentially convergent trapezoidal rule: periodic integrals and integrals over $\mathbb{R}$ [3], [7]

In the previous chapters we have introduced the composite trapezoidal rule, showing its error estimate and its convenient behaviour for integrals of periodic functions over their period. In this chapter, we will show some of the most important results of the trapezoidal rule: the exponential decrease of the error as a function of the number of nodes (equivalently, the step h) for the trapezoidal rule for periodic functions and for integrals over the whole real line.

4.1 The exponential convergence in periodic integrands

The Euler-Maclaurin formula 3.1 already hints that the error of the composite trapezoidal rule improves drastically when integrating periodic functions over their period.

In this section we will show that, under some conditions, the convergence with the number of nodes (equivalently, the number of subintervals) is way better. In fact, the error decreases exponentially with the number of nodes. For the sake of simplicity, we will consider functions that are 2π periodic and the integral:

$$I = \int_0^{2\pi} v(\theta) d\theta \tag{4.1}$$

Where $v(\theta)$ is 2π-periodic but is not necessarily a real function. If we consider the composite trapezoidal rule with N nodes, the trapezoidal approximation to the previous integral is:

$$I_N = \frac{2\pi}{N} \sum_{k=1}^{N} v(\theta_k) \overset{\theta_k = \frac{2\pi k}{N}}{=} \frac{2\pi}{N} \sum_{k=1}^{N} v\left(\frac{2\pi k}{N}\right) \tag{4.2}$$

Assuming a period of 2π is not a loss of generality because any other integral can be transformed conveniently with a correct change of variable. Therefore, we can focus strictly on this case. We will follow the steps in [7][section 3] with more detail.

Theorem 4.1. ***(Exponential decrease of the error for periodic, bounded and analytic functions in a half plane)*** *Suppose v is a 2π-periodic and analytic function in the half-plane set*

$$H = \{z \in \mathbb{C} \mid Im(z) > -a\} \tag{4.3}$$

For some $a > 0$. Suppose $|v(\theta)| \leq M$ in H where M is a nonegative real number. Let N be the number of nodes used in the trapezoidal approximation (4.2) of the integral (4.1). Then:

$$|I_N - I| \leq \frac{2\pi M}{e^{aN} - 1} \tag{4.4}$$

Proof. Firstly, note that the set H verifies the property 8.2. Since v is 2π-periodic and analytic, we can write its Fourier series using theorem 8.2:

$$v(\theta) = \sum_{j=-\infty}^{j=\infty} c_j e^{ij\theta} \tag{4.5}$$

Where the coefficients are:

$$c_j = \frac{1}{2\pi}\int_0^{2\pi} e^{-ij\theta} v(\theta) d\theta \tag{4.6}$$

We will show now that this Fourier series is absolutely and uniformly convergence. We can start by showing that all coefficients with $j < 0$ are zero. To do so, we can shift upwards the path of integration in (4.6) a distance b. This new path remains in the interior of the set H. Then:

$$\begin{aligned} |c_j| &= \left|\frac{1}{2\pi}\int_0^{2\pi} e^{-ij(\theta+ib)} v(\theta+ib) d\theta\right| = \frac{e^{jb}}{2\pi}\left|\int_0^{2\pi} e^{-ij\theta} v(\theta+ib) d\theta\right| \leq \\ &\leq \frac{e^{jb}}{2\pi}\int_0^{2\pi} |v(\theta+ib)|\, d\theta \overset{j<0}{=} \frac{e^{-|j|b}}{2\pi}\int_0^{2\pi} |v(\theta+ib)|\, d\theta \overset{|v(\theta)\leq M \text{ in } H}{\leq} e^{-|j|b} M \end{aligned} \tag{4.7}$$

But this can be done for all $b > 0$ so $|c_j| \leq e^{-|j|b} M \;\forall\; b > 0$ and therefore, $c_j = 0 \;\forall\; j < 0$.

Now we can proceed in a similar manner with those coefficients with $j \geq 0$. We can shift the path of integration in (4.6) downwards a quantity $\beta < a$ so that the new path still lies in the set H. Then:

$$\begin{aligned} |c_j| &= \frac{1}{2\pi}\left|\int_0^{2\pi} e^{-ij(\theta-i\beta)} v(\theta-i\beta) d\theta\right| = \frac{e^{-j\beta}}{2\pi}\left|\int_0^{2\pi} e^{-ij\theta} v(\theta-i\beta) d\theta\right| \leq \\ &\leq \frac{e^{-j\beta}}{2\pi}\int_0^{2\pi} |v(\theta-i\beta)|\, d\theta \overset{|v(\theta)\leq M \text{ in } H}{\leq} e^{-j\beta} M \quad j > 0 \end{aligned} \tag{4.8}$$

With the previous inequality, we can bound the series of the coefficients:

$$\sum_{j=\infty}^{\infty} |c_j| = \sum_{j=0}^{\infty} |c_j| \leq \sum_{j=0}^{\infty} e^{-j\beta} M = M \sum_{j=0}^{\infty} \frac{1}{(e^{\beta})^j} < \infty \tag{4.9}$$

Where the last bound follows from the fact that the sum is a geometric series with $r = e^{\beta} > 0$. In the end, we have that the coefficients in (4.6) satisfy the hypothesis in theorem 8.3 and therefore, the Fourier series for $v(\theta)$ is uniformly and absolutely convergent. We will need this later.

Now, from (4.6):

$$I = \int_0^{2\pi} v(\theta) d\theta = 2\pi c_0 \tag{4.10}$$

And from (4.1) and (4.5) we can write:

$$I_N = \frac{2\pi}{N}\sum_{k=1}^{N} v(\theta_k) = \frac{2\pi}{N}\sum_{k=1}^{N}\sum_{j=0}^{\infty} c_j e^{\frac{ij2\pi k}{N}} \tag{4.11}$$

The absolute convergence shown in (4.9) allows to interchange the sums:

$$I_N = \frac{2\pi}{N}\sum_{j=0}^{\infty} c_j \sum_{k=1}^{N} e^{\frac{2\pi ikj}{N}} = 2\pi \sum_{j=0}^{\infty} c_j \delta_{j,mN} \tag{4.12}$$

Where the last equality comes from the fact that the last sum gives N if j is a multiple of N and zero otherwise[1]. Consequently, we have:

$$I_N = 2\pi \sum_{m=0}^{\infty} c_{mN} \tag{4.13}$$

[1]Recall that $\sum_{k=1}^{N} r^k = \frac{r(r^N-1)}{r-1}$. In this case, $r = e^{\frac{2\pi ij}{N}}$. Then, $\sum_{k=1}^{N}\left(e^{\frac{2\pi ij}{N}}\right)^k = \frac{e^{\frac{2\pi ij}{N}}(e^{2\pi ij}-1)}{e^{\frac{2\pi ij}{N}}-1} = 0$ since $e^{2\pi ij} = 1$. On the other hand, if $j = mN$ with $m \in \mathbb{N}$, then $\sum_{k=1}^{N}\left(e^{\frac{2\pi ij}{N}}\right)^k = \sum_{k=1}^{N} 1 = N$.

Which combined with (4.10) gives:

$$I_N - I = 2\pi \sum_{m=1}^{\infty} c_{mN} \tag{4.14}$$

We can bound this difference the same way we did for (4.8):

$$\begin{aligned} |I_N - I| = 2\pi \left| \sum_{m=1}^{\infty} c_{mN} \right| \leq 2\pi \sum_{m=1}^{\infty} |c_{mN}| \overset{(4.8)}{\leq} 2\pi \sum_{m=1}^{\infty} M e^{-mN\beta} = 2\pi M \sum_{m=1}^{\infty} \left(\frac{1}{e^{N\beta}} \right)^m = \\ = 2\pi M \frac{e^{-N\beta}}{1 - e^{-N\beta}} = \frac{2\pi M}{e^{N\beta} - 1} \end{aligned} \tag{4.15}$$

Finally, when obtaining the bound in (4.8), we chose a real β such that $\beta < a$. Nevertheless, note that we can repeat the exact same process for values of β arbitrarily close to a i.e for all $\beta \in (0, a)$. This results in $|c_j| \leq e^{-ja} M$ for all $j > 0$ and therefore, we can write from the last inequality:

$$|I_N - I| \leq \frac{2\pi M}{e^{Na} - 1} \tag{4.16}$$

And the proof is now complete. □

Remark 4.1. *The previous theorem can be used for both complex and real functions as long as the requirements are satisfied. However, the hypothesis of being analytic in a half plane is a bit restrictive. This restriction is even more irritating if we are just interested in integrating a function $f \in \mathcal{C}^\infty(\mathbb{R})$ such that $f(x)$ is real for real x. Consider the example of $f(z) = \frac{1}{\sin(z)+2}$. This function is in $\mathcal{C}^\infty(\mathbb{R})$ and real for real z, but it is not analytic in a half-plane due to its poles. We can find their location by solving $sin(x) = -2$. Writing the sine in terms of exponentials, the equality is equivalent to:*

$$e^{ix} - e^{-ix} = -4i \Longrightarrow e^{2ix} - 1 = -4ie^{ix} \tag{4.17}$$

And by writing $z = e^{ix}$, we obtain the algebraic equation:

$$z^2 + 4iz - 1 = 0 \tag{4.18}$$

With solutions $z = i(-2 \pm \sqrt{3})$. Therefore

$$e^{ix} = i(-2 + \sqrt{3}) \quad or \quad e^{ix} = i(-2 - \sqrt{3}) \tag{4.19}$$

And taking the logarithm on both sides, we get:

$$\begin{aligned} ix = \ln\left(i(-2+\sqrt{3})\right) + 2k\pi \quad or \quad ix = \ln\left(i(-2-\sqrt{3})\right) + 2k\pi \\ ix = \ln(i) + \ln\left(-2+\sqrt{3}\right) + 2k\pi \quad or \quad ix = \ln(i) + \ln\left(-2-\sqrt{3}\right) + 2k\pi \\ ix = \frac{i\pi}{2} + \ln\left(-2+\sqrt{3}\right) + 2k\pi \quad or \quad ix = \frac{i\pi}{2} + \ln\left(-2-\sqrt{3}\right) + 2k\pi \\ x = \frac{\pi}{2} - i\ln\left(-2+\sqrt{3}\right) + 2k\pi \quad or \quad x = \frac{\pi}{2} - i\ln\left(-2-\sqrt{3}\right) + 2k\pi \end{aligned} \tag{4.20}$$

With $k \in \mathbb{Z}$. We just have to simplify $\ln\left(-2+\sqrt{3}\right)$ and $\ln\left(-2-\sqrt{3}\right)$:

$$\begin{aligned} \ln\left(-2+\sqrt{3}\right) = \ln\left(-(2-\sqrt{3})\right) = \ln(-1) + \ln\left(2-\sqrt{3}\right) = i\pi + \ln\left(2-\sqrt{3}\right) \\ \ln\left(-2-\sqrt{3}\right) = \ln\left(-(2+\sqrt{3})\right) = \ln(-1) + \ln\left(2+\sqrt{3}\right) = i\pi + \ln\left(2+\sqrt{3}\right) \end{aligned} \tag{4.21}$$

So in the end:

$$x = \frac{3\pi}{2} + 2k\pi - i\ln\left(2-\sqrt{3}\right) \quad or \quad x = \frac{3\pi}{2} + 2k\pi - i\ln\left(2+\sqrt{3}\right) \tag{4.22}$$

We can make a final and very convenient simplification:

$$\begin{gathered}\left(2-\sqrt{3}\right)\left(2+\sqrt{3}\right)=1\\ \Longrightarrow \ln\left((2-\sqrt{3})(2+\sqrt{3})\right)=\ln\left(2-\sqrt{3}\right)+\ln\left(2+\sqrt{3}\right)=0\\ \Longrightarrow \ln\left(2-\sqrt{3}\right)=-\ln\left(2+\sqrt{3}\right)\end{gathered} \tag{4.23}$$

And we can write the solutions as:

$$x=\frac{3\pi}{2}+2k\pi+i\ln\left(2+\sqrt{3}\right) \quad or \quad x=\frac{3\pi}{2}+2k\pi-i\ln\left(2+\sqrt{3}\right) \tag{4.24}$$

Consequently, the function $f(x)=\frac{1}{\sin(x)+2}$ *is not analytic in a half plane, but it is analytic in the strip:*

$$H=\{z\in\mathbb{C}\mid -\ln\left(2+\sqrt{3}\right)<Im(z)<\ln\left(2+\sqrt{3}\right)\} \tag{4.25}$$

There are many use cases like this in which the function we want to integrate is analytic in a strip around real axis instead on a half-plane. For these cases in which theorem 4.1 does not apply, we have its strip-domain counterpart in the following theorem.

Theorem 4.2. ***(Exponential decrease of the error for periodic, bounded and analytic functions in a strip)*** *Suppose* v *is a* 2π*-periodic and analytic function in the strip defined by the set*

$$H=\{z\in\mathbb{C}\mid -a<Im(z)<a\} \tag{4.26}$$

For some $a>0$*. Suppose* $|v(\theta)|\leq M$ *in* H *where* M *is a nonegative real number. Let* N *be the number of nodes used in the trapezoidal approximation (4.2) of the integral (4.1). Then:*

$$|I_N-I|\leq\frac{4\pi M}{e^{aN}-1} \tag{4.27}$$

Proof. The proof of this theorem follow the same ideas as those for theorem 4.1. The set H verifies the property 8.2. Since v is 2π-periodic and analytic, we can write its Fourier series using theorem 8.2:

$$v(\theta)=\sum_{j=-\infty}^{j=\infty}c_j e^{ij\theta} \tag{4.28}$$

Where the coefficients are:

$$c_j=\frac{1}{2\pi}\int_0^{2\pi}e^{-ij\theta}v(\theta)d\theta \tag{4.29}$$

We will bound the coefficients to show that the Fourier series in (4.28) converges uniformly and absolutely. For the terms with $j<0$, we can shift upwards the path of integration in (4.29) a quantity $b<a$. This path is still in H and it serves us to obtain:

$$\begin{gathered}|c_j|=\left|\frac{1}{2\pi}\int_0^{2\pi}e^{-ij(\theta+ib)}v(\theta+ib)d\theta\right|=\frac{e^{jb}}{2\pi}\left|\int_0^{2\pi}e^{-ij\theta}v(\theta+ib)d\theta\right|\leq\\ \leq\frac{e^{jb}}{2\pi}\int_0^{2\pi}|v(\theta+ib)|\,d\theta\overset{j\leq 0}{=}\frac{e^{-|j|b}}{2\pi}\int_0^{2\pi}|v(\theta+ib)|\,d\theta\overset{|v(\theta)\leq M\text{ in }H}{\leq}e^{-|j|b}M\end{gathered} \tag{4.30}$$

But this process can be repeated for all $b<a$. Hence, $|c_j|\leq e^{-|j|a}M$ for all $j<0$. For the coefficients with $j\geq 0$, we can shift downwards the path in (4.29) a quantity $\beta<a$. The path is still in H and it serves us to obtain:

$$\begin{gathered}|c_j|=\frac{1}{2\pi}\left|\int_0^{2\pi}e^{-ij(\theta-i\beta)}v(\theta-i\beta)d\theta\right|=\frac{e^{-j\beta}}{2\pi}\left|\int_0^{2\pi}e^{-ij\theta}v(\theta-i\beta)d\theta\right|\leq\\ \leq\frac{e^{-j\beta}}{2\pi}\int_0^{2\pi}|v(\theta-i\beta)|\,d\theta\overset{|v(\theta)\leq M\text{ in }H}{\leq}e^{-j\beta}M\quad j>0\end{gathered} \tag{4.31}$$

But this process can be repeated for all $\beta < a$. Therefore, $|c_j| \leq e^{-ja}M$ for all $j \geq 0$. With these bounds

$$\begin{aligned}\sum_{j=\infty}^{\infty}|c_j| = \sum_{j=-1}^{-\infty}|c_j| + \sum_{j=0}^{\infty}|c_j| \leq \sum_{j=-1}^{-\infty} e^{-|j|a}M + \sum_{j=0}^{\infty} e^{-ja}M = M\sum_{j=1}^{\infty} e^{-ja} + M\sum_{j=0}^{\infty} e^{-ja} = \\ = M\sum_{j=1}^{\infty}\frac{1}{(e^a)^j} + M\sum_{j=0}^{\infty}\frac{1}{(e^a)^j} < \infty\end{aligned} \tag{4.32}$$

Where the last bound follows from the fact that both sums are a geometric series with $r = e^a > 0$. In the end, we have that the coefficients in (4.29) satisfy the hypothesis in theorem 8.3 and therefore, the Fourier series for $v(\theta)$ is uniformly and absolutely convergent. We will need this to interchange sums later as we did in the proof of theorem (4.1). Now, from (4.29)

$$I = \int_0^{2\pi} v(\theta)d\theta = 2\pi c_0 \tag{4.33}$$

And from (4.2) and (4.28) we can write:

$$I_N = \frac{2\pi}{N}\sum_{k=1}^{N} v(\theta_k) = \frac{2\pi}{N}\sum_{k=1}^{N}\sum_{j=-\infty}^{\infty} c_j e^{\frac{ij2\pi k}{N}} \tag{4.34}$$

The absolute convergence shown in (4.32) allows to interchange the sums:

$$I_N = \frac{2\pi}{N}\sum_{j=-\infty}^{\infty} c_j \sum_{k=1}^{N} e^{\frac{2\pi ikj}{N}} = 2\pi\sum_{j=-\infty}^{\infty} c_j\delta_{j,mN} \tag{4.35}$$

