UNIVERSIDAD
DE LA RIOJA

TRABAJO FIN DE ESTUDIOS

Título

Aproximantes de Padé: irracionalidad y trascendencia

Autor/es

CARLOS CARBONELL URTUBIA

Director/es

JUDIT MÍNGUEZ CENICEROS

Facultad

Facultad de Ciencia y Tecnología

Titulación

Grado en Matemáticas

Departamento

MATEMÁTICAS Y COMPUTACIÓN

Curso académico

2019-20

Aproximantes de Padé: irracionalidad y trascendencia, de CARLOS CARBONELL URTUBIA
(publicada por la Universidad de La Rioja) se difunde bajo una Licencia Creative Commons Reconocimiento-NoComercial-SinObraDerivada 3.0 Unported.
Permisos que vayan más allá de lo cubierto por esta licencia pueden solicitarse a los titulares del copyright.

© El autor, 2020
© Universidad de La Rioja, 2020
publicaciones.unirioja.es
E-mail: publicaciones@unirioja.es

UNIVERSIDAD DE LA RIOJA

Facultad de Ciencia y Tecnología

TRABAJO FIN DE GRADO

Grado en Matemáticas

Aproximantes de Padé: irracionalidad y trascendencia

Realizado por:

Carlos Carbonell Urtubia

Tutelado por:

Judit Mínguez Ceniceros

Logroño, julio de 2020

Resumen/Abstract

Resumen

En este Trabajo de Fin de Grado se hace una breve introducción a la teoría de los aproximantes de Padé, que surgen como una generalización de los polinomios de Taylor y permiten aproximar funciones más allá del disco de convergencia. Además, también se introducen los aproximantes de Hermite-Padé, que nos permiten la aproximación simultánea de varias funciones.

Primero se estudian los aproximantes de Padé y sus propiedades. Se plantea el problema de momentos y se deducen de forma razonada los aproximantes de Padé en el infinito, así como sus condiciones de ortogonalidad. Continuamos estudiando los aproximantes de Hermite-Padé, sus dos tipos y sus condiciones de ortogonalidad. Se explican los sistemas de Angelesco y los sistemas algebraicos de Chebyshev.

Una vez expuesta toda la teoría de los aproximantes, se aborda la aplicación práctica de estos como herramienta para la construcción de pruebas de irracionalidad de ciertos números como $\log 2$, π y $\zeta(3)$. Asimismo, se usan los aproximantes de Hermite-Padé para la demostración de la trascendencia del número e.

Abstract

In this Final Degree Project, we first find a brief introduction to the theory of Padé approximants, which arise as a generalization of Taylor polynomials, and which allow the approximation of functions beyond the convergence disc. Additionally, this introduction will also include the Hermite-Padé approximants, which give us the opportunity to carry out a simultaneous approximation of various functions.

First, we study Padé approximants and their properties. The moment problem is raised and Padé approximants near infinity are reasonably inferred, as well as their conditions of orthogonality. We continue studying Hermite-Padé approximants, their two types and their orthogonality conditions. We also explain the Angelesco systems and those of Chebyshev.

Once the full theory of the approximants has been presented, we deal with the practical application of these concepts as a tool for constructing irrationality proofs for certain numbers, such as $\log 2$, π and $\zeta(3)$. Also, Hermite-Padé approximants are used to demonstrate the transcendence of the number e.

Índice general

Resumen/Abstract **III**

1. Introducción **1**

2. Aproximantes de Padé **5**
2.1. Definición de los aproximantes de Padé 7
2.2. Ortogonalidad 8
2.3. Aproximantes de Padé en el infinito 9
2.4. El problema de momentos 12
2.5. Ceros y polos 13
2.6. Convergencia 14

3. Aproximantes de Hermite-Padé **15**
3.1. Definición de los aproximantes de Hermite-Padé 15
3.2. Ortogonalidad 16
3.3. Sistemas de Angelesco 20
3.4. Sistemas algebraicos de Chebyshev 21

4. Aproximantes de Padé y Hermite-Padé: irracionalidad y trascendencia **25**
4.1. Introducción 25
4.2. Irracionalidad 25
4.2.1. Resultados previos auxiliares 26
4.2.2. Irracionalidad del número e 28
4.2.3. Irracionalidad de $\log 2$ 28
4.2.4. Irracionalidad de π 32
4.2.5. Irracionalidad de $\zeta(3)$ 39
4.3. Trascendencia 44
4.3.1. Trascendencia del número e 45
4.3.2. Trascendencia de π 50
4.4. Otros resultados y problemas abiertos 50

5. Conclusiones **53**

Bibliografía **55**

Capítulo 1

Introducción

Vamos a jugar. Espero que te apetezca, va a ser divertido. Para ello tan solo vas a necesitar un papel y un bolígrafo. ¿Ya los tienes?, perfecto. Ahora coge el bolígrafo y dibuja la recta real. Selecciona al azar un número de la recta que has pintado. No estoy ahí delante para ver cuál has elegido. Pero, has de saber, querido amigo y lector, que con probabilidad uno, acabas de seleccionar un número que no puede ser escrito como una fracción de números enteros y, además, no es solución de ninguna ecuación polinómica con coeficientes racionales.

¿Cómo podemos realizar tan categóricamente esta afirmación? Detrás de ella hay una larguísima historia que se remonta a los orígenes de la humanidad y que comienza a tomar forma en una de las etapas más fecundas de la historia para el pensamiento científico: la Grecia clásica.

El origen del concepto de número se pierde en la historia y está relacionado con la necesidad de conocer el tamaño de las colecciones o la de ordenar sucesos o ceremonias.

En Grecia se comienza a preguntar por su significado profundo, por su naturaleza. Pitágoras, tras conocer que las figuras geométricas, los movimientos de los astros y la armonía musical están ordenados por números, concluye que "todo está formado de acuerdo con el número".

En la escuela formada por él en Crotona, se estudia la naturaleza del número y se produce la primera gran crisis de conocimientos de la historia de la humanidad: el descubrimiento de las *magnitudes inconmensurables*.

Se cuenta que alrededor del año 500 a.C en Crotona o Metaponto, en el Golfo de Tarento (actualmente sur de Italia), vivió Hipaso de Metaponto, miembro de la primera escuela de los pitagóricos. Hipaso realizó varios descubrimientos matemáticos notables. Por ejemplo, demostró cómo se puede construir un dodecaedro regular inscrito en una esfera.

No obstante, por lo que es más conocido Hipaso es por haber divulgado (no está claro que lo descubriera él) la inconmensurabilidad de $\sqrt{2}$ o de $\sqrt{5}$, no se sabe con certeza cuál de los dos fue. Algunos afirman que Hipaso descubrió que la diagonal del cuadrado unitario era inconmensurable (ver [13]). Mientras que otros autores como [8] creen que también pudo ser $\sqrt{5}$, al estudiar la relación entre el lado y la diagonal en un pentágono regular. Esta relación es el *número áureo* y se suele representar por la letra griega *phi*: $\phi = \frac{1+\sqrt{5}}{2}$.

Como resultado de hacer público este descubrimiento, Hipaso fue condenado y expulsado de la famosa sociedad. No se sabe con exactitud cuál fue su castigo. Algunas fuentes afirman que los miembros de la misma le condenaron al ostracismo, lo que llevó a Hipaso al suicidio. Mientras que también hay autores que aseguran que fue arrojado al mar, incluso se adorna la historia añadiendo que fue el propio Pitágoras quien lo lanzó por la borda. En los escritos de Jámblico (s. III d.C.) podemos leer:

> "Hipaso era un pitágorico, pero al haber divulgado por escrito cómo se podía construir una esfera a partir de doce pentágonos, pereció en el mar por haber cometido ese acto de impiedad. Recibió el mérito por este descubrimiento pero en realidad todo provenía de Él (Pitágoras)".

Es también en la Grecia clásica, cuando se planteó por primera vez el problema de la cuadratura del círculo. Se trataba de determinar si, mediante regla y compás, se podía construir un cuadrado de área igual a la de un círculo dado. En términos actuales, suponiendo que el círculo tiene radio 1, esto es equivalente a construir un cuadrado de lado $\sqrt{\pi}$. Con regla y compás, sólo se pueden efectuar operaciones algebraicas; más concretamente, sumas, restas, multiplicaciones, divisiones y raíces cuadradas.

Antes de continuar, traduzcamos al lenguaje actual los conceptos de los que estamos hablando. Los números *irracionales* son aquellos que no pueden expresarse como cociente de números enteros. Se denominan *algebraicos* a los números que son raíz de un polinomio no nulo con coeficientes enteros, mientras que llamamos *trascendentes* a los números que no son algebraicos. Notar que todos los números racionales son algebraicos y, por lo tanto, todos los números trascendentes son irracionales.

Los números irracionales, desde el punto de vista de su conocimiento, son más antiguos que los trascendentes. Como ya hemos visto, en la época griega ya se conocían las magnitudes inconmensurables que podrían corresponder a lo que hoy llamamos números irracionales. Pero, tenemos que esperar hasta 1844 para afirmar la existencia de los números trascendentes. Sería un sacrilegio no agradecer la gran contribución de algunos matemáticos como Pitágoras y los miembros destacados de su escuela, Arquímedes, Leonhard Euler y Johann Heinrich Lambert a las pruebas sobre irracionalidad. Y a Joseph Liouville, Charles Hermite, Ferdinand von Lindemann y Georg Cantor, entre otros, por su aportación a la trascendencia.

Pero, saber si un número es irracional o trascendente no es fácil. Si así pudiera parecer, merece la pena leer las siguientes palabras sobre la trascendencia de π del propio Hermite en 1873, en una carta a Carl Wilhelm Borchardt, tras haber probado con gran esfuerzo la trascendencia del número e:

> "No me aventuraré a la búsqueda de una demostración de la trascendencia del número π. Que otros intenten esa empresa, nadie será más feliz que yo de su éxito. Pero créame, amigo mío, que no les costará poco esfuerzo".