Where the last equality comes from the fact that the last sum gives N if j is a multiple of N and zero otherwise. Consequently, we have:

$$I_N = 2\pi c_0 + 2\pi\sum_{m=1}^{\infty}(c_{mN} + c_{-mN}) \tag{4.36}$$

Which combined with (4.33) gives:

$$I_N - I = 2\pi\sum_{m=1}^{\infty}(c_{mN} + c_{-mN}) \tag{4.37}$$

And we can bound this quantity as a consequence of (4.30) and (4.31):

$$\begin{aligned}|I_N - I| \leq 2\pi\left|\sum_{m=1}^{\infty}(c_{mN} + c_{-mN})\right| \leq 2\pi\sum_{m=1}^{\infty}|(c_{mN} + c_{-mN})| \leq \\ \leq 2\pi\sum_{m=1}^{\infty}|c_{mN}| + 2\pi\sum_{m=1}^{\infty}|c_{-mN}| = 2\pi\sum_{m=1}^{\infty}|c_{mN}| + 2\pi\sum_{m=-1}^{-\infty}|c_{mN}| \overset{(4.30),(4.31)}{\leq} \\ \leq 2\pi\sum_{m=1}^{\infty} e^{-mNa}M + 2\pi\sum_{m=-1}^{-\infty} e^{-|mN|a}M = 2\pi M\left[\sum_{m=1}^{\infty} e^{-mNa} + \sum_{m=1}^{\infty} e^{-mNa}\right] = \\ = 4\pi M\sum_{m=1}^{\infty} e^{-mNa} = 4\pi M\sum_{m=1}^{\infty}\frac{1}{(e^{Na})^m} = 4\pi M\frac{e^{-Na}}{1-e^{-Na}} = \frac{4\pi M}{e^{Na}-1}\end{aligned} \tag{4.38}$$

And the proof is complete. □

Remark 4.2. ***(Does the largest*** *a* ***give the best error bound?*** ***[7][section 4])*** *The bounds for* $|I_N - I|$ *in theorems 4.1 and 4.2 are respectively:*

$$|I_N - I| \leq \frac{2\pi M}{e^{aN}-1} \quad , |I_N - I| \leq \frac{4\pi M}{e^{aN}-1} \tag{4.39}$$

We could think that in principle, these results suggest that the larger the value of a, the better error bound we get. However, this is not the case (in general) since the quantity M that bounds the function is dependent on a generally. We can illustrate this with an example. Consider the function $f(\theta) = e^{\cos(\theta)}$. *This function is analytic in all the complex plane and therefore we can apply then theorem 4.2. We can use theorem 4.1 as well, lets stick to 4.2 for convenience.* f *is analytic in the strip* $-a < Im\theta < a$ *for all* $a > 0$. *We also note that in this strip, we can bound* f *as:*

$$|e^{\cos(\theta)}| \leq |e^{\cos(\pm ia)}| = e^{\cosh(a)} \tag{4.40}$$

Therefore, direct application of theorem 4.32 yields:

$$|I_N - I| \leq \frac{4\pi e^{\cosh(a)}}{e^{aN} - 1} \tag{4.41}$$

Which illustrates that the error bound can be dependent on a via M. In this situation, the largest a does not give the best error bound and to obtain the best a, we have to optimize

$$r(a) = \frac{e^{\cosh(a)}}{e^{aN} - 1} \tag{4.42}$$

With derivative

$$r'(a) = \frac{e^{\cosh(a)} \cdot \sinh(a) \cdot (e^{aN} - 1) - e^{\cosh(a)} \cdot N \cdot e^{aN}}{(e^{aN} - 1)^2} = \frac{e^{\cosh(a)} \left[\sinh(a) \cdot (e^{aN} - 1) - N \cdot e^{aN}\right]}{(e^{aN} - 1)^2} \tag{4.43}$$

Forcing this derivative to be 0 implies:

$$\begin{gathered}\left[\sinh(a) \cdot (e^{aN} - 1) - N \cdot e^{aN}\right] = 0 \Longrightarrow e^{aN}(\sinh(a) - N) = \sinh(a) \Longrightarrow e^{aN} = \frac{\sinh(a)}{\sinh(a) - N} \\ \Longrightarrow a = \frac{1}{N} \ln\left(\frac{\sinh(a)}{\sinh(a) - N}\right)\end{gathered} \tag{4.44}$$

And solving this equation numerically for a would give the a that minimizes the error bound.

These last results show that the trapezoidal rule is really powerful. In fact, the trapezoidal rule can give an exact result when integrating trigonometric polynomials:

Definition 4.1. *Let* $n \in \mathbb{N}$. *We call the function:*

$$f(\theta) = \sum_{j=-n}^{j=n} c_j e^{ij\theta} \quad c_i \in \mathbb{C} \tag{4.45}$$

a trigonometric polynomial of degree n. The values of c_j *are not necessarily non-zero.*

And with this definition, we can write the following corollary.

Corollary 4.1. *If v is a trigonometric polynomial of degree n (a linear combination of* $e^{-in\theta}, \ldots, e^{in\theta}$*), the N-point trapezoidal rule (3.2) is exact for all* $N > n$.

Proof. We start the same way as we did when proving theorem 4.1. The Fourier expansion of the trigonometric polynomial of degree n is itself:

$$v(\theta) = \sum_{j=-n}^{j=n} c_j e^{ij\theta} \tag{4.46}$$

Where the coefficients are:

$$c_j = \frac{1}{2\pi} \int_0^{2\pi} e^{-ij\theta} v(\theta) d\theta \tag{4.47}$$

And we also have that

$$I = \int_0^{2\pi} v(\theta)d\theta = 2\pi c_0 \tag{4.48}$$

Using equation (4.46) with (4.2):

$$\begin{aligned} I_N = \frac{2\pi}{N}\sum_{k=1}^{N} v\left(\frac{2\pi k}{N}\right) = \frac{2\pi}{N}\sum_{k=1}^{N}\sum_{j=-n}^{j=n} c_j e^{\frac{ij2\pi k}{N}} = \frac{2\pi}{N}\sum_{j=-n}^{j=n} c_j \sum_{k=1}^{N} e^{\frac{ij2\pi k}{N}} \\ = 2\pi \sum_{j=-n}^{j=n} c_j \delta_{j,mN} \end{aligned} \tag{4.49}$$

Where the last equality comes from the fact that the last sum gives N if j is a multiple of N and zero otherwise. At this point, we can use the fact that $N > n$. By doing so, the only index j ($j = -n, \ldots, n$) that contributes is c_0 because the only multiple of N in $[-n, n]$ is zero. Consequently:

$$I_N = 2\pi \sum_{j=-n}^{j=n} c_j \delta_{j,mN} = 2\pi c_0 \tag{4.50}$$

And therefore, we trivially obtain:

$$I_N - I = 2\pi c_0 - 2\pi c_0 = 0 \tag{4.51}$$

□

Corollary 4.2. *As we mentioned at the beginning of this section, assuming 2π-periodic functions is not a loss of generality since we can always use a change of variable on a T-periodic function to make it 2π-periodic. Let v be a T-periodic function analytic on any of the H sets defined in the previous theorems. Suppose $|v(t)| \leq M$ in H, where M is a nonnegative real number. The integral we would like to calculate is:*

$$I = \int_0^T v(t)dt \overset{t=\frac{\theta T}{2\pi}}{=} \frac{T}{2\pi}\int_0^{2\pi} v\left(\frac{\theta T}{2\pi}\right) d\theta \tag{4.52}$$

The function $v\left(\frac{\theta T}{2\pi}\right)$ is 2π-periodic and it satisfies the hypothesis on a set H' defined consequently after the change of variable. Let N be the number of nodes in the trapezoidal approximation (4.2) of (4.1). We can repeat the same proof in theorems 4.1 and 4.2 for the right-hand side of (4.52) and we obtain:

$$\begin{aligned} |I_N - I| &\leq \frac{TM}{e^{aN} - 1} \quad \textit{For T-periodic, bounded and analytic functions in a strip (Analog to theorem 4.1)} \\ |I_N - I| &\leq \frac{2TM}{e^{aN} - 1} \quad \textit{For T-periodic, bounded and analytic functions in a half plane (Analog to theorem 4.2)} \end{aligned} \tag{4.53}$$

4.2 Error of the trapezoidal rule in $\mathbb{R}$

There are many situations in which we would like to compute integrals over the real line. In these situations, we usually have a function w (not necessarily a real function) which decays sufficiently fast[2], and the integral

$$I = \int_{-\infty}^{\infty} w(x)dx \tag{4.54}$$

would have a trapezoidal approximation given by:

$$I_h = h \sum_{k=-\infty}^{k=\infty} w(kh) \tag{4.55}$$

[2]Sufficiently fast so that the integral converges absolutely.

Where h is just the step of the trapezoidal rule. Of course, in practice this expression has to be truncated at some point. We introduce the following notation to denote the truncated approximation:

$$I_h^{[n^-,n^+]} = h \sum_{k=-n^-}^{k=n^+} w(kh) \tag{4.56}$$

With this notation, we can introduce the following definitions.

Definition 4.2. ***(Discretization error [7][sections 5 and 6])***
We call the quantity $|I_h - I|$ the discretization error of I.

Definition 4.3. ***(Truncation error [7][sections 5 and 6])***
We call the quantity $|I_h - I_h^{[n^-,n^+]}|$ the truncation error of I.

In the following, we will introduce results that address the convergence of the discretization error. The truncation error will be adressed in the next section. We will follow the scheme in [7][sections 5 and 6], starting with the following theorem.

Theorem 4.3. ***(Exponential decrease of the error for rapidly decaying analytic functions in a strip)***
Let w be an analytic function in the strip defined by the set

$$H = \{z \in \mathbb{C} \mid -a < Im(z) < a\} \tag{4.57}$$

for some $a > 0$. Suppose further that $w(x) \to 0$ uniformly as $|x| \to \infty$ in the strip, and for some M, it satisfies

$$\int_{-\infty}^{\infty} |w(x+ib)|dx \leq M \tag{4.58}$$

for all $b \in (-a, a)$. Then, for any $h > 0$, I_h as defined by (4.55) exists and satisfies

$$|I_h - I| \leq \frac{1}{e^{2\pi b/h} - 1} \left[\int_{-\infty}^{\infty} |w(x+ib)|dx + \int_{-\infty}^{\infty} |w(x-ib)|dx \right] \leq \frac{2M}{e^{2\pi b/h} - 1} \quad \text{for all } b \in (-a, a) \tag{4.59}$$

$$|I_h - I| \leq \frac{2M}{e^{2\pi a/h} - 1} \tag{4.60}$$

Proof. We can start by considering the function

$$m(z) = -\frac{i}{2} \cot\left(\frac{\pi z}{h}\right) \tag{4.61}$$

This function has simple poles at points in which $\sin\left(\frac{\pi z}{h}\right)$ vanishes i.e when $z = kh$ where k is an integer number. The residues of this function at the singularities are:

$$\begin{aligned} Res(m, z=0) &= \lim_{z\to 0} -\frac{iz}{2} \frac{\cos\left(\frac{\pi z}{h}\right)}{\sin\left(\frac{\pi z}{h}\right)} = -\frac{ih}{2\pi} = \frac{h}{2\pi i} \\ Res(m, z=\pm kh) &= \lim_{z\to \mp kh} -\frac{i(z \mp kh)}{2} \frac{\cos\left(\frac{\pi z}{h}\right)}{\sin\left(\frac{\pi z}{h}\right)} \overset{\text{l'Hôpital rule 8.9}}{=} \\ = -\frac{i}{2} \lim_{z\to \mp kh} \frac{\cos\left(\frac{\pi z}{h}\right) - \frac{\pi}{h}\sin\left(\frac{\pi z}{h}\right)(z \mp nh)}{\frac{\pi}{h}\cos\left(\frac{\pi z}{h}\right)} &= -\frac{i}{2}\frac{h}{\pi} - \lim_{z\to \mp kh} \frac{\sin\left(\frac{\pi z}{h}\right)(z \mp nh)}{\cos\left(\frac{\pi z}{h}\right)} = \frac{h}{2\pi i} \end{aligned} \tag{4.62}$$

Since $w(x)$ is analytic (and continuous) in H:

$$\begin{aligned} Res(m \cdot w, 0) &= \lim_{z\to 0} z m(z) w(z) = \left(\lim_{z\to 0} z m(z)\right)\left(\lim_{z\to 0} w(z)\right) \overset{\text{Using (4.62)}}{=} \frac{h}{2\pi i} w(0) \\ Res(m \cdot w, \pm kh) &= \lim_{z\to \mp kh} (z \mp kh) m(z) w(z) = \left(\lim_{z\to \mp kh} (z \mp kh) m(z)\right)\left(\lim_{z\to \mp kh \to 0} w(z)\right) \overset{\text{Using (4.62)}}{=} \\ &= \frac{h}{2\pi i} w(\pm kh) \end{aligned} \tag{4.63}$$

Knowing the residues of $m \cdot w$ is very useful to use the residue theorem 8.5 with the product of these two functions. Let us define the path Γ as the rectangle defined by the vertices $(n^+ + 1/2)h + bi$, $-(n^- + 1/2)h + bi$, $-(n^- + 1/2)h - bi$, $(n^+ + 1/2)h - bi$ with $b < a$. This rectangle has in its interior the poles of $m(z)$ ranging from $-n^- h$ to $n^+ h$. Then, if we use the residue theorem 8.5 with the integral of $m \cdot w$ along this path, we obtain:

$$\int_\Gamma m(z)w(z)dz = h \sum_{k=n^-}^{k=-n^+} w(kh) \overset{(4.56)}{=} I_h^{[n^-,n^+]} \tag{4.64}$$

This chain of equalities will be useful to compute I_h by taking the limit $n^+, n^- \longrightarrow \infty$ and via the integral along Γ.

Now, let us call Γ^+ and Γ^- the upper and lower half of Γ respectively with positive orientation. Using Cauchy's integral theorem 8.6, we have:

$$\begin{aligned} \int_{-(n^-+1/2)h}^{(n^++1/2)h} w(z)dz + \int_{\Gamma^+} w(z)dz = 0 \Longrightarrow \int_{-(n^-+1/2)h}^{(n^++1/2)h} w(z)dz = -\int_{\Gamma^+} w(z)dz \\ -\int_{-(n^-+1/2)h}^{(n^++1/2)h} w(z)dz + \int_{\Gamma^-} w(z)dz = 0 \Longrightarrow \int_{-(n^-+1/2)h}^{(n^++1/2)h} w(z)dz = \int_{\Gamma^+} w(z)dz \end{aligned} \tag{4.65}$$

With these equalities, we can write:

$$\begin{aligned} I_h^{[n^-,n^+]} - \int_{-(n^-+1/2)h}^{(n^++1/2)h} m(z)w(z)dz \overset{(4.64)}{=} \int_\Gamma m(z)w(z)dz - \int_{-(n^-+1/2)h}^{(n^++1/2)h} m(z)w(z)dz \overset{(4.65)}{=} \\ = \int_{\Gamma^+} m(z)w(z)dz + \int_{\Gamma^-} m(z)w(z)dz + \frac{1}{2}\int_{\Gamma^+} w(z)dz - \frac{1}{2}\int_{\Gamma^+} w(z)dz = \\ = \int_{\Gamma^+} w(z)\left(\frac{1}{2} + m(z)\right)dz + \int_{\Gamma^-} w(z)\left(-\frac{1}{2} + m(z)\right)dz \overset{(4.61)}{=} \\ = \frac{1}{2}\int_{\Gamma^+} w(z)\left(1 - i\cot\left(\frac{\pi z}{h}\right)\right)dz - \frac{1}{2}\int_{\Gamma^-} w(z)\left(-\frac{1}{2} + m(z)\right)dz \overset{1 \pm i\cot\left(\frac{\pi z}{h}\right) = \frac{2}{1-e^{\pm\frac{2\pi i z}{h}}}}{=} \\ = \int_{\Gamma^+} \frac{w(z)}{1 - e^{-\frac{2\pi i z}{h}}}dz - \int_{\Gamma^-} \frac{w(z)}{1 - e^{\frac{2\pi i z}{h}}}dz = \text{Dividing the integral in segments} = \\ = \int_0^b \frac{w\left(\left(n^+ + \frac{1}{2}\right)h + iy\right)}{1 - e^{-\frac{2\pi i}{h}\left(\left(n^+ + \frac{1}{2}\right)h + iy\right)}}dy + \int_{\left(n^+ + \frac{1}{2}\right)h}^{-\left(n^- + \frac{1}{2}\right)h} \frac{w\left(x + bi\right)}{1 - e^{-\frac{2\pi i}{h}(x + bi)}}dx + \int_b^0 \frac{w\left(\left(n^- + \frac{1}{2}\right)h + yi\right)}{1 - e^{-\frac{2\pi i}{h}\left(\left(n^- + \frac{1}{2}\right)h + yi\right)}}dy \\ - \int_0^b \frac{w\left(\left(n^- + \frac{1}{2}\right)h + yi\right)}{1 - e^{\frac{2\pi i}{h}\left(\left(n^- + \frac{1}{2}\right)h + yi\right)}}dy - \int_{-\left(n^- + \frac{1}{2}\right)h}^{\left(n^+ + \frac{1}{2}\right)h} \frac{w\left(x - bi\right)}{1 - e^{\frac{2\pi i}{h}(x - bi)}}dx - \int_b^0 \frac{w\left(\left(n^+ + \frac{1}{2}\right)h + yi\right)}{1 - e^{\frac{2\pi i}{h}\left(\left(n^+ + \frac{1}{2}\right)h + yi\right)}}dy \end{aligned} \tag{4.66}$$

The next step would be taking the limit $\lim_{n^+, n^- \longrightarrow \infty}$ in (4.66). It is essential to realize that the integrals corresponding to vertical segments vanish in this limit. To verify it, note that:

$$\begin{aligned} 1 - e^{-\frac{2\pi i}{h}\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)} = 1 - e^{-2\pi i n^\pm} e^{-i\pi} e^{\frac{2\pi y}{h}} \overset{e^{-2\pi i n^\pm} = 1}{=} 1 + e^{\frac{2\pi y}{h}} \geq 2 \\ 1 - e^{\frac{2\pi i}{h}\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)} = 1 - e^{2\pi i n^\pm} e^{i\pi} e^{\frac{2\pi y}{h}} \overset{e^{2\pi i n^\pm} = 1}{=} 1 + e^{-\frac{2\pi y}{h}} \geq 1 \end{aligned} \tag{4.67}$$

And therefore, we can do the following for the integrals on the vertical segments:

$$\begin{aligned} \left|\int_0^b \frac{w\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)}{1 - e^{-\frac{2\pi i}{h}\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)}}dy\right| \overset{(4.67)}{\leq} \frac{1}{2}\int_0^b \left|w\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)\right|dy \underset{n^\pm \longrightarrow \infty}{\longrightarrow} 0 \\ \left|\int_0^b \frac{w\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)}{1 - e^{\frac{2\pi i}{h}\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)}}dy\right| \overset{(4.67)}{\leq} \int_0^b \left|w\left(\left(n^\pm + \frac{1}{2}\right)h + iy\right)\right|dy \underset{n^\pm \longrightarrow \infty}{\longrightarrow} 0 \end{aligned} \tag{4.68}$$

Where the last limit is a consequence of the hypothesis that $w(x) \to 0$ uniformly as $|x| \to \infty$ in the strip H. We can also bound the contributions on the horizontal segments thanks to the inverse triangle inequality:

$$\begin{aligned}\left|1-e^{-\frac{2\pi i}{h}(x+bi)}\right| &= \left|e^{-\frac{2\pi ix}{h}}e^{\frac{2\pi b}{h}}-1\right| \overset{|x-y|\geq||x|-|y||}{\geq} \left|\left|e^{-\frac{2\pi ix}{h}}\right|e^{\frac{2\pi b}{h}}-1\right| = \left|e^{\frac{2\pi b}{h}}-1\right| = e^{\frac{2\pi b}{h}}-1 \\ \left|1-e^{\frac{2\pi i}{h}(x-bi)}\right| &= \left|e^{\frac{2\pi ix}{h}}e^{\frac{2\pi b}{h}}-1\right| \overset{|x-y|\geq||x|-|y||}{\geq} \left|\left|e^{\frac{2\pi ix}{h}}\right|e^{\frac{2\pi b}{h}}-1\right| = \left|e^{\frac{2\pi b}{h}}-1\right| = e^{\frac{2\pi b}{h}}-1\end{aligned} \tag{4.69}$$

Finally, by taking absolute values and the limit $\lim_{n^+,n^-\longrightarrow\infty}$ in (4.66):

$$\begin{aligned}&\left|\lim_{n^+,n^-\longrightarrow\infty} I_h^{[n^-,n^+]} - \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+1/2)h}^{(n^++1/2)h} m(z)w(z)dz\right| \overset{(4.56)(4.54)}{=} |I_h - I| \overset{(4.66)(4.68)}{=} \\ &= \left|\lim_{n^+,n^-\longrightarrow\infty}\int_{(n^++\frac{1}{2})h}^{-(n^-+\frac{1}{2})h}\frac{w(x+bi)}{1-e^{-\frac{2\pi i}{h}(x+bi)}}dx - \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x-bi)}{1-e^{\frac{2\pi i}{h}(x-bi)}}dx\right| = \\ &= \left|-\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x+bi)}{1-e^{-\frac{2\pi i}{h}(x+bi)}}dx - \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x-bi)}{1-e^{\frac{2\pi i}{h}(x-bi)}}dx\right| = \\ &= \left|\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x+bi)}{1-e^{-\frac{2\pi i}{h}(x+bi)}}dx + \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x-bi)}{1-e^{\frac{2\pi i}{h}(x-bi)}}dx\right| \leq \\ &\leq \left|\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x+bi)}{1-e^{-\frac{2\pi i}{h}(x+bi)}}dx\right| + \left|\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{w(x-bi)}{1-e^{\frac{2\pi i}{h}(x-bi)}}dx\right| \leq \\ &\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{|w(x+bi)|}{\left|1-e^{-\frac{2\pi i}{h}(x+bi)}\right|}dx + \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{|w(x-bi)|}{\left|1-e^{\frac{2\pi i}{h}(x-bi)}\right|}dx \overset{(4.69)}{\leq} \\ &\leq \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{|w(x+bi)|}{\left|1-e^{-\frac{2\pi i}{h}(x+bi)}\right|}dx + \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}\frac{|w(x-bi)|}{\left|1-e^{\frac{2\pi i}{h}(x-bi)}\right|}dx \overset{(4.69)}{\leq} \\ &\frac{1}{e^{\frac{2\pi b}{h}}-1}\left[\lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}|w(x+bi)|dx + \lim_{n^+,n^-\longrightarrow\infty}\int_{-(n^-+\frac{1}{2})h}^{(n^++\frac{1}{2})h}|w(x-bi)|dx\right] = \\ &= \frac{1}{e^{\frac{2\pi b}{h}}-1}\left[\int_{-\infty}^{\infty}|w(x+bi)|dx + \int_{-\infty}^{\infty}|w(x-bi)|dx\right]\end{aligned} \tag{4.70}$$

And we have almost shown (4.59). In order to simplify more, we need to get rid of the integral with $|w(x-bi)|$. To do so, we can just use the hypothesis given by (4.58), which applies for all $b \in (-a,a)$ and therefore it applies to $-b$ as well. Consequently, we can proceed with (4.70):

$$|I_h - I| \leq \frac{1}{e^{\frac{2\pi b}{h}}-1}\left[\int_{-\infty}^{\infty}|w(x+bi)|dx + \int_{-\infty}^{\infty}|w(x-bi)|dx\right] \overset{(4.58)}{\leq} \frac{2M}{e^{\frac{2\pi b}{h}}-1} \tag{4.71}$$

And the proof of (4.59) is now complete. In order to prove (4.60), we just need to realize that the previous inequality can be obtained for all $0 < b < a$. Therefore:

$$|I_h - I| \leq \frac{2M}{e^{\frac{2\pi a}{h}}-1} \tag{4.72}$$

That proves (4.60) and the proof is finally complete. □

Corollary 4.3. *If we add the hypothesis to theorem 4.3 that $w(z)$ is real for real z, then we have:*

$$|I_h - I| \leq \frac{2}{e^{2\pi b/h}-1}\int_{-\infty}^{\infty}|w(x+ib)|dx \leq \frac{2M}{e^{2\pi b/h}-1} \quad \textit{for all } b \in (-a,a) \tag{4.73}$$

Proof. We can start from (4.70):

$$|I_h - I| \leq \frac{1}{e^{\frac{2\pi b}{h}} - 1} \left[\int_{-\infty}^{\infty} |w(x + bi)| dx + \int_{-\infty}^{\infty} |w(x - bi)| dx \right] \tag{4.74}$$

And we aim to try to get rid of the term with $|w(x - ib)|$. The extra hypothesis of $w(z)$ for real z is key to apply the Schwarz reflection principle 8.10. With the assumptions we have, we know that w is C-analytic in the upper half plane region:

$$D = \{x + iy \mid 0 \leq y \leq b\} \tag{4.75}$$

And real on the real axis. Therefore, we can use Schwarz reflection principle 8.10 to obtain that $w(x - yi) = \overline{w(x + yi)} \implies |w(x - yi)| = |w(x + yi)|$ And therefore, we can simplify the last inequality as:

$$|I_h - I| \leq \frac{2}{e^{2\pi b/h} - 1} \int_{-\infty}^{\infty} |w(x + ib)| dx \tag{4.76}$$

And that ends the proof. □

Remark 4.3. ***(Does the largest b give the best error bound? [3][pages 151-152])*** *We can introduce a remark similar to the one in remark 4.2. Corollary 4.3 gives a bound for the discretization error as:*

$$|I_h - I| \leq \frac{2}{e^{2\pi b/h} - 1} \int_{-\infty}^{\infty} |w(x + ib)| dx \tag{4.77}$$

If we naively ignore the dependence on b of the integral, we could think that the larger the value of b, the better bound of the integral we have. Unfortunately, this is not always the case, and we can exemplify it with the function

$$f(x) = e^{-\alpha x^2} g(x) \quad w > 0 \tag{4.78}$$

Where $g(x)$ is an analytic function over the strip $-a < Imz < a$ and real for real x. Now, note that:

$$\int_{-\infty}^{\infty} |f(x + ib)| dx = e^{\alpha b^2} \int_{-\infty}^{\infty} e^{-\alpha x^2} |g(x + ib)| dx \leq e^{\alpha b^2} \int_{-\infty}^{\infty} |g(x + ib)| dx \tag{4.79}$$

For this example, we need to assume that all these integrals are convergent and that the value of $\int_{-\infty}^{\infty} |g(x + ib)| dx$ is independent of b. If we do so, using corollary (4.3), we get:

$$|I_h - I| \leq \frac{2e^{\alpha b^2}}{e^{2\pi b/h} - 1} \int_{-\infty}^{\infty} |g(x + ib)| dx \tag{4.80}$$

And with the assumptions we are taking, the bound of the error is going to depend only on:

$$\beta(b) = \frac{e^{\alpha b^2}}{e^{2\pi b/h} - 1} \overset{b \gg 1}{\approx} e^{\alpha b^2 - \frac{2\pi b}{h}} \tag{4.81}$$

Hence, the best error bound (approximately) can be estimated by minimizing

$$\gamma(b) = \alpha b^2 - \frac{2\pi b}{h} \tag{4.82}$$

Which has a minimum at:

$$b = \frac{\pi}{\alpha h} \tag{4.83}$$

And if this b is valid in the sense that the function $f(x)$ is analytic in:

$$\{z = x + iy \mid x \in \mathbb{R}, -\frac{\pi}{\alpha h} \leq y \leq \frac{\pi}{\alpha h}\} \tag{4.84}$$

Then, the value $b = \frac{\pi}{wh}$ is the best error bound, illustrating that sometimes the best error bound is not given by the largest b possible. Again, this happens when the value of the integral is dependent on b.

There are several versions of theorem 4.3 for different sets in which the function w is analytic and decays rapidly enough. The proofs are very similar and they do put more interesting mathematics on the table. Therefore, we will write the theorems and a sketch of the proof.

Theorem 4.4. ***(Exponential decrease of the error for rapidly decaying analytic functions in a half-plane)*** *Let w be an analytic function in the half-plane defined by the set*

$$H = \{z \in \mathbb{C} \mid -a < Im(z)\} \tag{4.85}$$

for some $a > 0$. Suppose further that $w(x) \to 0$ uniformly as $|x| \to \infty$ in half-plane, and for some M, it satisfies

$$\int_{-\infty}^{\infty} |w(x+ib)|dx \leq M \tag{4.86}$$

for all $b > -a$. Then, for any $h > 0$, I_h as defined by (4.55) exists and satisfies

$$\begin{aligned} |I_h - I| \leq= \frac{1}{e^{\frac{2\pi b_+}{h}} - 1} \int_{-\infty}^{\infty} |w\left(x + b_+ i\right)| dx + \frac{1}{e^{\frac{2\pi b_-}{h}} - 1} \int_{-\infty}^{\infty} |w\left(x - b_- i\right)| dx \\ |I_h - I| \leq \frac{M}{e^{\frac{2\pi b_+}{h}} - 1} + \frac{M}{e^{\frac{2\pi b_-}{h}} - 1} \end{aligned} \tag{4.87}$$

Where $b_+ > 0$ and $0 < b_- < a$. Additionally:

$$|I_h - I| \leq \frac{M}{e^{\frac{2\pi a}{h}} - 1} \tag{4.88}$$

Proof. The proof of this theorem is very similar to the one for 4.3. The main difference is the fact that H is a upper half plane. This allows to modify the rectangle Γ, setting its vertices to $(n^+ + 1/2)h + b_+ i$, $-(n^- + 1/2)h + b_+ i$, $-(n^- + 1/2)h - b_- i$, $(n^+ + 1/2)h - b_- i$ with $b_+ > 0$, $0 < b_- < a$. With this new rectangle and following the same process as in (4.70), we obtain:

$$|I_h - I| \leq \frac{1}{e^{\frac{2\pi b_+}{h}} - 1} \int_{-\infty}^{\infty} |w\left(x + b_+ i\right)| dx + \frac{1}{e^{\frac{2\pi b_-}{h}} - 1} \int_{-\infty}^{\infty} |w\left(x - b_- i\right)| dx \tag{4.89}$$

Using the fact that $\int_{-\infty}^{\infty} |w(x+ib)|dx \leq M$ for all $b > -a$ on the previous inequality:

$$|I_h - I| \leq \frac{M}{e^{\frac{2\pi b_+}{h}} - 1} + \frac{M}{e^{\frac{2\pi b_-}{h}} - 1} \tag{4.90}$$

But this inequality can be obtained for all $b_+ > 0$ and for all $b < a$, consequently:

$$|I_h - I| \leq \frac{M}{e^{\frac{2\pi a}{h}} - 1} \tag{4.91}$$

□

Theorem 4.5. ***(Exponential decrease of the error for rapidly decaying analytic functions in an asymmetric strip)*** *Let w be an analytic function in the strip defined by the set*

$$H = \{z \in \mathbb{C} \mid -a_- < Im(z) < a_+\} \tag{4.92}$$

for some $a^-, a^+ > 0$. Suppose further that $w(x) \to 0$ uniformly as $|x| \to \infty$ in the strip, and for some M, it satisfies

$$\int_{-\infty}^{\infty} |w(x - ib_-)|dx \leq M_-, \quad \int_{-\infty}^{\infty} |w(x + ib_+)|dx \leq M_+ \tag{4.93}$$

for all $b_- \in (0, a_-)$ and $b_+ \in (0, a_+)$. Then, for any $h > 0$, I_h as defined by (4.55) exists and satisfies

$$|I_h - I| \leq \frac{M_+}{e^{2\pi a_+/h} - 1} + \frac{M_-}{e^{2\pi a_-/h} - 1} \tag{4.94}$$

Proof. The proof of this theorem is very similar to the one for 4.3. The main difference is the fact that H is now an asymmetric strip. This allows to modify the rectangle Γ, setting its vertices to $(n^+ + 1/2)h + b_+ i$, $-(n^- + 1/2)h + b_+ i$, $-(n^- + 1/2)h - b_- i$, $(n^+ + 1/2)h - b_- i$ with $a_+ > b_+ > 0$, $0 < b_- < a$. With this new rectangle and following the same process as in (4.70), we obtain:

$$|I_h - I| \leq \frac{1}{e^{\frac{2\pi b_+}{h}} - 1} \int_{-\infty}^{\infty} |w(x + b_+ i)| dx + \frac{1}{e^{\frac{2\pi b_-}{h}} - 1} \int_{-\infty}^{\infty} |w(x - b_- i)| dx \tag{4.95}$$

Using the fact that $\int_{-\infty}^{\infty} |w(x + ib_+)| dx \leq M_+$ for all $b_+ \in (0, a_+)$ and $\int_{-\infty}^{\infty} |w(x - ib_-)| dx \leq M_-$ for all $b_- \in (0, a_-)$ on the previous inequality:

$$|I_h - I| \leq \frac{M_+}{e^{\frac{2\pi b_+}{h}} - 1} + \frac{M_-}{e^{\frac{2\pi b_-}{h}} - 1} \tag{4.96}$$

But this inequality can be obtained for all $a_+ > b_+ > 0$ and for all $0 < b_- < a_-$, consequently:

$$|I_h - I| \leq \frac{M_+}{e^{\frac{2\pi a_+}{h}} - 1} + \frac{M_-}{e^{\frac{2\pi a_-}{h}} - 1} \tag{4.97}$$

□

Remark 4.4. ***(Practical tips in implementations [7][Section 4], [3][Page 152])*** *The theorems we have provided in this section analyze the convergence of the discretization error 4.2 which involves the term I_h in (4.55). In practice, this term cannot be computed due to its infinite amount of terms. The usual approach is to truncate (4.55), which leads to (4.56). The number of nodes is usually chosen by analyzing how small the integrand gets as a function of x with the ultimate goal of obtaining accuracy while avoiding an excessive number of function evaluations. After fixing the number of nodes, the step h remains as a parameter to choose. We will explain how to choose h conveniently in the following section. Other more complex approaches include keeping track of the integrals over subintervals. This way, one can stop increasing the interval of integrations when these contributions start being less than a threshold.*

4.3 Control of the truncation error

Theorems 4.3, 4.4 and 4.5 state the convergence of the discretization error $|I_h - I|$. Unfortunately, I_h (4.55) cannot be computed and we have to approximate it with its truncated approximation $I_h^{[n^-, n^+]}$ given by (4.56). In this section, we will consider the truncation error 4.3. Firstly, we can give a bound for the truncation error [7][Section 4]. For simplicity, let us take $n^- = n^+ = n$:

$$|I_h^{[n,n]} - I| \leq \underbrace{|I_h - I|}_{\text{Discretization error 4.2}} + \underbrace{|I_h^{[n,n]} - I_h|}_{\text{Truncation error 4.3}} \tag{4.98}$$

Remark 4.5. ***(The discretization error decreases with decreasing h while the truncation error usually increases with decreasing h)*** *From theorems 4.3, 4.4 and 4.5, we know that the discretization error gets smaller with decreasing step. On the other hand, if we consider a fixed number of nodes n, the truncation error will likely increase with decreasing h unless the support of the function has an extremely short length. We can get an intuition about it by imagining the Gaussian distribution:*

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \tag{4.99}$$

Where μ is the mean and σ^2 is the variance. For simplicity, let us set $\mu = 0$ and analyze how the truncation error would behave with the variance. Consider a high value of the variance and a fixed number of nodes n. To have a low truncation error, we would need to cover as much length as possible of the set

$$\{x \in \mathbb{R} \mid f(x) \geq \epsilon\} \tag{4.100}$$

Where ϵ is a nonzero positive small real number. However, the number of nodes is limited and so is the length of the set we can cover with the trapezoidal rule. If we start making the step smaller and smaller, we are covering just a small length of the set and therefore, we are leaving out of the truncation an important contribution of the total integral. This is clearly exemplified in figure 4.1 for the Gaussian distribution with $\mu = 0$ and $\sigma^2 = 100$ (read caption).