La irracionalidad de e fue probada por Leonhard Euler, en 1737, antes incluso que la irracionalidad de π, que no fue probada hasta 1761 por Lambert. En 1882, Lindemann probó la trascendencia de π, dando respuesta al célebre y antiguo problema de la cuadratura del círculo: ni π ni —por tanto— $\sqrt{\pi}$ son algebraicos, luego no se pueden construir con regla y compás. Así que la cuadratura del círculo es imposible. ¡Un razonamiento analítico de teoría de números consiguió dar la solución a un problema de geometría propuesto dos mil años atrás en la historia!

Además, previamente, en 1874 ya había saltado la sorpresa cuando Cantor dio un argumento según el cual «casi todos» los números son trascendentes. Cantor probó que $\mathbb{Q}$ es numerable, que el conjunto de los números algebraicos también lo es y que $\mathbb{R}$ es no numerable. Entonces, los números trascendentes forman un conjunto no numerable. ¡Demostró que los números trascendentes existían sin haber encontrado ninguno! ¡Casi como un truco de magia! Este descubrimiento también implicaba que existían más (muchos más) números irracionales que racionales. Pero, ¿sabemos reconocerlos? Es decir, si nos encontramos con un número cualquiera, ¿sabemos "clasificarlo"? Desgraciadamente, estamos muy lejos de ello.

Por ejemplo, una función que tiene una importancia significativa en la teoría de números y que está relacionada con uno de los problemas del milenio, es la función *zeta de Riemann*,

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s},$$

nombrada en honor a Bernhard Riemann (1826-1866). Euler probó que para los números pares su valor es

$$\zeta(2n) = \frac{2^{2n-1} B_n}{(2n)!} \, \pi^{2n},$$

donde B_n son los números de Bernoulli, que son racionales. De esta expresión y el resultado de Lindemann, deducimos que la trascendencia de π implica que los números de la forma $\zeta(2n)$, con n natural, son irracionales y trascendentes.

Sin embargo, no conocemos mucho sobre los valores $\zeta(2n+1)$. Sí sabemos que $\zeta(3)$ es irracional, como veremos en el capítulo 4, pero, ni siquiera sabemos si es trascendente.

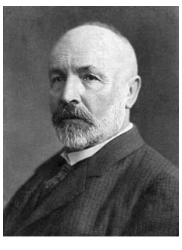

Figura 1.1: Georg Cantor (1845-1919).

Figura 1.2: Carl Louis Ferdinand von Lindemann (1852-1939).

Otro resultado muy conocido sobre trascendencia es el séptimo problema de Hilbert, que lo enunció en el Congreso Internacional de Matemáticos de 1900 y afirma que si

$\alpha \neq 0, 1$ es algebraico y β es un irracional algebraico, entonces α^β es trascendente. Hilbert llegó a decir en una clase que ninguno de sus alumnos allí presentes viviría para ver su solución. Afortunadamente, Hilbert no fue un gran profeta y años más tarde, de forma independiente, lo demostraron A. Gelfond (1934) y T. Schneider (1935).

Como vemos, el campo de la irracionalidad y trascendencia sigue guardando muchos misterios. Se conocen algunos resultados como los que veremos a lo largo del capítulo 4. Pero, aún quedan muchos problemas abiertos. Además de la función zeta de Riemann evaluada en los impares, también tenemos otros ejemplos, como la constante de Euler-Mascheroni γ, o la constante de Catalan G, de los cuales no sabemos si son trascendentes y tampoco si son irrracionales. Para una ampliación sobre irracionalidad y trascendencia, ver capítulo 6 de [24].

El objetivo principal de este trabajo es mostrar cómo los aproximantes de Padé y de Hermite-Padé han contribuido a demostrar irracionalidad y trascendencia de algunos números muy conocidos, como el número e.

Los aproximantes de Padé son funciones racionales que, en cierto sentido, generalizan los polinomios de Taylor y en algunos casos, aproximan mejor a determinadas funciones. Estos aproximantes deben su nombre a Henri Padé, que los estudió en su tesis doctoral en 1892. Sin embargo, no fue Padé el primero que trabajó con ellos. Hermite en 1873, [12], ya conocía las expresiones explícitas de estas funciones racionales para la función exponencial y le sirvieron para probar la trascendencia del número e, como veremos en el capítulo 4.

Figura 1.3: Charles Hermite (1822-1901).

Figura 1.4: Henri E. Padé (1863-1953).

Por último, el trabajo está estructurado en 5 capítulos, incluida esta introducción. En el capítulo 2 introduciremos los aproximantes de Padé y sus propiedades. El capítulo 3 lo dedicaremos a los aproximantes de Hermite-Padé, que nos servirán para aproximar simultáneamente varias funciones. Finalmente, en el capítulo 4, usaremos estos aproximantes para demostrar la irracionalidad y trascendencia de ciertos números. En particular, demostraremos que $\log 2$, π^2 y $\zeta(3)$ son irracionales; y el número e es trascendente (ver [20], [21], [22] y para una lectura más didáctica [23]). El capítulo 5 lo dedicaremos a las conclusiones.

Antes de acabar esta introducción recordamos la película *El declive del imperio americano* de 1986:

> “Hay tres cosas importantes en la historia. En primer lugar, los números; en segundo lugar, los números, y en tercer lugar, los números.”

Capítulo 2

Aproximantes de Padé

El objetivo en la teoría de la aproximación es aproximar funciones por otras más simples de tal forma que la diferencia entre la aproximación y la función sea mínima en la norma en la que estamos trabajando. La ventaja que tiene esto es que la aproximación es más simple y se puede manejar sin demasiadas dificultades, pero, la desventaja es que se pierde información ya que la función y la aproximación no son iguales.

Sea

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^k, \tag{2.1}$$

un desarrollo formal en serie de potencias centrada en $z = a$. El teorema clásico de Cauchy establece que (2.1) define una función analítica en un entorno de $z = a$ si y sólo si

$$\frac{1}{R} = \limsup_{k\to\infty} |c_k|^{1/k} < \infty.$$

En este caso, la sucesión de las sumas parciales, que son precisamente los polinomios de Taylor de f, convergen a dicha función uniformemente sobre cada subconjunto compacto del disco

$$D(a,R) = \{z \in \mathbb{C} \, : \, |z-a| < R\},$$

y divergen en el exterior del disco, $\mathbb{C}\backslash\overline{D(a,R)}$. Es decir, f no puede ser aproximada más allá del disco de convergencia por los polinomios de Taylor. Recordemos que los polinomios de Taylor P_n están caracterizados por

$$f(z) - P_n(z) = \mathcal{O}\left((z-a)^n\right), \quad z \to a.$$

Esta condición es una condición de interpolación que afirma que la diferencia $f - P_n$ tiene un cero de multiplicidad n en el punto a. Es decir, $\exists R > 0$ y $\{b_k\}_{k=n}^{\infty}$ tales que

$$f(z) - P_n(z) = \sum_{k=n}^{\infty} b_k (z-a)^k, \quad \forall z \in D(a,R) \text{ con } b_n \neq 0.$$

Pero, ¿qué ocurre si f tiene prolongación analítica más allá de de $D(a,R)$? ¿Qué ocurre si f tiene singularidades aisladas? Recordemos que f tiene una *singularidad aislada* en un punto b, si existe $S > 0$ tal que f es holomorfa en $D(b, S)\setminus\{b\}$. Además, f tiene un *polo* en b si $\lim_{z\to b} f(z) = \infty$.

Para afrontar este problema surgen los *aproximantes de Padé*, que van a ser funciones racionales, es decir, las funciones más simples con singularidades.

Antes de dar la definición, vamos a mostrar un par de ejemplos donde podemos apreciar las ventajas y desventajas de usar funciones racionales para aproximar funciones.

Ejemplo 2.1. Tomamos la función $f(z) = e^z$, que es entera (holomorfa en $\mathbb{C}$). Como sabemos, los polinomios de Taylor covergen uniformemmente a f en todo $\mathbb{C}$. Sin embargo, cuando intentamos aproximar f por una función racional R_m/Q_n, la convergencia ya no es tan buena debido a los polos del aproximante. En la figura 2.1 se muestra la función exponencial, el polinomio de Taylor de grado 4 y un aproximante de Padé donde los grados del numerador y denominador son 1.

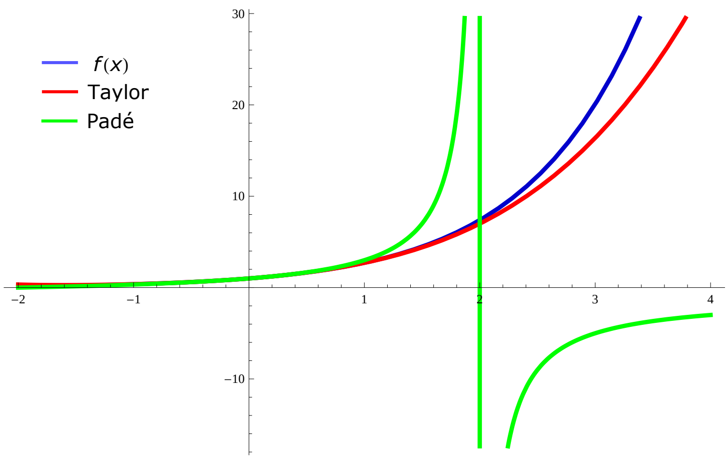

Figura 2.1: Ejemplo 2.1. Polinomio de Taylor y aproximantes de Padé de la función exponencial.

La razón de que los aproximantes de Padé aproximen peor a la función $f(z) = e^z$, es que esta es entera, mientras que los aproximantes tienen polos.

En el siguiente ejemplo vamos a tomar una función que tiene dos polos. Veamos cómo en este caso, las funciones racionales aproximan mejor a la función que los polinomios de Taylor.

Ejemplo 2.2. La función $f(z) = \frac{\operatorname{sen} z}{(3-z)^2(z+2)}$ es analítica en $\mathbb{C}\setminus\{3, -2\}$. En este caso, los polinomios de Taylor centrados en $z = 0$ solo convergen en $D(0, 2)$; mientras que los aproximantes de Padé convergen en $\mathbb{C}$. En la figura 2.2, se muestra $f(z)$, el polinomio de Taylor de grado 6 y el aproximante de Padé donde el numerador tiene grado 5 y el denominador grado 6. Podemos apreciar cómo los polos del aproximante convergen a los polos de f.