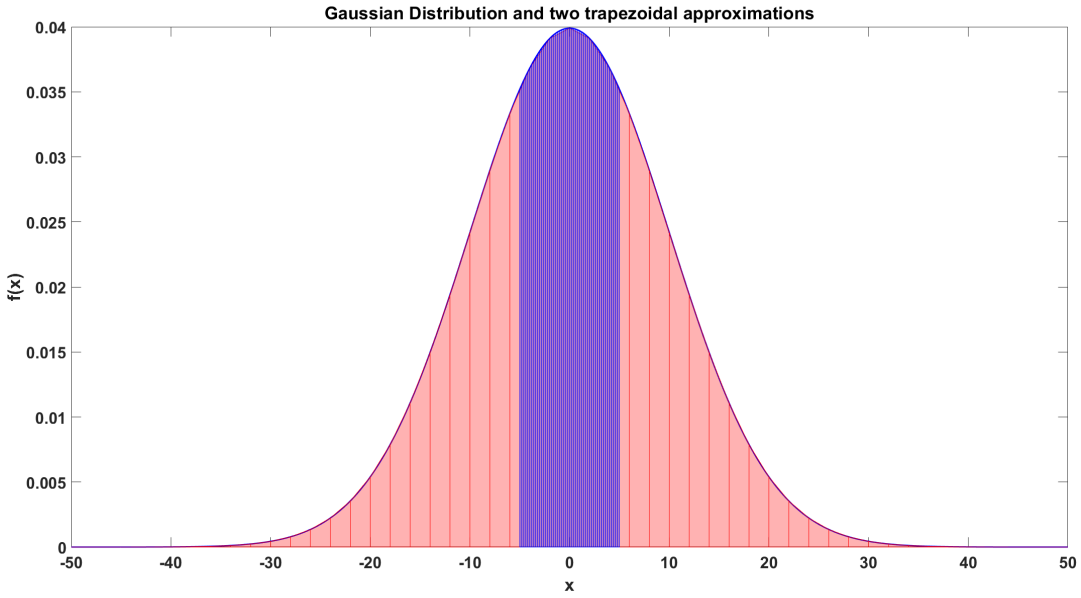

Figure 4.1: Plot of the Gaussian distribution with $\mu = 0$ and $\sigma^2 = 100$. The shaded areas represent the trapezoidal approximations given by $I_1^{[25,25]}$ (red) and $I_{0.1}^{[25,25]}$ (blue). We can clearly see how keeping the same number of nodes while reducing the step leaves an important contribution to the total integral out of the trapezoidal approximation and therefore, the truncation error increases when reducing the step h.

The unfortunate situation in which the discretization error decreases as $h \longrightarrow 0$ and the truncation error increases as $h \longrightarrow 0$ motivates the careful selection of the number of nodes and the step.

The simplest solution to sort this issue is finding a compromise between both errors i.e taking the number of nodes and the step in such a way both errors are equal [15]. To do so, we firstly need to estimate both errors. The discretization error can be estimated with theorems 4.3, 4.4 and 4.5. The truncation error is usually estimated as [7][Page 404], [3][Page 152] :

$$\max_{\{x=-nh\,,\,x=nh\}} |f(x)| \tag{4.101}$$

Remark 4.6. ***(Why does the estimation of the truncation error (4.101) make sense?)*** *This estimation makes sense only under the assumption that the integrand decays fast enough. In fact, faster than $\lambda e^{-|x|}$ for some constant λ. To see this, note that the part of the integral that is not considered is $\int_{nh}^{\infty} f(x)dx + \int_{-\infty}^{-nh} f(x)dx$. The function $g(x) = \frac{|f(sign(x)nh)|}{e^{-nh}} e^{-|x|}$ is such that $g(\pm nh) = f(\pm nh)$. As a result, if $f(x)$ decays faster than $\lambda e^{-|x|}$ in such a way that $|f(x)| \leq g(x)$ in $[nh, \infty) \cup (-\infty, -nh]$, we have:*

$$\begin{aligned} \left| \int_{nh}^{\infty} f(x)dx \right| &\leq \int_{nh}^{\infty} |f(x)|dx \leq \int_{nh}^{\infty} g(x)dx = \int_{nh}^{\infty} \frac{|f(nh)|}{e^{-nh}} e^{-x}dx = |f(nh)| \\ \left| \int_{-\infty}^{-nh} f(x)dx \right| &\leq \int_{-\infty}^{-nh} |f(x)|dx \leq \int_{-\infty}^{-nh} g(x)dx = \int_{-\infty}^{-nh} \frac{|f(-nh)|}{e^{-nh}} e^{x}dx = |f(-nh)| \end{aligned} \tag{4.102}$$

And these inequalities imply that:

$$|I_h^{[n,n]} - I_h| = \left|\int_{nh}^{\infty} f(x)dx + \int_{-\infty}^{-nh} f(x)dx\right| \leq \left|\int_{nh}^{\infty} f(x)dx\right| + \left|\int_{-\infty}^{-nh} f(x)dx\right| \overset{(4.102)}{\leq} \\ \leq |f(nh)| + |f(-nh)| \leq 2 \max_{\{x=-nh\,,\,x=nh\}} |f(x)| \tag{4.103}$$

The factor 2 in the previous inequality is absent in (4.101) and hence, the estimation of the truncation error is just a more optimistic bound than the one in (4.103).

Now that we can estimate the truncation error, we can exemplify how to balance the discretization error and the truncation error to make them more or less equal. This balancing can be done by fixing the step and choosing the number of nodes or the other way around although it makes more sense to fix the number of nodes because that way we fix beforehand the computational cost of the calculation. We will proceed by fixing the number of nodes in the following examples. Consider the functions

$$f_1(x) = \frac{e^{-x\tanh(x)}}{1+x^2} \tag{4.104}$$

$$f_1(x) = \frac{e^{-x^2}}{\sqrt{1+x^2}} \tag{4.105}$$

$$f_1(x) = e^{-x^2} \tag{4.106}$$

The functions $f_1(x)$ and $f_2(x)$ have singularities at $x = \pm i$ and therefore they are analytic in the strip given by $-1 < Im(z) < 1$. Using theorem 4.3, a bound for their discretization error is $\frac{2M}{e^{\frac{2\pi}{h}}-1}$ (M is different for each function) but for the sake of simplicity let us write it as $\mathcal{O}(e^{-s\frac{2\pi}{h}})$. Using (4.101), the truncation error of $f_1(x)$ is $\frac{e^{-nh}e^{\tanh(nh)}}{1+n^2h^2}$ which can be roughly written as $\mathcal{O}(e^{-nh})$ since $|e^{\tanh(nh)}| \leq e$. Making the discretization error and the truncation error more or less equal means setting:

Optimal step for $f_1(x)$

$$\mathcal{O}\left(e^{-\frac{2\pi}{h}}\right) = \mathcal{O}\left(e^{-nh}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \sqrt{\frac{2\pi}{n}} \tag{4.107}$$

And this step h is the optimal step in the sense that balances the discretization error and the truncation error. If we proceed the same way with $f_2(x)$, we obtain a truncation error $\mathcal{O}\left(e^{-n^2h^2}\right)$ and the optimal step is then:

Optimal step for $f_2(x)$

$$\mathcal{O}\left(e^{-\frac{2\pi}{h}}\right) = \mathcal{O}\left(e^{-n^2h^2}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \left(\frac{2\pi}{n^2}\right)^{1/3} \tag{4.108}$$

The function $f_3(x)$ is analytic in the whole complex plane and the obtention of its discretization error has to be obtained following similar guidelines to those in remark 4.3. Firstly, we shall use theorem 4.4 to obtain that the bound for the discretization error is $\frac{M}{e^{\frac{2\pi a}{h}}-1}$. We can get the explicit dependence of M on a by noting that:

$$\int_{-\infty}^{\infty} |e^{-(x-ai)^2}|dx = \int_{-\infty}^{\infty} |e^{-x^2}e^{a^2}e^{2axi}|dx = e^{a^2}\int_{-\infty}^{\infty} e^{-x^2}dx = e^{a^2}\sqrt{\pi} \tag{4.109}$$

And consequently, the bound is $\frac{e^{a^2}\sqrt{\pi}}{e^{\frac{2\pi a}{h}}-1}$ which can be roughly written as $\mathcal{O}(e^{a^2-2\pi a/h})$. The error is minimized for $a = \pi/h$ and then, the discretization error is roughly $\mathcal{O}(e^{-s\frac{\pi^2}{h^2}})$. The estimation of the truncation error is trivially $\mathcal{O}(e^{-n^2h^2})$ and finally, the optimal step is:

Optimal step for $f_3(x)$

$$\mathcal{O}\left(e^{-\frac{\pi^2}{h^2}}\right) = \mathcal{O}\left(e^{-n^2h^2}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \sqrt{\frac{\pi}{n}} \tag{4.110}$$

With the steps we have followed, the computational cost is fixed because we are obtaining the optimal step in terms of a given number of nodes. The final convergence rate can be obtained simply by pluggin the optimal step in either the discretization error or the truncation error. By doing so, we obtain the following optimal convergence rates:

Optimal convergence rate for $f_1(x)$

$$h = \sqrt{\frac{2\pi}{n}};\ \text{Discretization error: } \mathcal{O}\left(e^{-\frac{2\pi}{h}}\right) \Longrightarrow\ \text{Optimal convergence rate: } \mathcal{O}\left(e^{-\sqrt{2\pi n}}\right) \tag{4.111}$$

Optimal convergence rate for $f_2(x)$

$$h = \left(\frac{2\pi}{n^2}\right)^{1/3};\ \text{Discretization error: } \mathcal{O}\left(e^{-\frac{2\pi}{h}}\right) \Longrightarrow\ \text{Optimal convergence rate: } \mathcal{O}\left(e^{-(2\pi n)^{2/3}}\right) \tag{4.112}$$

Optimal convergence rate for $f_3(x)$

$$h = \sqrt{\frac{\pi}{n}};\ \text{Discretization error: } \mathcal{O}\left(e^{-\frac{\pi^2}{h^2}}\right) \Longrightarrow\ \text{Optimal convergence rate: } \mathcal{O}\left(e^{-\pi n}\right) \tag{4.113}$$

We have introduced in this section several results that show how strong the trapezoidal rule is when integrating over the real line. In fact, these results are much stronger than those for the composite trapezoidal rule in compact intervals. In the next chapter, we will show how we can convert an integral over a compact interval into an integral over the real line through changes of variable in order to exploit the results of the trapezoidal rule in the real line.

Chapter 5

Changes of variables and the doubly exponential formulas [3], [15]

Theorems 4.3, 4.4 and 4.5 in the previous chapter show how strong the trapezoidal rule is when integrating over the whole real line. Then, we can think about transforming any integral over any interval into an integral over the real line by means of a change of variable in order to exploit these theorems.

A change of variable can also be motivated by how fast the integrand decays. The trapezoidal rule would need a huge amount of nodes to approximate a slowly decaying integrand. But by doing a change of variable, we can get an integrand that decays faster, possibly improving the performance of the trapezoidal rule.

These two reasons are the main motivation of the search of convenient changes of variable to transform an integral into another one in which the trapezoidal rule outstands. These ideas were heavily pushed forward by Takahasi and Mori starting in the 1970's [15] [16] although the idea of using changes of variables to speed up the convergence of the trapezoidal rule is attributed to C. Schwartz in 1969 [17]. In this chapter, we will introduce useful changes of variables for the trapezoidal rule following references [3] [7][section 14] [15] [16].

5.1 Using change of variables to move from $[-1,1]$ to $\mathbb{R}$

We will start considering the integral

$$I = \int_{-1}^{1} f(x)dx \tag{5.1}$$

Where $f(x)$ is analytic in $\mathbb{R}$. In this section we will only consider integrals between -1 to 1 although any integral between a and b can always be moved to $[-1,1]$ via the change of variable [16][page 908]

$$\int_a^b f(y)dy \quad \overset{y=\frac{(b-a)x}{2}+\frac{b+a}{2}}{=} \quad \frac{b-a}{2}\int_{-1}^{1} f\left(\frac{(b-a)x}{2}+\frac{b+a}{2}\right)dx \tag{5.2}$$

And hence, there is no loss of generality. We will be analyzing changes of variables $x = \phi(u)$ such that the original integral converts to:

$$\begin{aligned} I &= \int_{-\infty}^{\infty} g(u)du \\ g(u) &= f(\phi(u))\phi'(u) \end{aligned} \tag{5.3}$$

The trapezoidal approximation of the integral is then

$$I_h = h\sum_{k=-\infty}^{\infty} f(\phi(kh))\phi'(kh) \tag{5.4}$$

And we can use now theorems 4.3, 4.4 and 4.5 for the integral of $g(u)$ over the real line.

5.1.1 The erf-rule

When using a change of variable (5.3) (5.4), it would be very convenient that the resulting function $g(u)$ does not have more singularities than those $f(x)$ already has. A change of variable with this feature is the so called erf-rule.

Definition 5.1. ***(Error function** [3][page 154]**)*** *We call Error function the function given by:*

$$\operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-z^2} dz \quad , \quad z \in \mathbb{C} \tag{5.5}$$

Which is an entire function. Note that erf$(+\infty) = 1$ and erf$(-\infty) = -1$.

Definition 5.2. ***(The** erf**-rule** [3][pages 154-156], [7][page 429], [15][page 736]**)*** *Following the notation in (5.3) and definition 5.1, we define the erf-rule as the change of variable*

$$x = \operatorname{erf}(u) \tag{5.6}$$

The erf*-rule converts (5.1) into*

$$I = \int_{-\infty}^{\infty} g(u)du = \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} f(\operatorname{erf}(u)) e^{-u^2} du \tag{5.7}$$

And we also have that

$$|g(u)| \approx e^{-u^2} \quad (u \longrightarrow \pm\infty) \tag{5.8}$$

This change of variable results in the following trapezoidal approximation:

$$I_h = \sum_{k=-\infty}^{\infty} f(\operatorname{erf}(kh)) \exp(-k^2h^2) \tag{5.9}$$

In order to study the optimal convergence rate with the erf-rule, we shall start with the simple case: assuming that $f(x)$ is analytic and therefore has no singularities. Since erf(z) is entire, $f(\operatorname{erf}(u))$ is analytic too and does not have singularities as well. Then, the integrand with the erf-rule, $g(u)$ in (5.7) is an analytic function and we can use the exponential convergence theorems 4.3, 4.4 and 4.5. The discretization error for a function like that in (5.7) was estimated already in remark 4.3, being $\mathcal{O}(e^{-\frac{\pi^2}{h^2}})$. From (5.8) and (4.101), we can estimate the truncation error to be $\mathcal{O}(e^{-n^2h^2})$. As done in the previous chapter, we balance the discretization and truncation error by trying to make them equal. Doing so, we obtain:

Optimal step for the erf-rule (For analytic $f(z)$)

$$\mathcal{O}\left(e^{-\frac{\pi^2}{h^2}}\right) = \mathcal{O}\left(e^{-n^2h^2}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \left(\frac{\pi}{n}\right)^{1/2} \tag{5.10}$$

Optimal convergence rate for the erf-rule (For analytic $f(z)$)

$$h = \left(\frac{\pi}{n}\right)^{1/2}; \text{ Discretization error: } \mathcal{O}\left(e^{-\frac{\pi^2}{h^2}}\right) \Longrightarrow \text{ Optimal convergence rate: } \mathcal{O}\left(e^{-\pi n}\right) \tag{5.11}$$

And this result shows that the erf-rule is indeed a very good choice to integrate entire functions.

The situation gets more complicated when $f(z)$ has a singularity. Obtaining the explicit expression for $u = \operatorname{erf}^{-1}(z)$ is not possible due to the implicit definition of the error function. Therefore, without loss of generality, we write that $u = \operatorname{erf}^{-1}(z)$ maps the singularity z of f into a point u at a distance d of the real axis. This would result in a discretization error given by one of theorems

4.3, 4.4 and 4.5. One way or the other, we can estimate the discretization error to be $\mathcal{O}(e^{-\frac{2\pi d}{h}})$. Once again, the truncation error from (4.101) is $\mathcal{O}(e^{-n^2h^2})$ and we can obtain the optimal convergence rate as usual.

Optimal step for the erf-rule (For $f(z)$ with singularities)

$$\mathcal{O}\left(e^{-\frac{2\pi d}{h}}\right) = \mathcal{O}\left(e^{-n^2h^2}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \left(\frac{2\pi d}{n^2}\right)^{1/3} \tag{5.12}$$

Optimal convergence rate for the erf-rule (For $f(z)$ with singularities)

$$h = \left(\frac{2\pi d}{n^2}\right)^{1/3}; \text{ Discretization error: } \mathcal{O}\left(e^{-\frac{2\pi d}{h}}\right) \Longrightarrow \text{ Optimal convergence rate: } \mathcal{O}\left(e^{-(2\pi d)^{1/3}n^{2/3}}\right) \tag{5.13}$$

Note that the optimal convergence rate is worse that one for the case of $f(z)$ without singularities as we would intuitively think. In the following subsection, we will introduce a different change of variable that will serve to deduce how to obtain an optimal one.

5.1.2 The $\tanh -m$ rule

Definition 5.3. ***(The*** $\tanh -m$ ***rule*** *[15][page 725], [3][page 154], [7][page 428])* *Following the notation in (5.3), we define the* $\tanh -m$ *rule as the change of variable*

$$x = \tanh(u^m) \tag{5.14}$$

The $\tanh -m$ *rule converts (5.1) into*

$$I = \int_{-\infty}^{\infty} g(u)du = \int_{-\infty}^{\infty} f(\tanh(u^m))\frac{mu^{m-1}}{\cosh^2(u^m)}du \tag{5.15}$$

And we also have that

$$|g(u)| \approx e^{-2|u|^m} \quad (u \longrightarrow \pm\infty) \tag{5.16}$$

This change of variable results in the following trapezoidal approximation

$$I_h = h\sum_{n=-\infty}^{\infty} f(\tanh(n^mh^m))\frac{m\,n^{m-1}h^{m-1}}{\cosh^2(n^mh^m)} \tag{5.17}$$

We can estimate the optimal step and convergence rate the same way we did in the previous chapter. Firstly, we note that $f(x)$ can have its own singularities in the $\mathbb{C}$ plane and that the $\tanh -m$ rule introduces more singularities that arise from the zeros of $\cosh^2 u^m$ which are $u = i\pi(k+1/2)$, $k \in \mathbb{Z}$. Let us focus on these last singularities. The closest zero of $\cosh^2 u^m$ to the real axis is at $w = i\pi/2$. If we assume that the singularities of $f(\tanh u^m)$ are not closer to the real axis than $w = i\pi/2$, then, theorem 4.3 predicts a discretization error of $\mathcal{O}(-\pi^2/h)$. The truncation error can be estimated by noting that $\frac{1}{\cosh^2((nh)^m)} = \frac{2}{\left(e^{(nh)^m}+e^{-(nh)^m}\right)^2} = \frac{2}{e^{2(nh)^m}\left(1+e^{-2(nh)^m}\right)^2} = \frac{2e^{-2(nh)^m}}{\left(1+e^{-2(nh)^m}\right)^2} = \mathcal{O}(e^{-2(nh)^m})$ as $nh \longrightarrow \infty$. Now, we must balance the discretization error and the truncation error to obtain the optimal step and the optimal convergence rate:

Optimal step for the $\tanh -m$ rule (Without considering the singularities of $f(z)$)

$$\mathcal{O}\left(e^{-\frac{\pi^2}{h}}\right) = \mathcal{O}\left(e^{-2(nh)^m}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \left(\frac{\pi^2}{2n^m}\right)^{\frac{1}{m+1}} \tag{5.18}$$

Optimal convergence rate for the tanh $-m$ rule (Without considering the singularities of $f(z)$)

$$h = \left(\frac{\pi^2}{2n^m}\right)^{\frac{1}{m+1}} \; ; \; \text{Discretization error: } \mathcal{O}\left(e^{-\frac{\pi^2}{h}}\right) \Longrightarrow \text{Optimal convergence rate: } \mathcal{O}\left(e^{-\pi^2\left(\frac{2n^m}{\pi^2}\right)^{\frac{1}{m+1}}}\right) \tag{5.19}$$

Now we can focus on the complicated case, that is, when $f(\tanh u^m)$ has singularities closer to the real axes than $w = {}^{i\pi}/_{2}$. Let z_s be a singularity of f. The corresponding point in the u-plane can be obtained as:

$$\begin{aligned} u_s = \operatorname{atanh}^{\frac{1}{m}}(z_s) \overset{\operatorname{atanh}(z)=\frac{1}{2}\ln\left(\frac{1+z}{1-z}\right)}{=} \left(\frac{1}{2}\ln\left(\frac{1+z_s}{1-z_s}\right)\right)^{\frac{1}{m}} = \\ = \left(\frac{1}{2}\ln\left(\left|\frac{1+z_s}{1-z_s}\right| e^{iArg\left(\frac{1+z_s}{1-z_s}\right)+2k\pi i}\right)\right)^{\frac{1}{m}} = \left(\frac{1}{2}\ln\left|\frac{1+z_s}{1-z_s}\right| + \frac{i}{2}Arg\left(\frac{1+z_s}{1-z_s}\right) + k\pi i\right)^{\frac{1}{m}} = \\ = (s+i(t+k\pi))^{\frac{1}{m}} \quad k \in \mathbb{Z} \end{aligned} \tag{5.20}$$

Where we have used the following notation:

$$s = \frac{1}{2}\ln\left|\frac{1+z_s}{1-z_s}\right| \quad t = \frac{1}{2}Arg\left(\frac{1+z_s}{1-z_s}\right) \tag{5.21}$$

With this new compact notation, it is trivial to obtain the modulus and the phase of u_s:

$$\begin{gathered} u_s = \operatorname{atanh}^{\frac{1}{m}}(z_s) = \left(s^2+(t+k\pi)^2\right)^{\frac{1}{2m}} \exp\left[\frac{i}{m}\left(\arctan\left(\frac{t+k\pi}{s}\right)+2\pi l\right)\right] \\ k \in \mathbb{Z} \quad l \in \mathbb{Z} \end{gathered} \tag{5.22}$$

The singularity closest to the real axis is that with $k = l = 0$. Let us call it u_0. The imaginary part of u_0 is:

$$Im(u_0) = \left(s^2+t^2\right)^{\frac{1}{2m}} \exp\left[\frac{i}{m}\arctan\left(\frac{t}{s}\right)\right] = \sigma^{\frac{1}{2m}}\sin\left(\frac{\tau_1}{m}\right) \tag{5.23}$$

Where we have used the following notation for consistency with [15]:

$$\sigma = s^2+t^2 \quad ; \quad \tau_1 = \arctan\left(\frac{t}{s}\right) \tag{5.24}$$

As a consequence, using theorem 4.3, we have that the discretization error is $\mathcal{O}\left(\exp\left(\frac{-2\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{h}\right)\right)$.

Now, we can balance the discretization error and the truncation error:

Optimal step for the tanh $-m$ rule (Considering the singularities of $f(z)$)

$$\mathcal{O}\left(\exp\left(\frac{-2\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{h}\right)\right) = \mathcal{O}\left(e^{-2(nh)^m}\right) \overset{\text{Solving for } h}{\Longrightarrow} h = \left(\frac{\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{n^m}\right)^{\frac{1}{m+1}} \tag{5.25}$$

Optimal convergence rate for the tanh $-m$ rule (Considering the singularities of $f(z)$)

$$h=\left(\frac{\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{n^m}\right)^{\frac{1}{m+1}};\ \text{Discretization error: } \mathcal{O}\left(\exp\left(\frac{-2\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{h}\right)\right)\Longrightarrow$$

$$\text{Optimal convergence rate: } \mathcal{O}\left(\exp\left(\frac{-2\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|}{\left(\pi\sigma^{\frac{1}{2m}}\left|\sin\left(\frac{\tau_1}{m}\right)\right|\right)^{\frac{1}{m+1}}}n^{\frac{m}{m+1}}\right)\right)\overset{m\gg1}{\approx}\mathcal{O}\left(e^{\left(\frac{-2\pi|\tau_1|}{m}\right)n^{\frac{m}{m+1}}}\right) \tag{5.26}$$

Remark 5.1. ***(Is the higher the m the better? [15][pages 726-729])*** *For the typical case $m=1$, the* tanh*-rule performs as $\mathcal{O}(e^{-cn^{1/2}})$, which is worse than the* erf*-rule. Note that the argument of the exponential in the optimal convergence rate in the high m regime behaves as:*

$$\frac{n^{\frac{m}{m+1}}}{m}\underset{m\longrightarrow\infty}{\longrightarrow}0 \tag{5.27}$$

And therefore:

$$e^{\left(\frac{-2\pi|\tau_1|}{m}\right)n^{\frac{m}{m+1}}}\underset{m\longrightarrow\infty}{\longrightarrow}1 \tag{5.28}$$

Which is not very convenient. Additionally, we can try to optimize $\frac{n^{\frac{m}{m+1}}}{m}$:

$$\begin{aligned}\left(\frac{n^{\frac{m}{m+1}}}{m}\right)'&=\left(-\frac{n^{\frac{m}{m+1}}}{m^2}+\frac{\log(n)n^{\frac{m}{m+1}}}{m(m+1)^2}\right)=0\Longrightarrow\frac{\log(n)}{m(m+1)^2}=\frac{1}{m^2}\\&\overset{m\gg1}{\Longrightarrow}\frac{\log(n)}{m^3}=\frac{1}{m^2}\Longrightarrow m=\log(n)\end{aligned} \tag{5.29}$$

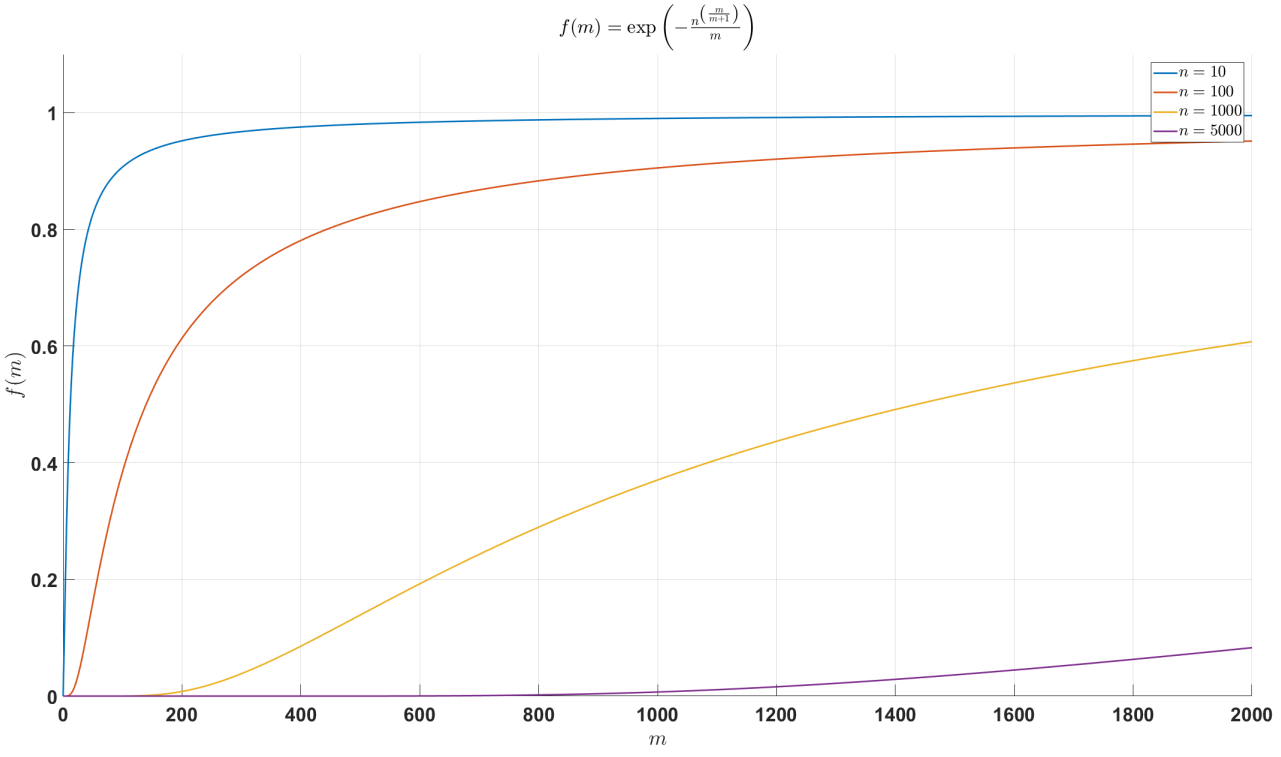

Figure 5.1: Evolution of the error for the tanh $-m$ rule in the high m regime for a fixed number of nodes n. The number of nodes is shown in the legend.

Which hints that in the high m regime, when the optimal convergence rate is $\mathcal{O}\left(e^{\left(\frac{-2\pi|\tau_1|}{m}\right)n^{\frac{m}{m+1}}}\right)$ there is in fact an optimal m minimizing the error, discarding the idea that the higher the m, the better. Now, recall that m in the definition 5.3 of the tanh $-m$ *rule, controls how fast the function decays after the change of variable is applied. Therefore, we can deduce that a faster decaying change of variable does not necessarily mean a better behaviour with the trapezoidal rule. These conclusions hint that there could be a sweet spot between fast decay and adequate truncation error that lead to an optimal convergence rate. We acknowledge that this reasoning is not rigorous at all. Nevertheless, a similar conclusion can be obtained by observing the behaviour of the error with m in the high m regime and for fixed number of nodes n. This is plotted in figure 5.1. The figure shows how the error remains low up to a certain value of m. From that m and above, the error increases a lot until it approaches its asymptotic limit 1. With all these arguments, we aim to conclude that for a fixed number of nodes n, the best value of m is not the greatest you can use.*

Remark 5.2. ***(About the distribution of the singularities u_s of the* tanh $-m$ *rule in the complex plane [15][pages 726-729])***
Recall that for a singularity z_s of $f(z)$, the tanh $-m$ *rule introduces the following singularities in the u-plane:*

$$u_s = atanh^{\frac{1}{m}}(z_s) = \left(s^2 + (t + k\pi)^2\right)^{\frac{1}{2m}} \exp\left[\frac{i}{m}\left(\arctan\left(\frac{t+k\pi}{s}\right) + 2\pi l\right)\right]$$
$$k \in \mathbb{Z} \quad l \in \mathbb{Z} \tag{5.30}$$
$$s = \frac{1}{2}\ln\left|\frac{1+z_s}{1-z_s}\right| \quad t = \frac{1}{2}Arg\left(\frac{1+z_s}{1-z_s}\right)$$

For $|k| \longrightarrow \infty$, these singularities tend to:

$$u_s \approx (|k|\pi)^{\frac{1}{m}} \exp\left[\left(\pm\frac{1}{2m} + \frac{2l}{m}\right)\pi i\right] \tag{5.31}$$

Indicating that they tend to group along $2m$ radial straight lines that go further and further from the real axis. We can observe how they distribute in figure 5.2.

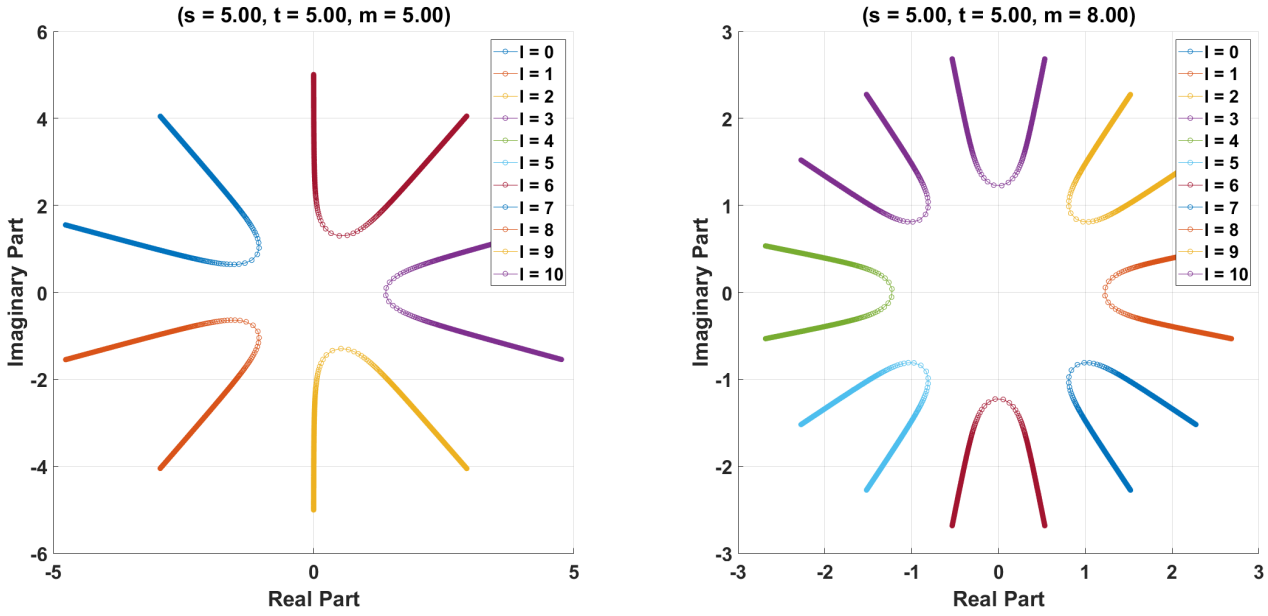

Figure 5.2: Plot of the singularities u_s following (5.30) for multiple values of l and k. The values of s, t and m are in the title of the plots.

We already discussed that the point that will likely control the convergence theorem 4.3 is that one for $k = l = 0$, which is the one closest to the real axis. The smart contribution of Takahase and Mori [15] is the

intuitive[1] *realization that this point should be the one that gives more error when computing numerically (5.17) since the others are getting further and further from the real axis. Consequently, the contribution of the singularities to the total error in (5.17) is not balanced in the sense that one point contributes a lot while the others that are far from the real axis give a negligible contribution. The next step of their rationale goes in the direction of optimization. In optimization, the error due to the contribution of different sources is usually minimized when all the sources contribute equally. With this idea in mind, the question is: is there any change of variable that makes all the singularities u_s contribute equally to the error of (5.17)? Such function would need to group all or most of the singularities at the same distance of the real axis. This question is the detonator of the birth of the doubly exponential formulas.*

5.2 The tanh − sinh rule and the doubly exponential formulas

Following the philosophy in remark 5.2, Takahase and Mori managed to find a change of variable that groups the singularities u_s into lines paralell to the real axis: the tanh − sinh formula.

Definition 5.4. ***(The* tanh − sinh *rule [15][page 730], [3][page 156], [7][page 430])*** *Following the notation in (5.3), we define the* tanh − sinh *rule as the change of variable*

$$x = \tanh\left(\frac{\pi}{2}\sinh(u)\right) \tag{5.32}$$

The tanh − sinh *rule converts (5.1) into*

$$I = \int_{-\infty}^{\infty} g(u)du = \int_{-\infty}^{\infty} f\left(\tanh\left(\frac{\pi}{2}\sinh(u)\right)\right)\frac{\cosh(u)}{\cosh^2\left(\frac{\pi}{2}\sinh(u)\right)}du \tag{5.33}$$

And we also have that

$$|g(u)| \approx e^{-\frac{\pi}{2}e^{|u|}} \quad (u \longrightarrow \pm\infty) \tag{5.34}$$

This change of variable results in the following trapezoidal approximation

$$I_h = \frac{\pi}{2}h\sum_{n=-\infty}^{\infty} f\left(\tanh\left(\frac{\pi}{2}\sinh(nh)\right)\right)\frac{\cosh(nh)}{\cosh^2\left(\frac{\pi}{2}\sinh(nh)\right)}. \tag{5.35}$$

Remark 5.3. ***(Why is there a factor* $1/2$*? [7][page 430])*** *One might wonder why the* tanh − sinh *rule has the $\frac{\pi}{2}$ factor in* $\tanh\left(\frac{\pi}{2}\sinh(u)\right)$*. To explain it, let us consider the rule as $x = \tanh\left(\mu\sinh(u)\right)$. Following (5.33), the rule would introduce poles where $\cosh^2\left(\mu\sinh(u)\right)$ vanishes i.e where $\mu\sinh(u) = \pm i\pi(k + 1/2)$, $k \in \mathbb{N}$. Let us focus on the poles above the real axis i.e those such that $\mu\sinh(u) = i\pi(k+1/2)$, $k \in \mathbb{N}$. The pole closest to the real axis of* cosh *is the one with $k = 0$. Solving for u leads to:*

$$u = \operatorname{arcsinh}\left(i\frac{\pi}{2\mu}\right) = i\arcsin\left(\frac{\pi}{2\mu}\right) \quad , k \in \mathbb{N} \tag{5.36}$$

We have now two possibilities. Firstly, if $0 < \frac{\pi}{2\mu} \leq 1$, equivalently $\frac{\pi}{2} \leq \mu$, then the pole can be at most a distance $\frac{\pi}{2}$ above the real axis. If we keep increasing μ over $\frac{\pi}{2}$, then the pole starts getting closer to the real axis, and this is something we would like to avoid. Secondly, if $\frac{\pi}{2\mu} \geq 1$, equivalently $0 < \mu \leq \frac{\pi}{2}$, we can rewrite the previous result using theorem 8.11:

$$\begin{aligned} u = \operatorname{arcsinh}\left(i\frac{\pi}{2\mu}\right) = i\arcsin\left(\frac{\pi}{2\mu}\right) \quad , k \in \mathbb{N} &\overset{\text{Theorem 8.11}}{=} i\left(\frac{\pi}{2} \pm i\ln\left[\frac{\pi}{2\mu} + \sqrt{\left(\frac{\pi}{2\mu}\right)^2 - 1}\right]\right) \quad , k \in \mathbb{N} = \\ &= \frac{i\pi}{2} \mp \ln\left[\frac{\pi}{2\mu} + \sqrt{\left(\frac{\pi}{2\mu}\right)^2 - 1}\right] \quad , k \in \mathbb{N} \end{aligned} \tag{5.37}$$

[1] We would like to emphasize that we are using the word **intuitive**.