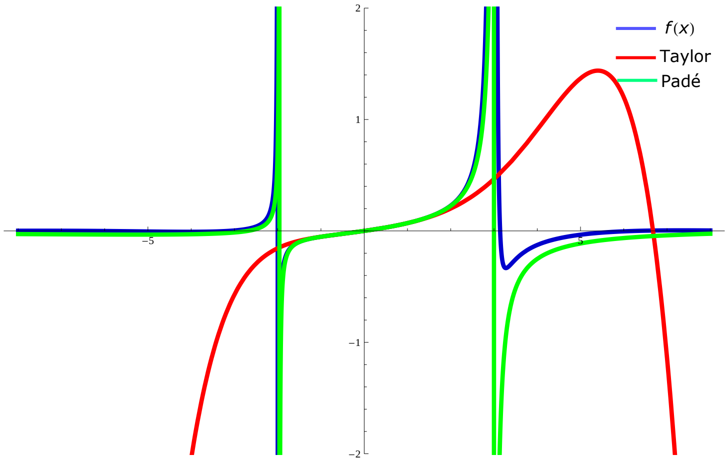

Figura 2.2: Ejemplo 2.1. Polinomio de Taylor y aproximantes de Padé de una función con polos.

2.1. Definición de los aproximantes de Padé

Definición 2.1. Sean $m, n \in \mathbb{Z}^+$ y f una función analítica en un dominio Ω tal que existe $R > 0$ con $D(a, R) \subset \Omega$. Llamamos *aproximante de Padé* de f de orden $[m, n]$ en el punto $z = a$, a la función racional $\pi_{m,n}(f) = P_m/Q_n$ tal que

i) $\operatorname{grad}(P_m) \leq m$, $\operatorname{grad}(Q_n) \leq n$, $Q_n \not\equiv 0$;

ii) $Q_n(z)f(z) - P_m(z) = \mathcal{O}\left((z-a)^{m+n+1}\right)$, donde $\mathcal{O}\left((z-a)^{m+n+1}\right)$ significa que la serie de potencias de $z - a$ comienza en la potencia $m + n + 1$.

Notemos que de la condición i) tenemos $m + n + 2$ coeficientes a determinar y de la condición ii) conseguimos $m+n+1$ ecuaciones, por lo que tenemos solución única salvo factor multiplicativo y el cociente $\pi_{m,n}$ es único. Al conjunto $\{\pi_{m,n}(f)\}_{n,m=0}^{\infty}$ se le llama *tabla de Padé de f*.

Si $n = 0$, $\pi_{m,0}(f)$ es el polinomio de Taylor de grado m en $z = a$. En este sentido, los aproximantes de Padé son una extensión de los polinomios de Taylor. No obstante, hay grandes diferencias ya que la presencia de polos puede ser un obstáculo para la convergencia.

Veamos qué supone la condición de interpolación ii) de la definición 2.1. Supongamos

$$Q_n(z)f(z) = \sum_{k=0}^{\infty} \hat{a}_k(z-a)^k, \qquad P_m(z) = \sum_{k=0}^{m} b_k(z-a)^k.$$

Entonces resulta que

$$\begin{aligned}
Q_n(z)f(z) - P_m(z) &= \hat{a}_0 + \hat{a}_1(z-a) + \cdots + \hat{a}_m(z-a)^m + \cdots + \hat{a}_{m+n+1}(z-a)^{m+n+1} + \cdots \\
&\quad - b_0 - b_1(z-a) - \cdots - b_m(z-a)^m \\
&= \hat{a}_{n+m+1}(z-a)^{n+m+1} + \hat{a}_{n+m+2}(z-a)^{n+m+2} + \cdots \\
&= \mathcal{O}\left((z-a)^{m+n+1}\right).
\end{aligned}$$

Así que los coeficientes de $(z-a)^k$ para $k = m+1, m+2, \ldots, m+n$ del polinomio de Taylor de $Q_n f$ son todos cero. El polinomio P_m corresponde al polinomio de Taylor de grado m de $Q_n f$.

Hay un grado de libertad ya que podemos multiplicar por una constante a ambos lados de la condición ii) de la definición 2.1.

Los aproximantes de Padé han sido ampliamente estudiados por Montessus de Ballore [16], Gonchar [10], Suetin [17] y [18], Markov [15], Baker y Graves-Morris [3]. Para una introducción más didáctica, ver [14], [22] y [23].

2.2. Ortogonalidad

Supongamos que f es analítica en un dominio Ω que contiene al punto a y cuyo desarrollo en serie en torno a este punto viene dado por (2.1). Tomamos un camino Γ que rodee al punto a una vez y en sentido positivo. Dividiendo ambos lados de la condición ii) de la definición 2.1 por $(z-a)^{m+k+2}$ e integrando, nos queda

$$\frac{1}{2\pi i}\int_\Gamma \frac{Q_n(z)f(z)}{(z-a)^{m+k+2}}dz - \frac{1}{2\pi i}\int_\Gamma \frac{P_m(z)}{(z-a)^{m+k+2}}dz = \sum_{j=m+n+1}^{\infty} \frac{b_{n,j}}{2\pi i}\int_\Gamma (z-a)^{j-m-k-2}dz,$$

donde los $b_{n,j}$'s son los coeficientes del polinomio de Taylor de $Q_n f - P_m$ en torno al punto a. La integral que contiene a P_m es cero para $k \geq 0$, ya que que por la fórmula de Cauchy para las derivadas, es proporcional a la derivada de orden $(m+k+1)$ de P_m (polinomio de grado a lo sumo m). En la suma de la parte de la derecha, por el teorema de los residuos, el único término distinto de cero resulta ser cuando $j = m+k+1$. Por la definición del sumatorio, se tiene que cumplir que $j \geq m+n+1$; y para que no sea nulo, $j = m+k+1$. Por lo tanto, si $k \leq n-1$, los sumandos son nulos. Esto implica que

$$\frac{1}{2\pi i}\int_\Gamma \frac{Q_n(z)}{(z-a)^{m+k+2}} f(z)dz = 0, \quad k = 0, 1, \ldots, n-1.$$

Tomando $Q_n(z) = \sum_{k=0}^{n} a_k(z-a)^{n-k}$, esto se traduce en

$$\sum_{j=0}^{n} a_j \frac{1}{2\pi i}\int_\Gamma (z-a)^{n-j-m-k-2} f(z)dz = 0, \quad k = 0, 1, \ldots, n-1.$$

Ahora, con la expresión (2.1) y la fórmula de Cauchy, tenemos que

$$\frac{1}{2\pi i}\int_\Gamma (z-a)^{n-j-m-k-2} f(z)dz = c_{m-n+k+j+1}.$$

Así que con las dos ecuaciones anteriores, podemos plantear el siguiente sistema

$$\begin{pmatrix} c_{m-n+1} & c_{m-n+2} & \cdots & c_{m+1} \\ c_{m-n+2} & c_{m-n+3} & \cdots & c_{m+2} \\ \vdots & \vdots & \cdots & \vdots \\ c_m & c_{m+1} & \cdots & c_{m+n} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.2}$$

Hay un grado de libertad, ya que tenemos $n+1$ incógnitas y n ecuaciones (homogéneas). La elección $a_0 = 1$ (si es posible) da el polinomio mónico Q_n, pero a veces se usa otra normalización, como veremos más adelante.

2.3. Aproximantes de Padé en el infinito

Veamos cómo se pueden obtener los aproximantes de Padé cerca del infinito. Esto se puede conseguir fácilmente a partir del aproximante de Padé en torno a cero y el cambio de variable $z \mapsto 1/z$. De hecho, si f^* tiene como expresión de Taylor

$$f^*(z) := \sum_{k=0}^{\infty} c_k z^k$$

cerca del origen, entonces $f(z) := f^*(1/z)/z$ tiene una expresión en series de Laurent cerca del infinito de la forma

$$f(z) = \sum_{k=0}^{\infty} \frac{c_k}{z^{k+1}}.$$

Como $f(z) = \mathcal{O}(1/z)$, la única elección razonable del grado del aproximante racional es coger $m = n-1$ para que $P_m(z)/Q_n(z)$ sea también $\mathcal{O}(1/z)$. En particular, esta situación ocurre cuando f es de la forma

$$f(z) = \int_{-\infty}^{\infty} \frac{d\mu(x)}{z-x}, \tag{2.3}$$

con μ medida real positiva. Entonces

$$f(z) = \frac{1}{z}\int_{-\infty}^{\infty} \frac{d\mu(x)}{1-\frac{x}{z}} = \frac{1}{z}\int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \left(\frac{x}{z}\right)^n d\mu(x),$$

y se ve que $f(z) = \mathcal{O}(1/z)$. Cuando f es de la forma (2.3), se dice que f es *la transformada de Stieltjes (o transformada de Cauchy).*

Los aproximantes de Padé cerca del infinito pueden ser obtenidos a partir de los aproximantes de Padé cerca del cero de la siguiente manera. El aproximante de Padé $[n-1, n]$ de la forma P^*_{n-1}/Q^*_n para f^* cerca del cero tiene la condición de interpolación

$$Q^*_{\mathrm{n}}(z) f^*(z) - P^*_{\mathrm{n}-1}(z) = \mathcal{O}\left(z^{2n}\right), \quad z \to 0.$$

Haciendo el cambio de variable $z \mapsto 1/z$ y usando que $f(z) := f^*(1/z)/z$, obtenemos

$$Q_n^*(1/z)\, zf(z) - P_{n-1}^*(1/z) = \mathcal{O}\left(z^{-2n}\right), \quad z \to \infty.$$

Para obtener polinomios, multiplicamos ambos lados por z^{n-1} y se tiene

$$Q_n(z)f(z) - P_{n-1}(z) = \mathcal{O}\left(z^{-n-1}\right), \quad z \to \infty, \tag{2.4}$$

donde $Q_n(z) := z^n Q_n^*(1/z)$ y $P_{n-1}(z) := z^{n-1}P_{n-1}^*(1/z)$ se han obtenido invirtiendo los polinomios Q_n^* y P_{n-1}^*. Así que las condiciones de interpolación en el infinito son las dadas en (2.4). Y el sistema de ecuaciones (2.2) para f^* y $m = n-1$ se convierte en

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{2.5}$$

y como

$$Q_n^*(x) = \sum_{k=0}^{n} a_k z^{n-k},$$

conseguiremos los $n+1$ coeficientes desconocidos de

$$Q_n(z) = z^n Q_n^*(1/z) = z^n \sum_{k=0}^{n} a_k z^{k-n} = \sum_{k=0}^{n} a_k z^k.$$

Tomemos ahora f una función cuya serie de Laurent en el ∞ viene dada por

$$f(z) = \sum_{k=0}^{\infty} \frac{c_k}{z^{k+1}}.$$

Definimos un funcional lineal $\mathcal{L}$ en el espacio lineal de los polinomios de la forma

$$\begin{aligned} \mathcal{L}\colon \mathbb{P}_n &\to \mathbb{R} \\ x^n &\mapsto \mathcal{L}(x^n) = c_n, \qquad n = 1, 2, \ldots. \end{aligned}$$

Notar que por linealidad, dado un polinomio $p(x) = \sum_{k=0}^{n} a_k x^k$, entonces se tiene que $\mathcal{L}(p) = \sum_{k=0}^{n} a_k c_k$. Si ahora miramos el sistema de ecuaciones (2.5), vemos que es equivalente a que los coeficientes de Q_n satisfagan las siguientes n ecuaciones

$$\sum_{j=0}^{n} a_j c_{k+j} = 0, \quad k = 0, 1, \ldots, n-1. \tag{2.6}$$

Lo que además equivale a

$$\mathcal{L}\left(x^k Q_n(x)\right) = 0, \quad k = 0, 1, \ldots, n-1. \tag{2.7}$$

Por lo que el polinomio Q_n es ortogonal a todos los polinomios de grado menor que n respecto al funcional lineal $\mathcal{L}$. También se puede normalizar exigiendo que $\mathcal{L}(Q_n^2(x)) = 1$, si el funcional es positivo. Cuando el funcional no es positivo, exigimos

$$\mathcal{L}\left(\left(\frac{Q_n(x)}{\sqrt{h_n}}\right)^2\right) = 1, \qquad \text{donde } h_n = \mathcal{L}(Q_n^2(x)) \neq 0.$$

Una vez obtenido el polinomio Q_n, los elementos restantes en el problema de aproximación de Padé pueden ser encontrados explícitamente en términos de Q_n (recordemos que P_{n-1} es el polinomio de Taylor de grado $n-1$ de $Q_n f$). En efecto, definiendo

$$P_{n-1}(z) := \mathcal{L}\left(\frac{Q_n(z) - Q_n(x)}{z - x}\right), \tag{2.8}$$

ya que $(Q_n(z) - Q_n(x))/(z - x)$ es un polinomio de grado $n-1$ en la variable z, y P_{n-1} es un polinomio de grado $n-1$, aplicando la linealidad del funcional resulta

$$\begin{aligned} P_{n-1}(z) &= \mathcal{L}\left(\frac{Q_n(z) - Q_n(x)}{z - x}\right) = \mathcal{L}\left(\frac{Q_n(z)}{z - x}\right) - \mathcal{L}\left(\frac{Q_n(x)}{z - x}\right) \\ &= Q_n(z)\mathcal{L}\left(\frac{1}{z - x}\right) - \mathcal{L}\left(\frac{Q_n(x)}{z - x}\right), \end{aligned}$$

y (2.8) es equivalente a

$$Q_{\mathrm{n}}(z)\mathcal{L}\left(\frac{1}{z - x}\right) - P_{n-1}(z) = \mathcal{L}\left(\frac{Q_n(x)}{z - x}\right).$$

Extendiendo formalmente la definición del funcional $\mathcal{L}$ para series de Laurent, si ahora expresamos $1/(z - x)$ como una serie de Laurent, obtenemos

$$\mathcal{L}\left(\frac{1}{z - x}\right) = \mathcal{L}\left(\sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}}\right) = \sum_{k=0}^{\infty} \frac{c_k}{z^{k+1}} = f(z).$$

Usando la serie de Laurent de $1/(z - x)$ y la linealidad del funcional $\mathcal{L}$, vamos a ver que

$$\mathcal{L}\left(\frac{Q_n(x)}{z - x}\right) = \mathcal{O}\left(z^{-n-1}\right).$$

Operamos aplicando la linealidad del funcional

$$\mathcal{L}\left(\frac{Q_n(x)}{z - x}\right) = \mathcal{L}\left(\left(\sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}}\right) Q_n(x)\right) = \sum_{k=0}^{\infty} \mathcal{L}\left(\frac{x^k}{z^{k+1}}\, Q_n(x)\right) = \sum_{k=0}^{\infty} \frac{1}{z^{k+1}}\, \mathcal{L}\left(x^k\, Q_n(x)\right)$$

y gracias a las condiciones de ortogonalidad (2.7), los términos del sumatorio anterior empiezan a ser distintos de cero para $k \geq n$ y así el sumatorio es una $\mathcal{O}\left(1/z^{\mathrm{n}+1}\right)$. Además, de esta prueba también deducimos que el error en la aproximación de Padé está dado explícitamente por

$$Q_n(z)f(z) - P_{n-1}(z) = \mathcal{L}\left(\frac{Q_n(x)}{z - x}\right),$$

el cual está en términos del polinomio Q_n.

2.4. El problema de momentos

El funcional lineal $\mathcal{L}$ está relacionado con la función f, pero nos gustaría saber algo más explícito. El teorema de representación de Riesz nos dice que todo funcional lineal positivo y acotado, en un espacio lineal de funciones continuas con soporte compacto en la recta real, puede ser representado por una medida μ positiva y finita en la recta real como

$$\mathcal{L}(f) = \int_{-\infty}^{\infty} f(x) d\mu(x).$$

Para conseguir la convergencia de los aproximantes de Padé, es conveniente trabajar con un funcional lineal, positivo y acotado $\mathcal{L}$, el cual está representado por una medida μ positiva y finita. En ese caso

$$c_k = \int_{-\infty}^{\infty} x^k d\mu(x) \tag{2.9}$$

serán los momentos de la medida μ positiva y la función f es la transformada de Cauchy (de Stieltjes) de la medida μ:

$$f(z) = \int_{-\infty}^{\infty} \frac{1}{z-x} d\mu(x).$$

Como es de esperar, no todas las sucesiones $c_0, c_1, c_2, \ldots$ darán lugar a un funcional positivo y acotado. El problema de momentos es obtener condiciones en esta sucesión $c_0, c_1, c_2, \ldots$ que garanticen que sus términos son los momentos de una medida positiva y finita en la recta real, como en (2.9). Una condición necesaria y suficiente para que $c_0, c_1, c_2, \ldots$ sean los momentos de una medida positiva en $(-\infty, \infty)$ es que todas las matrices

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & \cdots & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \end{pmatrix},$$

que se denominan *matrices de Hankel*, sean definidas positivas. Observar que estas son precisamente las matrices que aparecen en (2.5). Si añadimos una restricción más: que la medida esté soportada en un intervalo acotado $[a, b]$, se facilita el trabajo, ya que se evita la no compacidad del soporte. Así que, nuestra función f será de la siguiente forma, que es lo que se conoce como una *función de Markov*

$$f(z) = \int_a^b \frac{1}{z-x} d\mu(x). \tag{2.10}$$

Veamos que toda función de Markov es analítica en $\mathbb{C}\backslash[a, b]$. Sea $z_0 \in \mathbb{C}\backslash[a, b]$,

$$\begin{aligned} \frac{1}{z-x} &= \frac{1}{z - z_0 - x + z_0} = \frac{1}{z_0 - x} \frac{1}{1 - \frac{z-z_0}{x-z_0}} = \frac{-1}{x - z_0} \sum_{n=0}^{\infty} \left(\frac{z - z_0}{x - z_0} \right)^n \\ &= \sum_{n=0}^{\infty} \frac{-1}{(x - z_0)^{n+1}} (z - z_0)^n, \qquad \text{si } |z - z_0| < |x - z_0|. \end{aligned}$$

Así que eligiendo un radio de convergencia R tal que $R < \text{mín}\{|x - z_0| : x \in [a, b]\}$ y una sucesión

$$b_n = \int_a^b \frac{-1}{(x - z_0)^{n+1}} d\mu(x),$$

se tiene que

$$f(z) = \int_a^b \frac{1}{z - x} d\mu(x) = \sum_{n=0}^{\infty} b_n (z - z_0)^n.$$

Así f es analítica en $\mathbb{C} \backslash [a, b]$ y las singularidades de esta función se encuentran en el intervalo $[a, b]$.

En este caso, el funcional lineal está dado por

$$\mathcal{L}(g) = \int_a^b g(x) d\mu(x),$$

para toda función continua g en $[a, b]$. Y las condiciones de ortogonalidad (2.7) nos quedan

$$\int_a^b x^k Q_n(x) d\mu(x) = 0, \quad k = 0, 1, \ldots, n - 1.$$

Es decir, los polinomios del denominador en el problema de aproximación de Padé son los polinomios ortogonales respecto a la medida μ en el intervalo $[a, b]$.

La normalización $\mathcal{L}(Q_n^2(x)) = 1$ se convierte en

$$\int_a^b Q_n^2(x) d\mu(x) = 1.$$

Los polinomios del numerador vienen dados por

$$P_{n-1}(z) = \int_a^b \frac{Q_n(z) - Q_n(x)}{z - x} d\mu(x)$$

y el error es

$$Q_n(z) f(z) - P_{n-1}(z) = \int_a^b \frac{Q_n(x)}{z - x} d\mu(x). \tag{2.11}$$

2.5. Ceros y polos

Recordemos que la idea es que los polos de las funciones racionales tenderán a los polos de la función f, y por lo tanto, el dominio de convergencia podría ampliarse y se podrían descubrir singularidades de f usando los polos de las aproximaciones racionales. Esto es así cuando f es una función de Markov. Las singularidades de los aproximantes de Padé son polos en los ceros de Q_n. Una consecuencia de la ortogonalidad es que estos ceros son simples y todos están en el intervalo abierto (a, b).