And then, we get that if for $0 < \mu \leq \frac{\pi}{2}$, the pole can still be at most a distance $\frac{\pi}{2}$ over the real axis. We can conclude therefore that it is more convenient to choose $0 < \mu \leq \frac{\pi}{2}$, being $\mu = \frac{\pi}{2}$ the value that gives the fastest decay while keeping the pole at its maximum distance possible from the real axis.

We will obtain the error/optimal convergence rate for this rule as well. Firstly, the truncation error from (4.101) and (5.34) is $\mathcal{O}(\exp(-\frac{\pi}{2}e^{nh}))$. If $f(z)$ has no singularities, the only ones to pay attention to are those arising from $\cosh^2\left(\frac{\pi}{2}\sinh(u)\right)$. These zeros appear when $\frac{\pi}{2}\sinh(u) = \pm i\pi(k + 1/2)$, $k \in \mathbb{N}$. The pole closest to the real axis of cosh is the one with $k = 0$. Solving for u leads to:

$$u = \operatorname{arcsinh}(i) = i\arcsin(1) = \frac{i\pi}{2} \tag{5.38}$$

As a result, we can use theorem 4.3 to predict a discretization error $\mathcal{O}(e^{-\pi^2/h})$. In this case, in order to match the discretization and the truncation error, we need to solve for h:

$$2\pi n = nhe^{nh} \tag{5.39}$$

Its solution is given by:

$$h = \frac{W(2\pi n)}{n} \tag{5.40}$$

Where $W(x)$ is the Lambert W-function[2] [7][page 430]. This function behaves as $W(x) \approx \ln(x)$ as $x \longrightarrow \infty$ and consequently, for big n we can approximate $h = \frac{\ln(2\pi n)}{n}$ and we can obtain now the optimal convergence rate.

Optimal convergence rate for the tanh − sinh-rule (For $f(z)$ without singularities)

$$h = \frac{\ln(2\pi n)}{n};\ \text{Discretization error: } \mathcal{O}(e^{-\pi^2/h}) \Longrightarrow \text{Optimal convergence rate: } \mathcal{O}\left(e^{-\frac{\pi^2 n}{\ln(2\pi n)}}\right) \tag{5.41}$$

If $f(z)$ has singularities z_s, then we must analyze the map $u_s = \operatorname{arcsinh}\left(\frac{2}{\pi}\operatorname{arctanh}(z_s)\right)$. Due to the complexity and length of the full calculation, we include here only the result [15][page 730]:

$$\begin{aligned} u_s = l\pi i + (-1)^l \Bigg[& \operatorname{arcosh}\frac{1}{\pi}\left\{\sqrt{s^2 + \left(t + \frac{\pi}{2} + k\pi\right)^2}\right. \\ & + \sqrt{s^2 + \left(t - \frac{\pi}{2} + k\pi\right)^2} \pm i\arcsin\frac{1}{\pi}\left\{\sqrt{s^2 + \left(t + \frac{\pi}{2} + k\pi\right)^2}\right\} \\ & \left. - \sqrt{s^2 + \left(t - \frac{\pi}{2} + k\pi\right)^2}\right] \quad k, l \in \mathbb{Z} \\ s = \frac{1}{2}\ln\left|\frac{1+z_s}{1-z_s}\right| & \quad t = \frac{1}{2}Arg\left(\frac{1+z_s}{1-z_s}\right) \end{aligned} \tag{5.42}$$

With a similar analysis to the one in remark 5.3 about the behaviour of the arcsin, for $l = 0$, the singularities can be a distance $\frac{\pi}{2}$ at most from the real axis, independently of the value of k. Besides, note that

$$u_s \approx l\pi i + (-1)^l\left(\operatorname{arcosh} 2k \pm \frac{\pi}{2}i\right), \quad k \to \infty, \tag{5.43}$$

[2] Defined as the function following $W(x)e^{W(x)} = x$

And then, we extract that the point closest to the real axis is that with $k = l = 0$, namely u_0. We have that

$$\tau_2 := Im(u_0) = \operatorname{arcsin} \frac{1}{\pi} \left\{ \sqrt{s^2 + \left(t + \frac{\pi}{2}\right)^2} - \sqrt{s^2 + \left(t - \frac{\pi}{2}\right)^2} \right\} = \\ = \operatorname{arcsin} \left\{ \frac{2t}{\sqrt{s^2 + \left(t + \frac{\pi}{2}\right)^2} + \sqrt{s^2 + \left(t - \frac{\pi}{2}\right)^2}} \right\} \tag{5.44}$$

Using theorem 4.3, the discretization error is then $\mathcal{O}(e^{-\frac{2\pi|\tau_2|}{h}})$ and we can finally obtain the optimal step and the optimal convergence rate. Balancing the discretization error and the truncation error involves solving

$$4|\tau_2|n = nhe^{nh} \tag{5.45}$$

And the solution would be:

$$h = \frac{W(4|\tau_2|n)}{n} \tag{5.46}$$

In the same way we did before, for high n, we can make the approximation $W(x) \approx \ln(x)$ and then, the optimal step would be:

$$h = \frac{\ln(4|\tau_2|n)}{n} \tag{5.47}$$

Optimal convergence rate for the tanh − sinh-rule (Considering the singularities of $f(z)$)

$$h = \frac{\ln(4|\tau_2|n)}{n}; \text{ Discretization error: } \mathcal{O}(e^{-\pi^2/h}) \Longrightarrow \text{ Optimal convergence rate: } \mathcal{O}\left(e^{-\frac{\pi^2 n}{\ln(4|\tau_2|n)}}\right) \tag{5.48}$$

And we have obtained the remarkable result that the tanh − sinh rules behaves similarly independently of whether or not $f(z)$ has singularities. Also note that the tanh − sinh behaves better than the tanh $-m$ rule and the erf-rule when singularities are present.

Remark 5.4. ***(About the distribution of the singularities u_s of the*** tanh − sinh ***rule in the complex plane [15][pages 730-732])*** *In remark 5.2, we discussed how intuitively, the main contribution to the error of (5.17) would probably be due to the point closest to the real axis since the rest are getting further and further apart from the real line. Then, Takahase and Mori proposed that a better mapping function should map all or most singularities at the same distance of the real axis so that all of them contribute in equal manner to the error. With the* tanh − sinh *rule, we have obtained a better error estimate than with the* tanh $-m$ *rule. We can check now how the singularities are distributed in the u-plane. Firstly, recall that*

$$u_s \approx l\pi i + (-1)^l \left(\operatorname{arcosh} 2k \pm \frac{\pi}{2} i\right), \quad k \to \infty, \tag{5.49}$$

This means that the singularities are asymptotically lying on lines parallel to the real axis. We can see this visually by plotting (5.42) as shown in figure 5.3. From the figure, we can see how the most of the points are in fact at the same distance from the real axis and therefore, we can expect them to contribute a similar amount to the error of (5.35). Takahase and Mori attribute the better error estimate of this rule to this convenient distribution of the singularities and conclude the core of their paper by saying an optimal change of variable with respect to efficiency should follow one of these two conditions:

- *That* $|g(u)| \approx e^{-\frac{\pi}{2}e^{|u|}} \quad (u \longrightarrow \pm\infty)$
- *That the change of variable maps the singularities of $f(z)$ into an array of points that lie asymptotically parallel to the real axis.*

Takahase and Mori gave birth to the doubly exponential formulas: changes of variables satisfying the first condition. There are more doubly exponential formulas apart from the tanh − sinh *that apply for different intervals of integration [15][pages 733-735] [3][page 156].*

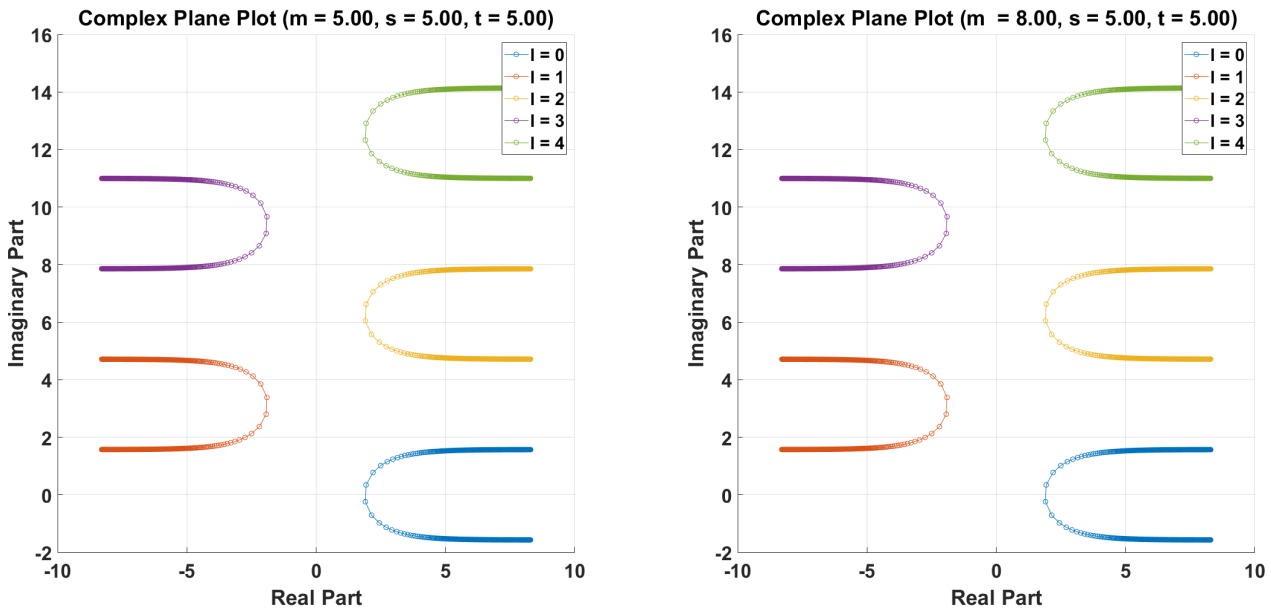

Figure 5.3: Plot of the singularities u_s following (5.42) for multiple values of l and k. The values of s, t and m are in the title of the plots.

Remark 5.5. ***(What happens if we increase the decay rate of the*** tanh − sinh ***rule? [15][page 731])*** *We still have to analyze if increasing the decay rate arbitrarily would improve the behaviour of the* tanh − sinh. *Consider:*

$$z = \tanh\left(\frac{\pi}{2}\sinh(u^m)\right) \tag{5.50}$$

With this rule, the asymptotical distribution of the singularities goes as:

$$u_s = (\operatorname{arcosh}(2k) \pm \frac{i\pi}{2})^{\frac{1}{m}} \quad k \longrightarrow \infty \tag{5.51}$$

And taking the limit as $m \longrightarrow \infty$ shows that all the points converge to the real axis (the value 1). This is not convenient at all since the convergence theorems 4.3, 4.4 and 4.5 work better when the singularities are far from the real axis and we have verified empirically that changes of variables have a better error estimate when the singularities are grouped in lines parallel to the real axis.

The good thing about computers is that they do what you tell them to do. The bad news is that they do what you tell them to do.

Ted Nelson

Chapter 6

Computational performance of Romberg integration VS trapezoidal rule

Up to this point, this text has been fully theoretical despite considering a tool such as the trapezoidal rule. To give a flavour of how the numerics work in practice, we briefly compare in this chapter the composite trapezoidal rule, Romberg integration and the usage of changes of variables for some simple cases.

6.1 Integrals for 2π-periodic functions

One of the most relevant use case scenarios for the trapezoidal rule is the integration of 2π-periodic functions. We have seen already with theorems 4.1 and 4.2 that the convergence should be exponential with the number of nodes. In figure 6.3 left, we compare the error of the trapezoidal rule computed with (4.2) and that of Romberg integration for the function $f = \frac{\sin(x)}{2+\cos(x)}$. For the Romberg integration, we show the values $R(n,n)$ where n is the number of nodes and we are following the notation of the Romberg Integration section. The theoretical value displayed is the one given by theorem 4.1.

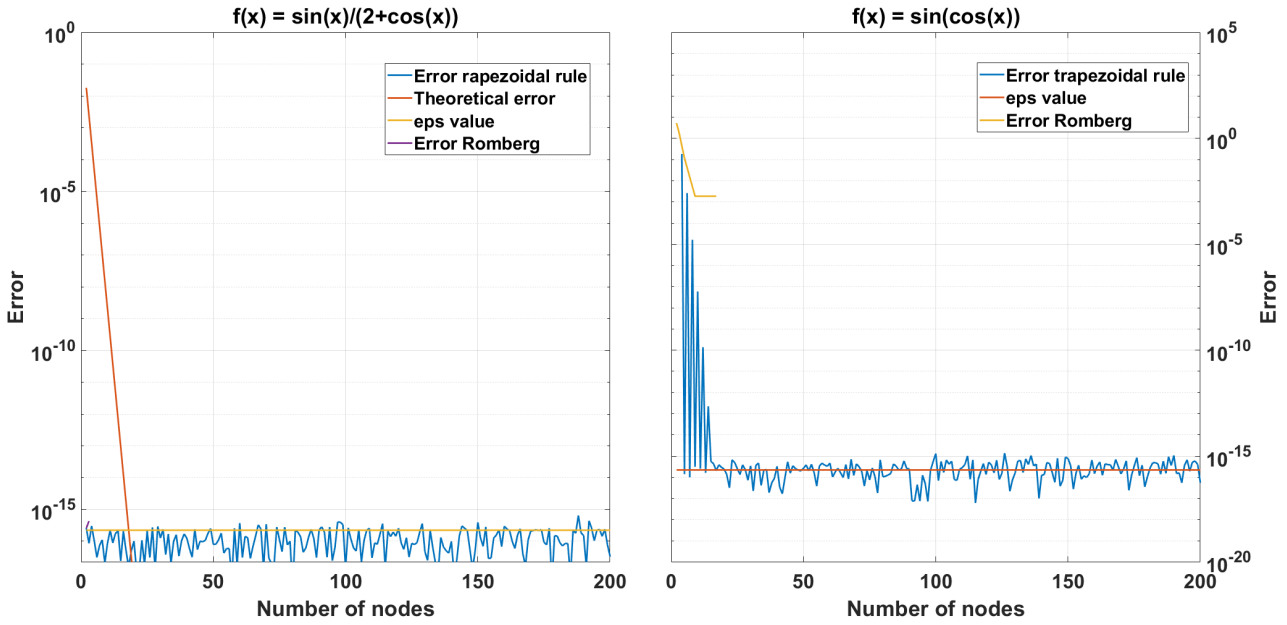

Figure 6.1: Evolution of the error of the trapezoidal rule and Romberg integration over $[0, 2\pi]$ for the functions shown on top of each plot as a function of the number of nodes. The value of the *eps* constant of `MATLAB` is shown for comparison. Both plots show an excellent convergence of the trapezoidal rule and Romberg integration, saturating to the *eps* value with just a bunch of nodes.

Note that the trapezoidal rule is giving a value of the order of *eps* of MATLAB with just a few nodes. The error of the Romberg integration is not even seen properly on the plot because algorithm 3.1 stops with just 3 nodes. The right side of the figure shows the same two errors but for the function $f = \sin(\cos(x))$. In this case, both rules perform worse than in the previous case, but the error still saturates around the *eps* value with a relative low number of nodes.

6.2 Integrals over $\mathbb{R}$

Another relevant case for benchmarking is the integration over the real line. In this section we will compare the performance of the trapezoidal rule against the sinh − sinh rule doubly exponential formula [15][page 735]:

$$\begin{aligned} x &= \sinh\left(\frac{\pi}{2}\sinh u\right) \\ I_h &= \frac{\pi}{2}h\sum_{n=-\infty}^{\infty} f\left(\sinh\left(\frac{\pi}{2}\sinh nh\right)\right)(\cosh nh)\cosh\left(\frac{\pi}{2}\sinh nh\right) \end{aligned} \tag{6.1}$$

We will compare the integration of functions $f(x) = \frac{e^{-x^2}}{1+x^2}$ and $f(x) = \frac{1}{1+x^2}$. The numerical integration is computed following (4.56) with nodes given in MATLAB notation as $[-nh : h : nh]$ where n is the number of nodes and the step $h = \left(\frac{2\pi}{n^2}\right)^{1/3}$ in consistency with the optimal step and convergence rate in (4.3). The results in figure 6.3 left for $f = \frac{e^{-x^2}}{1+x^2}$ show how the composite trapezoidal rule behaves following the estimated optimal error and surprisingly, it behaves better than the sinh − sinh rule. This observation follows the observation made by Takahase and Mori that we described in remark 5.4 that the optimal behaviour is when the function decays in a doubly exponential manner. However, when using the sinh − sinh, the decay would be triple exponential, which is not the optimal case according to this remark. On the right side of figure 6.3 for $f(x) = \frac{1}{1+x^2}$, we see the opposite trend. In this case, the sinh − sinh does improve the convergence since this time the transformed function decays in a doubly exponential manner. The composite trapezoidal rule has a very bad convergence as seen from the figure. Further inspection of its behaviour shows how the composite trapezoidal requires around 60000 nodes to achieve an error of the order of *eps*. In such cases, the usage of a doubly exponential formula is then essential.