Teorema 2.1. *Supongamos que el soporte de μ es un conjunto infinito en $[a, b]$ (no es una medida discreta). Entonces todas las raíces de Q_n son simples y están localizadas en (a, b).*

Demostración. Sean $x_1, \ldots, x_m$ los cambios de signo de Q_n en (a, b), entonces obviamente $m \leq n$ ya que cada cambio de signo es una raíz. Supongamos que Q_n tiene raíces en (a, b) que no son simples, entonces $m < n$. Definimos el polinomio $\pi_m(x) := (x - x_1)(x - x_2) \cdots (x - x_m)$. La función $Q_n(x)\pi_m(x)$ no cambia de signo en $[a, b]$ y el soporte de μ contiene infinitos puntos, por lo que

$$\int_a^b Q_n(x)\pi_m(x)d\mu(x) \neq 0.$$

Pero Q_n es ortogonal a todos los polinomios de grado $< n$, por lo tanto, esta integral es igual a 0. Esta contradicción implica que $m = n$. Así, Q_n tiene n ceros simples en (a, b). $\square$

2.6. Convergencia

Por la fórmula del error (2.11) tenemos que

$$f(z) - \frac{P_{n-1}(z)}{Q_n(z)} = \frac{1}{Q_n(z)} \int_a^b \frac{Q_n(x)}{z - x} d\mu(x).$$

Notar que

$$Q_n(z) \int_a^b \frac{Q_n(x)}{z - x} d\mu(x) = \int_a^b \frac{Q_n(x)(Q_n(z) - Q_n(x))}{z - x} d\mu(x) + \int_a^b \frac{Q_n^2(x)}{z - x} d\mu(x).$$

Como $Q_n(x)$ es ortogonal con respecto a la medida μ, la primera integral de la derecha es 0. Por tanto, nos queda que

$$Q_n(z) \int_a^b \frac{Q_n(x)}{z - x} d\mu(x) = \int_a^b \frac{Q_n^2(x)}{z - x} d\mu(x),$$

y entonces el error de aproximación será

$$f(z) - \frac{P_{n-1}(z)}{Q_n(z)} = \frac{1}{Q_n^2(z)} \int_a^b \frac{Q_n^2(x)}{z - x} d\mu(x).$$

Supongamos que $z \in K$ con $K \subset \mathbb{C} \setminus [a, b]$ compacto y que Q_n es el polinomio ortonormal con respecto a μ. Entonces,

$$\left| \int_a^b \frac{Q_n^2(x)}{z - x} d\mu(x) \right| \leq \int_a^b \frac{Q_n^2(x)}{|z - x|} d\mu(x) \leq \frac{1}{d_K},$$

donde

$$d_K := \inf\{|z - x| : z \in K, x \in [a, b]\}.$$

Esta cota no depende de n, por tanto la convergencia de los aproximantes de Padé a f está determinada por el comportamiento de Q_n.

Capítulo 3

Aproximantes de Hermite-Padé

Supongamos que en vez de querer aproximar una sola función mediante los aproximantes de Padé, se desea aproximar varias. Para cada una de ellas hay que calcular un numerador y un denominador. El proceso parece resultar bastante arduo y costoso. ¿Y si por ejemplo, algunos de los denominadores que calculamos para una función pudiéramos usarlos para otra? ¿O que una combinación lineal de los aproximantes ya calculados nos sirviera para calcular otros? Esta es la idea de los aproximantes de Hermite-Padé.

3.1. Definición de los aproximantes de Hermite-Padé

La aproximación de Hermite-Padé es la aproximación racional simultánea a un vector de r funciones $f_1, f_2, \ldots, f_r$, dadas como series de Taylor alrededor de un punto $a \in \mathbb{C}$, para las cuales vamos a exigir ciertas condiciones de interpolación en a. De forma análoga a la que se ha procedido con los aproximantes de Padé, invirtiendo los polinomios del numerador y denominador con el cambio de variable $z \mapsto 1/z$, se consigue la aproximación de Hermite-Padé alrededor de infinito, imponiendo condiciones de interpolación en el infinito. Por la limitación de extensión de esta memoria, nos centraremos solo en el estudio de los aproximantes de Hermite-Padé alrededor de infinito.

Supongamos que tenemos r funciones con sus series de Laurent

$$f_j(z) = \sum_{k=0}^{\infty} \frac{c_{k,j}}{z^{k+1}}, \quad j = 1, 2, \ldots, r.$$

En esencia, hay dos tipos diferentes de aproximación de Hermite-Padé. Aunque en la práctica se usan combinaciones que mezclan ambos tipos. Antes de explicar cada uno, necesitamos introducir el concepto de *multi-índice*. Se trata de un vector $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$ para el cual definimos también su norma o tamaño como $|\vec{n}| = n_1 + n_2 + \cdots + n_r$.

Definición 3.1. La *aproximación de Hermite-Padé tipo I para el vector* $(f_1, \ldots, f_r)$ cerca del infinito consiste en buscar un vector $\left(A_{\vec{n},1}, \ldots, A_{\vec{n},r}\right)$ de polinomios y un polinomio $B_{\vec{n}}$ tales que

i) $\mathrm{grad}(A_{\vec{n},j}) \leq n_j - 1$;

ii) $\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \mathcal{O}\left(\frac{1}{z^{|\vec{n}|}}\right), \quad z \to \infty.$

En la aproximación tipo I de Hermite-Padé se busca aproximar por un polinomio a una combinación lineal de las r funciones (cuyos coeficientes son polinomios).

El multi-índice $\vec{n}$ se dice *normal* para la aproximación tipo I si el grado de cada $A_{\vec{n},j}$ es exactamente $n_j - 1$.

Definición 3.2. La *aproximación tipo II de Hermite-Padé del vector* $(f_1, \ldots, f_r)$ cerca del infinito consiste en buscar polinomios $Q_{\vec{n}}$, $P_{\vec{n},j}$ $(j = 1, 2, \ldots, r)$ tales que

i) $\mathrm{grad}(Q_{\vec{n}}) \leq |\vec{n}|$;

ii) $Q_{\vec{n}}(z) f_j(z) - P_{\vec{n},j}(z) = \mathcal{O}\left(\frac{1}{z^{n_j+1}}\right), \quad z \to \infty, \quad j = 1, \ldots r.$

En la aproximación tipo II cada función f_j es aproximada de forma individual por funciones racionales con un denominador común $Q_{\vec{n}}$.

El multi-índice $\vec{n}$ se dice *normal* para el tipo II si el grado de $Q_{\vec{n}}$ es exactamente $|\vec{n}| = n_1 + \cdots + n_r$.

Notar que en ambos casos, cuando $r = 1$, los aproximantes de Hermite-Padé se reducen a los aproximantes de Padé. Sin embargo, estas generalizaciones son bastante más complejas, ya que la unicidad no está garantizada, a no ser que el multi-índice sea normal, que en nuestros casos lo será.

La aproximación de Hermite-Padé ha sido ampliamente estudiada para distintos sistemas de funciones [2], [6], [7] y [11].

3.2. Ortogonalidad

Si consideramos r funciones de Markov

$$f_j(z) = \int_{a_j}^{b_j} \frac{d\mu_j(x)}{z - x}, \quad j = 1, 2, \ldots, r,$$

entonces la aproximación de Hermite-Padé implica nuevas condiciones de ortogonalidad.

Vamos a considerar primero la aproximación tipo I. Multiplicando por z^k la condición ii) de la definición 3.1 e integrando sobre un camino Γ en cuyo interior estén contenidos todos los intervalos $[a_j, b_j]$ y recorrido en sentido positivo, resulta

$$\frac{1}{2\pi i} \int_\Gamma \left(\sum_{j=1}^{r} z^k A_{\vec{n},j}(z) f_j(z) \right) dz - \frac{1}{2\pi i} \int_\Gamma z^k B_{\vec{n}}(z) dz = \sum_{\ell=|\vec{n}|}^{\infty} b_{\vec{n},\ell} \frac{1}{2\pi i} \int_\Gamma z^{k-\ell} dz, \tag{3.1}$$

donde los $b_{\vec{n},\ell}$ son los coeficientes de la serie de Laurent de $\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z)$. Por el teorema de Cauchy, ya que $z^k B_{\vec{n}}(z)$ es una función entera, tenemos que

$$\frac{1}{2\pi i} \int_\Gamma z^k B_{\vec{n}}(z) dz = 0, \qquad \text{para todo } k \geq 0.$$

En la suma de la parte de la derecha de (3.1), por el teorema de los residuos, el único término distinto de cero es $\ell = k+1$. Por lo tanto

$$\frac{1}{2\pi i}\int_\Gamma \left(\sum_{j=1}^{r} z^k A_{\vec{n},j}(z) f_j(z)\right) dz = 0, \quad k = 0, 1, \ldots, |\vec{n}| - 2.$$

Ahora, como cada f_j es una función de Markov, $z^k A_{\vec{n},j}(z) f_j(z)$ es integrable e intercambiando el orden de integración por el teorema de Fubini, obtenemos

$$\begin{aligned}\frac{1}{2\pi i}\int_\Gamma z^k A_{\vec{n},j}(z) f_j(z) dz &= \frac{1}{2\pi i}\int_\Gamma z^k A_{\vec{n},j}(z)\left(\int_{a_j}^{b_j} \frac{1}{z-x} d\mu_j(x)\right) dz \\ &= \int_{a_j}^{b_j} d\mu_j(x)\, \frac{1}{2\pi i}\int_\Gamma \frac{z^k A_{\vec{n},j}(z)}{z-x}\, dz = \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) d\mu_j(x).\end{aligned}$$

Para la última igualdad hemos usado la fórmula de Cauchy.

Por lo que obtenemos las siguientes condiciones de ortogonalidad

$$\sum_{j=1}^{r} \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) d\mu_j(x) = 0, \quad k = 0, 1, \ldots, |\vec{n}| - 2. \tag{3.2}$$

Hay $|\vec{n}| - 1$ ecuaciones lineales y homogéneas. Como cada polinomio $A_{\vec{n},j}$ es de grado a lo sumo $n_j - 1$, tenemos n_j coeficientes. Ahora, como hay r polinomios, tendremos, como mucho, un total de $n_1 + n_2 + \ldots + n_r = |\vec{n}|$ coeficientes desconocidos. Por lo que podemos determinar estos polinomios salvo un factor multiplicativo, siempre que la matriz en este sistema tenga rango $|\vec{n}| - 1$. Lo que es equivalente a que $\vec{n}$ sea un multi-índice normal para el tipo I, es decir, el grado de cada $A_{\vec{n},j}$ sea exactamente $n_j - 1$.

Una vez los polinomios $(A_{\vec{n},1}, \ldots, A_{\vec{n},r})$ están determinados, tomamos $B_{\vec{n}}$ como sigue,

$$B_{\vec{n}}(z) = \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(z) - A_{\vec{n},j}(x)}{z-x} d\mu_j(x).$$

En efecto, con esta definición de $B_{\vec{n}}$, operando y teniendo presente la definición de cada $f_j(z)$ como una función de Markov, obtenemos

$$\begin{aligned}\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) &= \sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(z) - A_{\vec{n},j}(x)}{z-x} d\mu_j(x) \\ &= \sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - \left(\sum_{j=1}^{r} A_{\vec{n},j}(z) \int_{a_j}^{b_j} \frac{d\mu_j(x)}{z-x} - \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(x)}{z-x} d\mu_j(x)\right) \\ &= \sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - \left(\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(x)}{z-x} d\mu_j(x)\right) \\ &= \sum_{j=1}^{r} \int_{a_j}^{b_j} \frac{A_{\vec{n},j}(x)}{z-x} d\mu_j(x).\end{aligned}$$

Definimos por comodidad $R_1(z) = \sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z)$. Si ahora usamos la serie

$$\frac{1}{z-x} = \sum_{k=0}^{\infty} \frac{x^k}{z^{k+1}},$$

entonces $R_1(z)$ es

$$\begin{aligned} R_1(z) = & \int_{a_1}^{b_1} A_{\vec{n},1}(x) \frac{x^0}{z^{0+1}} d\mu_1(x) + \ldots + \int_{a_1}^{b_1} A_{\vec{n},1}(x) \frac{x^m}{z^{m+1}} d\mu_1(x) + \ldots \\ & + \int_{a_2}^{b_2} A_{\vec{n},2}(x) \frac{x^0}{z^{0+1}} d\mu_2(x) + \ldots + \int_{a_2}^{b_2} A_{\vec{n},2}(x) \frac{x^m}{z^{m+1}} d\mu_2(x) + \ldots \\ & \qquad \vdots \qquad\qquad\qquad\qquad \vdots \\ & + \int_{a_r}^{b_r} A_{\vec{n},r}(x) \frac{x^0}{z^{0+1}} d\mu_r(x) + \ldots + \int_{a_r}^{b_r} A_{\vec{n},r}(x) \frac{x^m}{z^{m+1}} d\mu_r(x) + \ldots \end{aligned}$$

y sumando por columnas

$$\begin{aligned} R_1(z) &= \sum_{j=1}^{r} \int_{a_j}^{b_j} A_{\vec{n},j}(x) \frac{x^0}{z^{0+1}} d\mu_j(x) + \ldots + \sum_{j=1}^{r} \int_{a_j}^{b_j} A_{\vec{n},j}(x) \frac{x^m}{z^{m+1}} d\mu_j(x) + \ldots \\ &= \sum_{k=0}^{\infty} \left(\sum_{j=1}^{r} \int_{a_j}^{b_j} A_{\vec{n},j}(x) \frac{x^k}{z^{k+1}} d\mu_j(x) \right) = \sum_{k=0}^{\infty} \left(\frac{1}{z^{k+1}} \sum_{j=1}^{r} \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) d\mu_j(x) \right). \end{aligned}$$

Las condiciones de ortogonalidad (3.2) muestran que la suma anterior de índice k empieza en $k = |\vec{n}| - 1$, por lo tanto

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \mathcal{O}\left(\frac{1}{z^{|\vec{n}|}}\right), \quad z \to \infty,$$

y se tiene lo exigido en la definición de aproximante de Hermite-Padé tipo I.

Consideramos ahora la aproximación tipo II de Hermite-Padé. Multiplicando la condición ii) de la definición 3.2 por z^k e integrando sobre un camino Γ en cuyo interior estén contenidos todos los intervalos $[a_j, b_j]$ y recorrido en sentido positivo, para cada $1 \leq j \leq r$, resulta

$$\frac{1}{2\pi i} \int_{\Gamma} z^k Q_{\vec{n}}(z) f_j(z) dz - \frac{1}{2\pi i} \int_{\Gamma} z^k P_{\vec{n},j}(z) dz = \sum_{\ell=n_j+1}^{\infty} b_{\vec{n},j,\ell} \frac{1}{2\pi i} \int_{\Gamma} z^{k-\ell} dz, \tag{3.3}$$

donde los $b_{\vec{n},j,\ell}$ son los coeficientes de la serie de Laurent de $Q_{\vec{n}}(z) f_j(z) - P_{\vec{n},j}(z)$.

Por el teorema de Cauchy, ya que $z^k P_{\vec{n},j}(z)$ es una función entera, tenemos que

$$\frac{1}{2\pi i} \int_{\Gamma} z^k P_{\vec{n},j}(z) dz = 0, \text{ para todo } k \geq 0.$$

En la suma de la parte de la derecha de (3.3), por el teorema de los residuos, el único término distinto de cero es $\ell = k+1$. Por lo que

$$\frac{1}{2\pi i}\int_\Gamma z^k Q_{\vec{n}}(z) f_j(z)dz = 0, \quad k = 0,1,\ldots,n_j-1.$$

Como cada f_j es una función de Markov y $z^k Q_{\vec{n}}(z) f_j(z)$ es integrable, intercambiando del orden de integración por el teorema de Fubini, obtenemos

$$\begin{aligned}\frac{1}{2\pi i}\int_\Gamma z^k Q_{\vec{n}}(z) f_j(z)dz &= \frac{1}{2\pi i}\int_\Gamma z^k Q_{\vec{n}}(z)\left(\int_{a_j}^{b_j}\frac{1}{z-x}d\mu_j(x)\right)dz \\ &= \int_{a_j}^{b_j} d\mu_j(x)\,\frac{1}{2\pi i}\int_\Gamma \frac{z^k Q_{\vec{n}}(z)}{z-x}\,dz = \int_{a_j}^{b_j} x^k Q_{\vec{n}}(x)d\mu_j(x).\end{aligned}$$

Habiendo usado para la última igualdad la fórmula de Cauchy.

Por lo que, para cada $1 \leq j \leq r$, obtenemos las siguientes condiciones de ortogonalidad

$$\int_{a_j}^{b_j} x^k Q_{\vec{n}}(x)d\mu_j(x) = 0, \quad k = 0,1,\ldots,n_j-1. \tag{3.4}$$

Hay $|\vec{n}|$ ecuaciones lineales y homogéneas en total. Como el polinomio $Q_{\vec{n}}$ es de grado a lo sumo $|\vec{n}|$, tendrá a lo sumo $|\vec{n}|+1$ coeficientes. Por lo que podemos determinar este polinomio salvo un factor multiplicativo, siempre que la matriz en este sistema tenga rango $|\vec{n}|$. Lo que es equivalente a que $\vec{n}$ sea un multi-índice normal para el tipo II, es decir, el grado de $Q_{\vec{n}}$ sea exactamente $|\vec{n}|$.

Una vez el polinomio $Q_{\vec{n}}$ es determinado, definimos $P_{\vec{n},j}$, para cada $1 \leq j \leq r$ como sigue,

$$P_{\vec{n},j}(z) = \int_{a_j}^{b_j}\frac{Q_{\vec{n}}(z)-Q_{\vec{n}}(x)}{z-x}d\mu_j(x).$$

En efecto, con esta definición de $P_{\vec{n},j}$, operando y teniendo presente que $f_j(z)$ es una función de Markov, obtenemos

$$\begin{aligned}Q_{\vec{n}}(z)f_j(z) - P_{\vec{n},j}(z) &= Q_{\vec{n}}(z)f_j(z) - \int_{a_j}^{b_j}\frac{Q_{\vec{n}}(z)-Q_{\vec{n}}(x)}{z-x}d\mu_j(x) \\ &= Q_{\vec{n}}(z)f_j(z) - \left(Q_{\vec{n}}(z)f_j(x) - \int_{a_j}^{b_j}\frac{Q_{\vec{n}}(x)}{z-x}d\mu_j(x)\right) = \int_{a_j}^{b_j}\frac{Q_{\vec{n}}(x)}{z-x}d\mu_j(x).\end{aligned}$$

Usando la expresión en serie de potencias de $1/(z-x)$ y sumando por columnas de forma análoga a la que hemos procedido en los aproximantes de Hermite-Padé tipo I, resulta

$$Q_{\vec{n}}(z)f_j(z) - P_{\vec{n},j}(z) = \int_{a_j}^{b_j}\frac{Q_{\vec{n}}(x)}{z-x}d\mu_j(x) = \sum_{k=0}^{\infty}\frac{1}{z^{k+1}}\int_{a_j}^{b_j} x^k Q_{\vec{n}}(x)d\mu_j(x).$$

Las condiciones de ortogonalidad (3.4) muestran que la suma anterior de índice k empieza en $k = n_j$ por lo tanto, para cada $1 \leq j \leq r$,

$$Q_{\vec{n}}(z)f_j(z) - P_{\vec{n},j}(z) = \mathcal{O}\left(\frac{1}{z^{n_j+1}}\right), \quad z \to \infty,$$

y se tiene lo exigido en la definición de aproximante de Hermite-Padé tipo II.

3.3. Sistemas de Angelesco

En la aproximación de Hermite-Padé, que el multi-índice sea normal, es de gran importancia. Para ello, hay varias formas de definir el sistema de funciones de Markov, para que el índice múltiple sea normal. Uno de estos sistemas es lo que se conoce por *sistema de Angelesco*, llamado así por Angelesco, que introdujo la noción en 1919.

Definición 3.3. Un *sistema de Angelesco* es un conjunto de r funciones de Markov $(f_1, f_2, \ldots, f_r)$ para las cuales los intervalos (a_j, b_j) son disjuntos dos a dos.