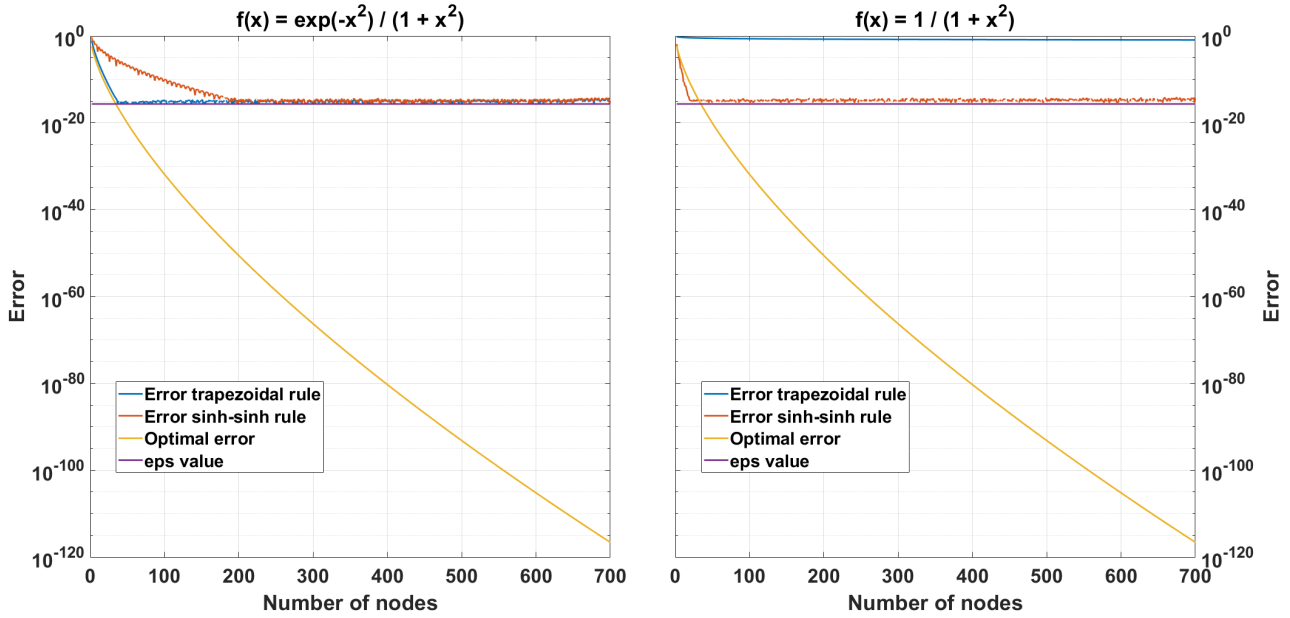

Figure 6.2: Evolution of the error of the trapezoidal rule and the sinh − sinh rule over $\mathbb{R}$ for the functions shown on top of each plot as a function of the number of nodes. The value of the *eps* constant of MATLAB is shown for comparison.

6.3 Integrals over $[-1,1]$

Last but not least, we compare the performance of Romberg integration, the composite trapezoidal rule and the tanh $-$ sinh rule when integrating over $[-1,1]$. We will do it for the functions $f(x) = \frac{e^{-x^2}}{1+x^2}$ and $f(x) = x^2$. The tanh $-$ sinh rule is approximation is calculated using (4.56) with nodes given by $[-3 : h : 3]$. The first case is shown in figure 6.3 left. In this case, the composite trapezoidal rule performs poorly. The Romberg approach improves the convergence, but the winner is the change of variable, which manages to obtain an error of the order of *eps* with less than a hundred nodes, outperforming the composite trapezoidal rule as expected. The situation is partially different for $f(x) = x^2$ in figure 6.3 right. In this case, the winner is Romberg integration because Simpson's rule has 3 as degree of exactness, giving therefore the exact result. As it happened before, the tanh $-$ sinh outperforms the composite trapezoidal rule.

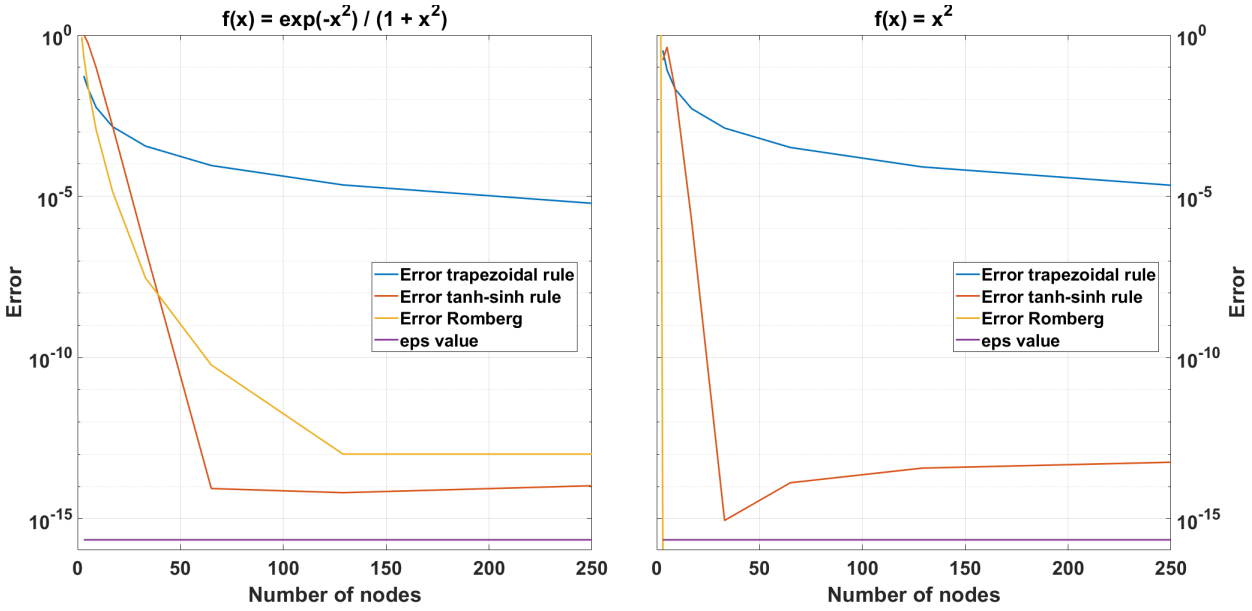

Figure 6.3: Evolution of the error of the trapezoidal rule and Romberg integration over $[-1,1]$ and the tanh $-$ tanh rule over $\mathbb{R}$ for the functions shown on top of each plot as a function of the number of nodes. The value of the *eps* constant of `MATLAB` is shown for comparison.

Nothing is particularly hard if you divide it into small jobs.

Henry Ford

Chapter 7

Appendix A: optional content

In this chapter, we will introduce (without proofs) some previous useful results. The interested reader can find proofs and information on the corresponding references.

7.1 Runge phenomenon

In this section we exemplify the *Runge phenomenon* with the analysis of the Lagrange interpolation of the function $f(x) = 1/(1+x^2)$ on the interval $[-5,5]$ [3, pages 54-55]. We denote by $P_n(x)$ the $n-$th degree Lagrange interpolating polynomial of the function $f(x)$. What Runge prooved is that:

$$\lim_{n\to\infty} ||f - P_n|| = \infty \tag{7.1}$$

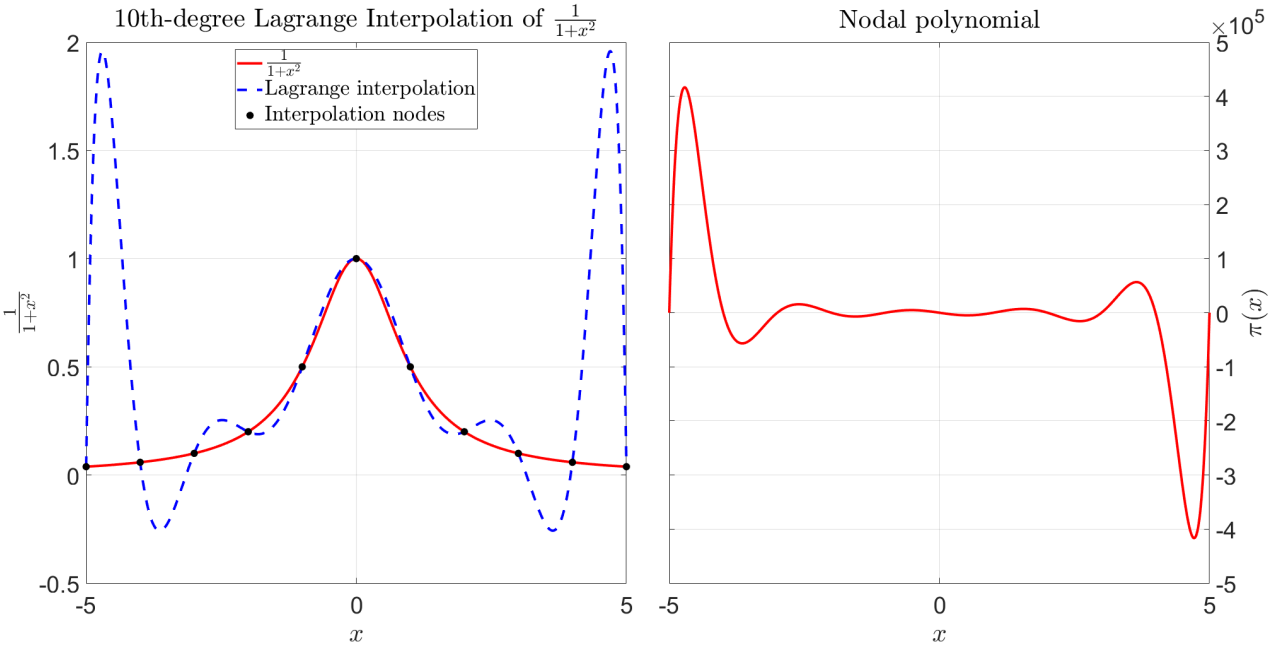

Figure 7.1: Left: The function $f(x) = 1/(1+x^2)$ and its 10-th degree Lagrange interpolation. The interpolation nodes are represented by black dots. Right: Nodal polynomial for the 10-th degree Lagrange interpolation of $f(x)$. The nodal polynomial is: $\pi(x) = x(x^2-1)(x^2-4)(x^2-3)(x^4-16)(x^2-25)$.

Which means that arbitrarily increasing the number of interpolation nodes in fact does not improve the interpolation for this function. The explanation for this phenomenon is obtained after inspecting the behaviour of the nodal polynomial close to the endpoints of the interval. For example, we

shall consider $P_{10}(x)$ with nodes in $\pm1, \pm2, \pm3, \pm4, \pm5$ and 0. The Lagrange interpolating polynomial is shown in figure 7.1 left. It observed that close to the edges of the interval, the interpolating polynomial presents strong oscillations and the approximation is poor on that area.
This behaviour is what forces $\lim_{n\to\infty} ||f - P_n|| = \infty$. This counterexample illustrates how increasing the number of equally spaced nodes does not always improve the interpolation.

7.2 More results for Bernouilli polynomials

Proposition 7.1. ***(Properties of the Bernouilli polynomials)*** *Let k be a natural number. The Bernouilli polynomials satisfy the following properties:*

1. $B_k(1-x) + B_k = (-1)^k [B_k(x) + B_k]$ *for* $k \geq 0$.
2. $B_{2k+1}\left(\frac{1}{2}\right) = 0$ *for* $k > 0$.
3. $(-1)^k(B_{2k-1}(x) + B_{2k-1}) > 0, 0 < x < 1/2$ *for* $k \geq 1$.
4. $(-1)^k B_{2k}(x) > 0, 0 < x < 1$ *for* $k \geq 1$.
5. $(-1)^{k+1} B_{2k} > 0$ *for* $k \geq 1$.
6. $B_{2k+1}(x)$ *can only vanish in* $[0,1]$ *at* $x = 0, \frac{1}{2}, 1$ *for* $k \geq 1$.
7. $B_{2k}(x)$ *can only vanish in* $[0,1]$ *at* $x = 0, 1$ *for* $k \geq 1$.

Proof. We will start step by step:

- 1. By using the definitions of the Bernouilli numbers and polynomials, 3.1 and 3.2, we have:

$$\frac{t(e^{xt}-1)}{e^t-1} + \frac{t}{e^t-1} = \frac{te^{xt}}{e^t-1} \overset{\text{Defintions 3.1 and 3.2}}{=} \sum_{k=0}^{\infty} (B_k(x) + B_k)\frac{t^k}{k!} \tag{7.2}$$

Using the definition of the Bernouilli polynomials 3.1 but evaluating at $1-x$ and for $-t$ instead of t:

$$\frac{-t(e^{-(1-x)t}-1)}{e^{-t}-1} = \frac{-t(e^{-t}e^{xt}-1)}{e^{-t}-1} \overset{\text{Definition 3.1}}{=} \sum_{k=0}^{\infty} B_k(1-x)(-1)^k\frac{t^k}{k!} \tag{7.3}$$

And then, using the definition of the Bernouilli numbers 3.7 with the previous result:

$$\frac{-t(e^{-t}e^{xt}-1)}{e^{-t}-1} - \frac{t}{e^{-t}-1} = \frac{-te^{-t}e^{xt}}{e^{-t}-1} \overset{\text{Definition 3.1}}{=} \sum_{k=0}^{\infty} (B_k(1-x) + B_k)(-1)^k\frac{t^k}{k!} \tag{7.4}$$

The first and last result in this proof imply:

$$\begin{aligned} \frac{te^{xt}}{e^t-1} + \frac{te^{-t}e^{xt}}{e^{-t}-1} = te^{xt}\left(\frac{1}{e^t-1} + \frac{e^{-t}}{e^{-t}-1}\right) = te^{xt}\left(\frac{e^{-t}-1+e^{-t}(e^t-1)}{(e^t-1)(e^{-t}-1)}\right) = \\ = e^te^{xt}\left(\frac{e^{-t}-1+1-e^{-t}}{(e^t-1)(e^{-t}-1)}\right) = 0 \end{aligned} \tag{7.5}$$

And this last equality implies:

$$\sum_{k=0}^{\infty} (B_k(1-x) + B_k)(-1)^k\frac{t^k}{k!} = \sum_{k=0}^{\infty} (B_k(x) + B_k)\frac{t^k}{k!} = 0 \tag{7.6}$$

And since the equality holds for every x and t, we obtain the final result:

$$(B_k(1-x) + B_k)(-1)^k = (B_k(x) + B_k) \tag{7.7}$$

- 2. It follows from 1. evaluated at $x = 1/2$ and the fact that $B_{2k+1} = 0$ for all $k > 0$.
- We are going to prove 3. 4. 5. following this process:

  A 3. is true for $k = 1$.

  B 3. $\Longrightarrow$ 4. $\Longrightarrow$ 5 for each k.

  C If 3. is true for $k \in \mathbb{N}$ (so are 4. and 5.), then it is true for $k + 1$.

  We can start now:

  A Using property 6. in proposition 3.1 and (3.6):

$$B_1 = -\int_0^1 B_1(x)dx = -\int_0^1 xdx = -\frac{1}{2} \tag{7.8}$$

  So now we have: $(-1)(B_1(x) + B_1) = -(x - 1/2) = -x + 1/2 > 0$ if $0 < x < 1/2$.

  B

3. $\Rightarrow$ 4. For $k = 1$, we have using (3.6) that $B_2(x) = x^2 - x$. Then, $-(x^2 - x) < 0$ if $0 < x < 1$ and the property is true. For $k > 1$, using property 4 in proposition (3.1):

$$B_{2k-1}(x) = \frac{1}{2k}B'_{2k}(x) \tag{7.9}$$

   And multiplying both sides by $(-1)^k$ and integrating from 0 to $0 < x \leq 1/2$:

$$\frac{(-1)^k}{2k}B_{2k}(x) = (-1)^k \int_0^x B_{2k-1}(t)dt > 0 \tag{7.10}$$

   Due to property 3. and the fact that $B_{2k+1} = 0$ for $k > 1$ (Property 3. in proposition 3.1). Therefore, $(-1)^k B_{2k}(x) > 0$ for $k > 0$ if $0 < x \leq 1/2$. But using property 1., we have that $B_{2k}(1 - x) = B_{2k}(x)$ and therefore, $(-1)^k B_{2k}(x) > 0$ if $1/2 \leq x < 1$. So finally, we have that $(-1)^k B_{2k}(x) > 0$ if $0 < x < 1$.

4. $\Rightarrow$ 5. By using property 6. in proposition 3.1, we have for $k \geq 1$:

$$-B_{2k} = \int_0^1 B_{2k}(x)dx \tag{7.11}$$

   And multiplying both sides by $(-1)^k$:

$$(-1)^{k+1}B_{2k} = (-1)^k \int_0^1 B_{2k}(x)dx > 0 \tag{7.12}$$

   Where the last inequality follows from property 4. Therefore, we have that $(-1)^{k+1}B_{2k} > 0$.