Todos los multi-índices son normales para el tipo II en un sistema de Angelesco. Lo probamos demostrando que el polinomio ortogonal $Q_{\vec{n}}$ tiene grado exactamente $|\vec{n}|$. Además, sabemos incluso algo más, como queda de manifiesto en el siguiente teorema.

Teorema 3.1. *Si $(f_1, \ldots, f_r)$ es un sistema de Angelesco con medidas μ_j que tienen infinitos puntos en su soporte (no son medidas discretas), entonces $Q_{\vec{n}}$ tiene n_j ceros simples en (a_j, b_j) para cada $j = 1, \ldots, r$.*

Demostración. Sean $x_1, \ldots, x_m$ los cambios de signo de $Q_{\vec{n}}$ en (a_j, b_j). Supongamos que $Q_{\vec{n}}$ tiene raíces en (a_j, b_j) que no son simples, entonces $m < n_j$. Definimos el polinomio $\pi_m(x) := (x - x_1)(x - x_2) \cdots (x - x_m)$. Entonces $Q_{\vec{n}}\pi_m$ no cambia de signo en $[a_j, b_j]$. Como el soporte de μ_j tiene infinitos puntos, tenemos

$$\int_{a_j}^{b_1} Q_{\vec{n}}(x)\pi_m(x)d\mu_j(x) \neq 0.$$

Sin embargo, la ortogonalidad (3.4) implica que $Q_{\vec{n}}$ es ortogonal a todos los polinomios de grado $\leq n_j - 1$ respecto a la medida μ_j en $[a_j, b_j]$, por lo que la integral es cero. Esta contradicción implica que $m \geq n_j$, y por lo tanto $Q_{\vec{n}}$ tiene al menos n_j ceros en (a_j, b_j). Esto se cumple para cada j, y ya que los intervalos (a_j, b_j) son disjuntos, $Q_{\vec{n}}$ tiene al menos $|\vec{n}|$ raíces reales. Pero el grado de $Q_{\vec{n}}$ es $\leq |\vec{n}|$, por lo tanto $Q_{\vec{n}}$ tiene exactamente n_j ceros simples en (a_j, b_j). □

Por lo tanto, el polinomio $Q_{\vec{n}}$ puede factorizarse como

$$Q_{\vec{n}}(x) = q_{n_1}(x)q_{n_2}(x) \cdots \cdot q_{n_r}(x),$$

donde cada q_{n_j}, es un polinomio de grado n_j con sus raíces en (a_j, b_j). Las condiciones de ortogonalidad (3.4) implican que

$$\int_{a_j}^{b_j} x^k q_{n_j}(x) \prod_{i \neq j} q_{n_i}(x)d\mu_j(x) = 0, \quad k = 0, 1, \ldots, n_j - 1.$$

El producto $\prod_{i \neq j} q_{n_i}(x)$ no cambia de signo en (a_j, b_j), así que la anterior expresión implica que q_{n_j} es un polinomio ortogonal de grado n_j en el intervalo $[a_j, b_j]$ respecto a la medida $\prod_{i \neq j} |q_{n_i}(x)| \, d\mu_j(x)$. La medida depende del multi-índice $\vec{n}$.

En [9], los autores probaron que todos los multi-índices son también normales para la aproximación de Hermite-Padé tipo I para un sistema de Angelesco, pero omitiremos la demostración por ser más complicada.

3.4. Sistemas algebraicos de Chebyshev

Definición 3.4. Un *sistema de Chebyshev* $(\varphi_1, \dots, \varphi_n)$ en $[a, b]$ es un sistema linealmente independiente de n funciones tales que cualquier combinación no trivial $\sum_{k=1}^{n} a_k\varphi_k$, con $a_k \in \mathbb{R}$, tiene como mucho $n - 1$ ceros en $[a, b]$.

Esto es equivalente a la condición

$$\det \begin{pmatrix} \varphi_1(x_1) & \varphi_1(x_2) & \cdots & \varphi_1(x_n) \\ \varphi_2(x_1) & \varphi_2(x_2) & \cdots & \varphi_2(x_n) \\ \vdots & \vdots & \cdots & \vdots \\ \varphi_n(x_1) & \varphi_n(x_2) & \cdots & \varphi_n(x_n) \end{pmatrix} \neq 0,$$

para cualesquiera n puntos distintos $x_1, \dots, x_n \in [a, b]$. En efecto, cuando $x_1, \dots, x_n$ son tales que su determinante es cero, entonces existe una combinación lineal de filas que da una fila de ceros, pero esto significa que esa combinación lineal $\sum_{k=1}^{n} a_k\varphi_k$ tiene ceros en $x_1, \dots, x_n$, siendo un total de n ceros, lo cual contradice la definición 3.4.

Veamos a continuación dos ejemplos útiles que usaremos más adelante y que además nos ayudarán a entender bien el concepto.

Ejemplo 3.1. El conjunto de funciones

$$\mathcal{S} = \{1, x, \dots, x^n, \log x, x\log x, \dots, x^n \log x\}$$

es un sistema de Chebyshev en el intervalo $[0, 1]$.

Una combinación no nula de los $2n + 2$ elementos que hay, es de la forma $p_n(x) + q_n(x)\log x$, siendo p_n y q_n polinomios de grado $\leq n$. Supongamos que la combinación tiene al menos $2n + 2$ ceros. Por el teorema de Rolle, al derivar la combinación $n + 1$ veces, la función resultante tendrá al menos $2n + 2 - (n + 1) = n + 1$ ceros. Puesto que p_n y q_n son polinomios de grado n a lo sumo, cuando los derivemos $n + 1$ veces, sus derivadas de orden $n + 1$ serán cero. Aplicando la fórmula de Leibniz para las derivadas k-ésimas de un producto y la fórmula para la derivada j-ésima del logaritmo, tenemos

$$\begin{aligned}(p_n(x) + q_n(x)\log x)^{(n+1)} = (q_n(x)\log x)^{(n+1)} &= \sum_{k=0}^{n} q_n(x)^{(k)}(\log x)^{(n+1-k)} \\ &= \sum_{k=0}^{n} q_n(x)^{(k)}(-1)^{n-k}(n-k)!\,\frac{1}{x^{n-k+1}}.\end{aligned}$$

Vamos a ver cada uno de los términos de esta suma.

- Para $k = 0$:

$$\frac{q_n(x)}{x^{n+1}}, \qquad \text{donde } q_n \text{ es el polinomio original, de grado } n.$$

- Para $k = 1$:

$$\frac{\hat{q}_{n-1}(x)}{x^n} = \frac{x\,\hat{q}_{n-1}(x)}{x^{n+1}}, \qquad \text{donde } \hat{q}_{n-1} \text{ es un polinomio de grado } n-1.$$

$\vdots$

- Para $k = n$:

$$\frac{\hat{q}_0}{x} = \frac{x^n\,\hat{q}_0}{x^{n+1}}, \qquad \text{donde } \hat{q}_0 \text{ es una constante, claramente, de grado } 0.$$

Por lo que

$$(p_n(x) + q_n(x)\log x)^{(n+1)} = \frac{R_n(x)}{x^{n+1}},$$

donde R_n es un polinomio de grado n a lo sumo. Entonces, el número máximo de ceros para esta derivada de orden $n+1$, será como mucho n. Lo que contradice el número mínimo de ceros que obteníamos por el teorema de Rolle. Como consecuencia de este absurdo, se tiene que el sistema planteado es un sistema de Chebyshev. $\Diamond$

Ejemplo 3.2. El conjunto de funciones

$$\{1, x, \ldots, x^n, \log x, x\log x, \ldots, x^n \log x, \log^2 x, x\log^2 x, \ldots, x^n \log^2 x\}$$

es un sistema de Chebyshev en el intervalo $[0, 1]$.

La idea es análoga a la expuesta en el ejemplo 3.1, añadiendo el detalle de conseguir la derivada n-ésima de $\log^2 x$, con la fórmula de Leibniz para las derivadas n-ésimas. $\Diamond$

Definición 3.5. Sea μ una medida positiva con soporte infinito en $[a, b]$. Sean u_j, $j = 1, \ldots, r$ funciones tales que $u_j > 0$, $x \in [a, b]$. Sea $(f_1, \ldots, f_r)$ el vector de funciones de Markov

$$f_j(z) = \int_a^b \frac{u_j(x)}{z - x}\, d\mu(x), \qquad j = 1, \ldots, r.$$

Decimos que $(f_1, \ldots, f_r)$ es un *sistema algebraico de Chebyshev (AT sistema)* para el multi-índice $\vec{n}$ si

$$\left\{u_1, xu_1, \ldots, x^{n_1-1}u_1, u_2, xu_2, \ldots, x^{n_2-1}u_2, \ldots, u_r, xu_r, \ldots, x^{n_r-1}u_r\right\}$$

es un sistema de Chebyshev en $[a, b]$.

Un vez definidos los AT sistemas, podemos enunciar los dos siguientes teoremas, cuyas demostraciones podemos encontrar en [22].

Teorema 3.2. *Supongamos que $\vec{n}$ es un multi-índice tal que $(f_1, \ldots, f_r)$ es un AT sistema en $[a, b]$ para todo multi-índice $\vec{m}$ tal que $m_j \leq n_j$ $(1 \leq j \leq r)$. Entonces $Q_{\vec{n}}$ tiene $|\vec{n}|$ ceros en (a, b) y $\vec{n}$ es un índice normal para el problema de aproximación de Hermite-Padé tipo II.*

Teorema 3.3. *Supongamos que $\vec{n}$ es un multi-índice tal que $(f_1, \ldots, f_r)$ es un AT sistema en $[a, b]$ para todo multi-índice $\vec{m}$ tal que $m_j \leq n_j$ $(1 \leq j \leq r)$. Entonces $\sum_{j=1}^{r} A_{\vec{n},j} u_j$ tiene $|\vec{n}| - 1$ ceros en (a, b) y $\vec{n}$ es un índice normal para el problema de aproximación de Hermite-Padé tipo I.*

Los AT sistemas van a ser de vital importancia en el siguiente capítulo para demostrar la irracionalidad de ciertos números. Para estos sistemas, se tienen similares condiciones de ortogonalidad que las que hemos deducido en la sección 3.2, pero simplemente teniendo en cuenta las distintas densidades para la misma medida y que trabajamos sobre un único intervalo de integración.