   C Let us assume that $(-1)^k(B_{2k-1}(x) + B_{2k-1}) > 0, 0 < x < 1/2$ for an integer $k > 1$. Using property 3. in proposition 3.1 that $B_{2k+1} = 0$ the assumption is equivalent to $(-1)^k B_{2k-1}(x) > 0, 0 < x < 1/2$ for an integer $k > 1$. We must check what happens with $k + 1$ i.e with $B_{2k+1}(x)$. Using properties 1 and 2 of preposition 3.1, we know that $B_{2k+1}(0) = B_{2k+1}(1) = 0$. Using property 2., we have that $B_{2k+1}\left(\frac{1}{2}\right) = 0$. As a result, if $B_{2k+1}(x)$ vanishes at some point in $(0, 1/2)$, then there is an inflection point y in $(0, 1/2$ such that $B''_{2k+1}(y) = 0$. But using properties 4. and 5. in proposition 3.1, this would mean that:

$$0 = B''_{2k+1}(y) = (2k + 1)B'_{2k}(y) \overset{\text{Property 4. in 3.1}}{=} (2k + 1)2kB_{2k-1}(y) \tag{7.13}$$

   Giving that $B_{2k-1}(y) = 0$ but this contradicts the hypothesis that $(-1)^k B_{2k-1}(x) > 0, 0 < x < 1/2$. Consequently, we have that $B_{2k+1} \neq 0$ in $(0, 1/2)$ and its derivative does not

change in this interval. Therefore, the sign of $B_{2k+1}(x)$ in this interval is the same as the sign of $B'_{2k+1}(0) =$ and we have that:

$$\begin{aligned} sign(B'_{2k+1}(0)) &\overset{\text{Property 5. in 3.1}}{=} sign((2k+1)(B_{2k}(0)+B_{2k})) \overset{\text{Property 1. in 3.1}}{=} \\ &= sign(B_{2k}) \overset{\text{Property 5.}}{=} (-1)^{k+1} \end{aligned} \tag{7.14}$$

And consequently, we have that $(-1)^{k+1}B_{2k+1}(x) > 0$ if $0 < x < 1/2$ and the property is proved for $k+1$.

And with all these steps, properties 3., 4. and 5. are proved.

- 6. It is a trivial consequence of 1. and 3.
- 7. It is trivial consequence of 4.

□

7.3 Rigorous development of the doubly exponential formulas

In the chapter about changes of variable, we arrived to the doubly exponential formulas following the original rationale by Takahasi and Mori [15]. We acknowledge that the logical process followed is not very rigorous: it relies on approximations, qualitative reasoning and observation. Nevertheless, the optimality of the doubly exponential formulas has been treated in a rigorous theoretical approach as narrated by Mori [16][pages 919-921]:

> In order for mathematicians to be convinced of the optimality of the DE formula Mori persuaded Sugihara to establish a theorem of optimality in a more mathematically rigorous manner because Mori knew that functional analysis is Sugihara's favorite subject and thought that it would be possible for Sugihara to bring it successful.
>
> Answering Mori's expectation Sugihara soon challenged this problem from his own stand point. Sugihara's idea can be summarized as follows. He first investigated the optimality in the function space whose elements decay single exponentially. Next he investigated it in the function space whose elements decay double exponentially. Then he found that the trapezoidal formula is optimal in each of both function spaces and that the error of the trapezoidal formula approaches zero faster in the function space with double exponential decay than in the space with single one as the number of function evaluations increases. Then, in order to find a more efficient transformation, it is natural to consider next a function space whose elements decay faster than double exponentially. To this problem he gave a negative answer by proving a theorem that there exists no function that decays faster than double exponentially with an exponent $\frac{\pi}{2d}$ as $Re(w) \to \pm\infty$ as long as we think of functions regular in the strip region $|Im(w)| < d$. Thus Sugihara succeeded in establishing a theorem for the optimality of the trapezoidal formula combined with the double exponential decay of integrands, although the conclusion was on the whole the same as that by Takahasi and Mori. Incidentally it should be noted that Sugihara's nonexistence theorem stated above seems to contradict the analysis by Takahasi and Mori in that they considered functions whose decay is faster than double exponential. But it is not the case because they could do it since they allowed the singularities of the weight function $\phi'(w)$ to approach to the real axis as $Re(w) \to \pm\infty$.
>
> Sugihara first presented the result orally at a RIMS symposium in 1986 [64] and it took more than ten years for him until he completed and published his theorem in 1997 in Numerische Mathematik [67]. Mori thinks that the reason why it took more than ten years for Sugihara to complete the work would be that he is very careful and prudent for everything.

Those who cannot remember the past are condemned to repeat it.

George Santayana

Chapter 8

Appendix B: previous results

8.1 Interpolation theory

Theorem 8.1. ***(Error of linear interpolation [1][pages 9,134-136])*** *Let a, b be two different real numbers, $a < b$, and let f be a real function with 2 continuous derivatives in $[a, b]$. Then:*

$$f(x) - \frac{(b-x)\,f(a) + (x-a)\,f(b)}{b-a} = (x-a)\,(x-b)\,\frac{f''(\eta_x)}{2} = (x-a)\,(x-b)\,f[a,b,x] \tag{8.1}$$

For some η_x in $[a, b]$.

8.2 Fourier analysis

Theorem 8.2. ***(Fourier series of a periodic function [13][pages 263-264])*** *Let f be an analytic and ω-periodic function in a region Ω that satisfies the following property:*

$$z \in \Omega \Longrightarrow z - \omega \in \Omega \text{ and } z + \omega \in \Omega \tag{8.2}$$

Then, we can write:

$$f(z) = \sum_{j=-\infty}^{\infty} c_j e^{\frac{2\pi i j z}{\omega}} \tag{8.3}$$

With the c_j terms given by:

$$c_j = \frac{1}{\omega}\int_0^{\omega} f(z) e^{-\frac{2\pi i j z}{\omega}} \tag{8.4}$$

Corollary 8.1. ***(Fourier series of a 2π-periodic function)*** *Let f be an analytic and 2π-periodic function in a region Ω that satisfies the following property:*

$$z \in \Omega \Longrightarrow z - 2\pi \in \Omega \text{ and } z + 2\pi \in \Omega \tag{8.5}$$

Then, we can write:

$$f(z) = \sum_{j=-\infty}^{\infty} c_j e^{ijz} \tag{8.6}$$

With the c_j terms given by:

$$c_j = \frac{1}{2\pi}\int_0^{2\pi} f(z) e^{-ijz} \tag{8.7}$$

Theorem 8.3. ***(Absolute and uniform convergence of the Fourier series [14][page 22])*** *Let c_j be the Fourier series coefficients defined in (8.4). If*

$$\sum_{j=-\infty}^{\infty} |c_j| < \infty \tag{8.8}$$

then, the Fourier series defined in (8.3) converges absolutely and uniformly to f.

8.3 Complex analysis results

Theorem 8.4. ***(Laurent Series [10, page 123])*** *If f is an analytic function in the annulus $R_1 < |z - z_0| < R_2$, then f has a unique representation:*

$$\begin{aligned} f(z) &= \sum_{k=-\infty}^{\infty} a_k (z - z_0)^k \\ a_k &= \frac{1}{2\pi i} \oint \frac{f(z)}{(z - z_0)^{k+1}} \end{aligned} \tag{8.9}$$

With $\mathcal{C} = \mathcal{C}(z_0; R)$ and $R_1 < R < R_2$

Definition 8.1. ***(Residue of a function [10, page 129]))*** *Let f be an analytic function in the annulus $R_1 < |z - z_0| < R_2$. With the notation of theorem 8.4, we call the residue of f at z_0 the coefficient $a_{-1} = Res(f; z_0)$.*

Definition 8.2. ***(Winding number [10, page 131]))*** *Let γ be a closed curve in the complex plane and z_0 a point of the complex plane. Then:*

$$n(\gamma, z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{dz}{z - z_0} \tag{8.10}$$

Is the winding number of γ around z_0. One can proof that the winding number is always an integer.

Theorem 8.5. ***(Residue theorem [10, page 134])*** *Let f be an analytic function in a simply connected domain D except for isolated singularities at $z_1, z_2, \ldots, z_m$. Let γ be a closed curve not intersecting any of the singularities. Then:*

$$\frac{1}{2\pi i} \oint_\gamma f(z) dz = \sum_{k=1}^{m} Res(f; z_k) n(\gamma, z_k) \tag{8.11}$$

Corollary 8.2. *If f has an isolated singularity in z_0, then it has a Laurent Series in a neighbourhood of z_0. Let us consider the terms in the Laurent series a_{-k} with $k > 0$.*

1. *$a_{-k} = 0$ for all $k > 0$ if and only if the singularity is removable.*
2. *There is a finite number of nonzero a_{-k} if and only if the singularity is a pole. If m is the order of the pole, the last nonzero term is a_{-m}.*

Theorem 8.6. ***(Cauchy integral theorem [10, page 56])*** *Let f be an analytic function in a simple connected set D of the complex plane. Let γ be a closed path in D. Then:*

$$\oint_\gamma f(z) dz = 0 \tag{8.12}$$

Theorem 8.7. ***(Cauchy integral formula [10, page 61])*** *Let f be an analytic function in a simple connected domain D. Let $\mathcal{C}(z_0, r)$ be a circumference in D with positive orientation. Then, for every α inside the circumference defined by $\mathcal{C}(z_0, r)$:*

$$\frac{1}{2\pi i} \oint_{\mathcal{C}(z_0, r)} \frac{f(z)}{z - \alpha} dz = f(\alpha) \tag{8.13}$$

If α is not inside the circumference defined by $\mathcal{C}(z_0, r)$, the integral vanishes as a consequence of the Cauchy integral theorem 8.6.

Remark 8.1. *The Cauchy integral formula is usually stated for positively oriented circumferences as done in the previous theorem, but the theorem still holds for any positively oriented simple closed curve. In fact, in [12, page 144], the theorem is introduced as done in the following theorem.*

Theorem 8.8. ***(Alternative version of the Cauchy integral formula [12, page 144])*** *Let f be an analytic function in a simple connected domain D. Let $\mathcal{C}$ be a simple closed curve in D with positive orientation. Then, for every α inside $\mathcal{C}$:*

$$\frac{1}{2\pi i}\oint_{\mathcal{C}}\frac{f(z)}{z-\alpha}dz = f(\alpha) \tag{8.14}$$

If α is not inside $\mathcal{C}$, the integral vanishes as a consequence of the Cauchy integral theorem 8.6.

Corollary 8.3. ***(Derivatives of an analytic function with the Cauchy integral formula [12, page 144])*** *Let f be an analytic function in a simple connected domain D. Let $\mathcal{C}$ be a simple closed curve in D with positive orientation. Then, f is in infinitely differentiable and for every α inside $\mathcal{C}$:*

$$f^{(n)}(\alpha) = \frac{n!}{2\pi i}\oint_{\mathcal{C}}\frac{f(z)}{(z-\alpha)^{n+1}}dz \tag{8.15}$$

Definition 8.3. ***(C-analytic function [10, page 86])*** *Let K be a connected and compact set of the complex plane. We say f is C-analytic in K if it is analytic in the interior of K and continuous in the boundary of K.*

Theorem 8.9. ***(l'Hôpital rule [5][page 241])*** *Let f and g be analytic on a neighbourhood of the point c, at which $f(c) = g(c) = 0$. Then*

$$\lim_{z\to c}\frac{f(z)}{g(z)} \text{ exists and equals } \lim_{z\to c}\frac{f'(z)}{g'(z)} \tag{8.16}$$

provided that this last limit exists.

Theorem 8.10. ***(Schwarz Reflection Principle [10, page 101])*** *Suppose f is C-analytic in a region D that is contained either in the upper or lower half plane and whose boundary contains a segment L on the real axis. Suppose f is real for real z. Then, we can define an analytic extension g of f to the region $D \cup L \cup D^*$ that is symmetric with respect to the real axis by setting:*

$$g(z) = \begin{cases} f(z) & z \in D \cup L \\ \overline{f(\overline{z})} & z \in D^* \end{cases} \tag{8.17}$$

Theorem 8.11. ***(Rewriting $z = \arcsin(a)$ for $a \in \mathbb{R}$, $|a| > 1$)*** *We can rewrite the equality $z = \arcsin(a)$ for $a \in \mathbb{R}$, $|a| > 1$ in a more convenient way as:*

$$z = \left(\frac{\pi}{2} + 2\pi m\right) \pm i \ln\left(a + \sqrt{a^2-1}\right) \quad m \in \mathbb{Z} \tag{8.18}$$

If we consider only the main branch of the logarithm (restricting the argument of a complex number to $[0, 2\pi)$), then the result is:

$$z = \frac{\pi}{2} \pm i \ln\left(a + \sqrt{a^2-1}\right) \tag{8.19}$$

Proof. The proof is just rearranging the initial equality:

$$\begin{gathered} z = \arcsin(a) \Longrightarrow a = \sin(z) = \frac{e^{iz} - e^{-iz}}{2i} \Longrightarrow 2ai = e^{iz} - e^{-iz} \Longrightarrow \\ \Longrightarrow 2aie^{iz} = (e^{iz})^2 - 1 \Longrightarrow (e^{iz})^2 - 2aie^{iz} - 1 = 0 \Longrightarrow \\ \Longrightarrow e^{iz} = \frac{2ai \pm \sqrt{-4a^2+4}}{2} = ai \pm \sqrt{-(a^2-1)} = i\left(a \pm \sqrt{(a^2-1)}\right) = e^{i\frac{\pi}{2}}\left(a \pm \sqrt{(a^2-1)}\right) = \\ = e^{i\frac{\pi}{2}}\left(a \pm \sqrt{(a^2-1)}\right) e^{2m\pi i} \quad m \in \mathcal{Z} \end{gathered} \tag{8.20}$$

The term $e^{2m\pi i} \quad m \in \mathcal{Z}$ is just to account for the possible rotations over the complex plane. Taking logarithms to both sides on the last equality:

$$\begin{aligned} iz &= \frac{i\pi}{2} + \ln\left[a \pm \sqrt{(a^2-1)}\right] + 2m\pi i \quad m \in \mathbb{Z} \Longrightarrow \\ z &= \left(\frac{\pi}{2} + 2m\pi\right) - i\ln\left[a \pm \sqrt{(a^2-1)}\right] \quad m \in \mathbb{Z} \end{aligned} \tag{8.21}$$

In order to simplify more, note that:

$$\begin{aligned} \left(a - \sqrt{a^2-1}\right)\left(a + \sqrt{a^2-1}\right) = a^2 - (a^2-1) = 1 \Longrightarrow \\ \Longrightarrow \ln\left(a + \sqrt{a^2-1}\right) = -\ln\left(a - \sqrt{a^2-1}\right) \end{aligned} \tag{8.22}$$

Which allows to express the final solution as:

$$z = \left(\frac{\pi}{2} + 2\pi m\right) \pm i\ln\left(a + \sqrt{a^2-1}\right) \quad m \in \mathbb{Z} \tag{8.23}$$

The final results follows from taking only the main branch of the logarithm (restricting the argument of a complex number to $[0, 2\pi)$). □

8.4 Classical results for integrals

Theorem 8.12. ***(Monotone convergence theorem for integrals [11, page 370])*** *Let (X, Σ, μ) be a measure space. Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of nonnegative measurable functions such that $f_n(x) \leq f_{n+1}(x)\ \forall x \in X$. Define f(x) = $\lim_{n\to\infty} f_n(x)$ for each $x \in X$. Then:*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_x f d\mu \tag{8.24}$$

Theorem 8.13. ***(Dominated convergence theorem [11, page 376])*** *Let (X, Σ, μ) be a measure space and $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of measurable functions on X for which $\{f_n\} \to f$ pointwise a.e on X and the function f is measurable. Assume there is a nonnegative function g integrable over X such that $|f_n(x)| \leq g(x)$ a.e on $X\ \forall\ n$. Then, f is integrable over X and:*

$$\lim_{n\to\infty} \int_X f_n d\mu = \int_X f d\mu \tag{8.25}$$

References

[1] Atkinson, K. E. (1989). *An introduction to numerical analysis* (2nd ed.). Wiley.

[2] José Javier Segura Sala. *Apuntes de la asignatura Cálculo Numérico*.

[3] Gil, A., Segura, J., & Temme, N. M. (2007). *Numerical Methods for Special Functions*. Society for Industrial and Applied Mathematics.

[4] Apostol, T. M. (2001). *Cálculo con funciones de una variable, con una introducción al álgebra lineal* (2nd ed.). Editorial Reverté.

[5] Brannan, D. A. (2006). *A first course in mathematical analysis* (Rev. ed.). Cambridge University Press.

[6] Ralston, A., & Rabinowitz, P. (2001). *A first course in numerical analysis* (2nd ed.). Dover Publications.

[7] Trefethen, L. N., & Weideman, J. A. C. (2014). *The exponentially convergent trapezoidal rule*. SIAM Review, 56(3), 385–458. https://doi.org/10.1137/130932132

[8] Whittaker, E. T. and Watson, G. N. *A Course in Modern Analysis*, 4th ed. Cambridge, England: Cambridge University Press, 1990.

[9] Weisstein, Eric W. "Bernoulli Polynomial." From MathWorld–A Wolfram Web Resource. https://mathworld.wolfram.com/BernoulliPolynomial.html

[10] Bak, J., & Newman, D. J. (2010). *Complex analysis* (3rd ed.). Springer.

[11] Royden, H. L., & Fitzpatrick, P. M. (2010). *Real analysis* (4th ed.). Pearson Education, Inc.

[12] Spiegel, M. R., Lipschutz, S., Schiller, J., & Spellman, D. (2009). *Complex variables* (2nd ed.). Schaum's Outlines.

[13] Ahlfors, L. V. (1979). *Complex analysis* (3rd ed.). McGraw-Hill.

[14] García Fernández, M. (2018). *Convergence and divergence of Fourier series* (Bachelor's thesis, Universitat de Barcelona). http://hdl.handle.net/2445/122536

[15] Hidetosi Takahasi, Masatake Mori, *Double Exponential Formulas for Numerical Integration*. Publ. Res. Inst. Math. Sci. 9 (1973), no. 3, pp. 721–741

[16] Masatake Mori, *Discovery of the Double Exponential Transformation and Its Developments*. Publ. Res. Inst. Math. Sci. 41 (2005), no. 4, pp. 897–935

[17] Schwartz, C. (1969). *Numerical integration of analytic functions*. Journal of Computational Physics, 4(1), 19–29. https://doi.org/10.1016/0021-9991(69)90037-0