Para el tipo I, las condiciones de ortogonalidad (3.2), se convierten en

$$\int_a^b \left(\sum_{j=1}^{r} A_{\vec{n},j}(x) u_j(x) \right) x^k d\mu(x) = 0, \quad k = 0, 1, 2, \ldots, |\vec{n}| - 2, \tag{3.5}$$

el polinomio $B_{\vec{n}}$ viene dado por

$$B_{\vec{n}}(z) = \int_a^b \left(\sum_{j=1}^{r} \frac{A_{\vec{n},j}(z) - A_{\vec{n},j}(x)}{z - x} u_j(x) \right) dx, \tag{3.6}$$

y el resto en la aproximación tipo I es

$$\sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \int_a^b \sum_{j=1}^{r} \frac{A_{\vec{n},j}(x)\, u_j(x)}{z - x} d\mu(x). \tag{3.7}$$

Para el tipo II, las condiciones de ortogonalidad (3.4), se convierten en

$$\int_a^b Q_{\vec{n}}(x) x^k u_j(x) d\mu(x) = 0, \quad k = 0, 1, \ldots, n_j - 1, \quad j = 1, 2, \ldots, r, \tag{3.8}$$

los numeradores vienen dados por

$$P_{\vec{n},j}(z) = \int_a^b \frac{Q_{\vec{n}}(z) - Q_{\vec{n}}(x)}{z - x} u_j(x) d\mu(x), \tag{3.9}$$

y el resto

$$Q_{\vec{n}}(z) f_j(z) - P_{\vec{n},j}(z) = \int_a^b \frac{Q_{\vec{n}}(x)}{z - x} u_j(x) d\mu_j(x), \quad j = 1, 2, \ldots, r. \tag{3.10}$$

Capítulo 4

Aproximantes de Padé y Hermite-Padé: irracionalidad y trascendencia

4.1. Introducción

Sabiendo que cualquier número trascendente es forzosamente un número irracional, se podría prescindir de alguna demostración; como por ejemplo la de la irracionalidad del número e, puesto que probaremos su trascendencia. Sin embargo, la belleza matemática de estas demostraciones no nos permite esta omisión.

La mayoría de las demostraciones mostradas no son las originales, puesto que estas se van refinando a lo largo del tiempo, con objeto de ser más simples.

Incluso con todas las aportaciones a los campos de la irracionalidad y trascendencia de las mentes más brillantes de la aritmética, todavía resta mucho camino por recorrer en estos ámbitos.

La aproximación por racionales juega un papel fundamental como herramienta de apoyo para argumentar sobre la trascendencia e irracionalidad de ciertos números. Es ahí donde podemos apreciar la potencia de la aplicación de los aproximantes de Padé y de Hermite-Padé.

4.2. Irracionalidad

Definición 4.1. Un número es *irracional* cuando no es racional, es decir, cuando no puede escribirse de la forma p/q, con $p \in \mathbb{Z}$ y $q \in \mathbb{N}$.

Ejemplo 4.1 ($\sqrt{2}$ es irracional)**.** Supongamos que $\sqrt{2} = h/k$, fracción irreducible. Elevando al cuadrado obtenemos $h^2 = 2k^2$, luego h par. Tomemos $h = 2m$, con lo cual $4m^2 = 2k^2$. Simplificando, $k^2 = 2m^2$, luego k par. Así pues, hemos probado que tanto h como k son pares. Pero esto no es posible, pues partíamos de una fracción h/k irreducible. Luego $\sqrt{2}$ es irracional. $\diamondsuit$

En el capítulo 6 de [24], además de este último ejemplo, podemos encontrar más resultados para establecer la irracionalidad de algunos números. Enunciamos varios de ellos a continuación.

Resultado 4.1 (Corolario 6.4 de [24]). *Sea a un entero positivo. Entonces, $\sqrt[n]{a}$ es racional si y sólo si a es la potencia n-ésima de un entero positivo.*

Resultado 4.2 (Teorema 6.5 de [24]). *Sea $\theta = r\pi$ con $r \in \mathbb{Q}$. Entonces,* $\cos(\theta)$, $\operatorname{sen}(\theta)$ *y* $\operatorname{tg}(\theta)$ *son números irracionales, con las siguientes excepciones:*

1. $\cos(\theta) = 0$, ± 1 *o* $\pm 1/2$.
2. $\operatorname{sen}(\theta) = 0$, ± 1 *o* $\pm 1/2$.
3. $\operatorname{tg}(\theta) = 0$, ± 1 *o no está definida.*

Ejemplo 4.2 (Logaritmos con base racional). Si tenemos que $a, b \in \mathbb{Q}$, $a, b > 0$, ¿cómo podemos descubrir si $\log_b(a)$ es racional o irracional?

Tan solo es necesario usar el teorema fundamental de la aritmética: la descomposición canónica en factores primos de cada entero es única. Mostremos un sencillo ejemplo que nos ayudará a comprender el procedimiento:

Comprobemos si $\log_6(9)$ es racional o irracional. Supongamos que $\log_6(9) = p/q$, un racional. Esto implica que $6^{p/q} = 9$, o $6^p = 9^q$. Descomponiendo 6 y 9 en factores primos, se tendría que $2^p\, 3^p = 3^{2q}$. Por la unicidad de la descomposición en factores primos, esto sólo es posible con $p = 0$ y $p = 2q$, luego $p = q = 0$, y por tanto p/q no estaría definido. Como consecuencia, $\log_6(9)$ es irracional. $\diamondsuit$

Sin embargo, ¿qué pasa para números que no sean de la forma de los ejemplos anteriores? ¿Qué pasa por ejemplo para algunos números como e o π? ¿Y para el ejemplo concreto de $\log 2$, que la base del logaritmo no es racional? A continuación, resolveremos estas tres incógnitas planteadas, haciendo uso, en algunos casos, de los aproximantes de Padé y de Hermite-Padé.

4.2.1. Resultados previos auxiliares

En algunas demostraciones de irracionalidad de los números que plantearemos recurriremos con cierta frecuencia a dos resultados auxiliares. Es por ello que ahora los enunciamos y probamos.

Lema 4.1. *Sea x un número real. Si existen enteros p_n y q_n tales que*

1. $q_n x - p_n \neq 0$ *para todo* $n \in \mathbb{N}$,
2. $\lim_{n\to\infty}(q_n x - p_n) = 0$,

entonces x es irracional.

Demostración. Supongamos que x es racional, es decir, $x = p/q$ para algunos p, q enteros. Asumiendo la primera condición, para todo $n \in \mathbb{N}$:

$$q_n x - p_n = q_n \frac{p}{q} - p_n = \frac{1}{q}(pq_n - qp_n) \neq 0.$$

Como p, q, p_n, q_n son enteros, entonces se tiene que $pq_n - qp_n \neq 0$, así que $|pq_n - qp_n| \geq 1$. Pero entonces $|q_n x - p_n| \geq 1/|q|$, por lo que la segunda condición del lema no es posible. Este absurdo implica que x tiene que ser irracional. $\square$

Observación. El lema 4.1 implica que si existen infinitos pares (p, q) tales que

$$\left| x - \frac{p}{q} \right| = o\left(\frac{1}{q}\right),$$

entonces x es irracional. Es decir, se tiene que

$$\left| x - \frac{p}{q} \right| < \frac{C}{q^2},$$

por lo que tenemos una aproximación de orden 2. Sin embargo, si x es un número racional, obtenemos que

$$\left| x - \frac{p}{q} \right| < \frac{C}{q},$$

por lo que la aproximación es de orden 1. Esto es, ¡los números racionales aproximan mejor números irracionales que racionales!

Lema 4.2. *Sea d_n el mínimo común múltiplo de* 1, 2, 3,…, n. *Entonces*

$$\limsup_{n\to\infty} d_n^{1/n} \leq e.$$

Demostración. El mínimo común múltiplo de 1, 2, 3,…, n es de la forma

$$d_n = p_1^{s_1} p_2^{s_2} \dots p_{\pi(n)}^{s_{\pi(n)}},$$

donde p_1, p_2, …, $p_{\pi(n)}$ son números primos, $\pi(n)$ es el número de primos menores o iguales que n; y s_k es el máximo exponente de p_k en la factorización de cada número $m \leq n$.

En particular, se tiene que $p_k^{s_k} \leq r$, siendo r el natural entre 1 y n (ambos incluidos) para el cual encontramos el primo p_k elevado a la máxima potencia. Es claro entonces que

$$d_n = p_1^{s_1} p_2^{s_2} \dots p_{\pi(n)}^{s_{\pi(n)}} \leq n\, n \dots n = n^{\pi(n)}$$

y obtenemos que

$$d_n^{1/n} \leq n^{\pi(n)/n} = e^{\log n\, \pi(n)/n}.$$

Ahora, aplicando el teorema de los números primos

$$\lim_{n\to\infty} \frac{\pi(n)}{n/\log n} = 1,$$

se tiene que $\limsup_{n\to\infty} d_n^{1/n} \leq e$. $\square$

4.2.2. Irracionalidad del número e

Leonhard Euler, en 1737, fue el primero en probar la irracionalidad de e (también probó la de e^2). Él mismo descubrió la serie de potencias de e^x y también introdujo el símbolo e para la base de los logaritmos naturales. No obstante, su demostración de la irracionalidad de e no estaba basada en la serie, sino en el desarrollo de e mediante la siguiente fracción continua:

$$\frac{e-1}{2} = \cfrac{1}{1+\cfrac{1}{6+\cfrac{1}{10+\cfrac{1}{14+\cfrac{1}{18+\cdots}}}}}.$$

Fue Joseph Fourier, en 1815, el que probó la irracionalidad de e haciendo uso de la serie de potencias de la exponencial.

No obstante, nosotros mostraremos una prueba más sencilla y refinada que podemos encontrar en [24].

Teorema 4.1. *El número e es irracional.*

Demostración. Sabemos que $e = \sum_{k=0}^{\infty} 1/k!$ y supongamos que e es racional, es decir, e es de la forma $e = m/n \in \mathbb{Q}$. Es evidente que $n!(m/n - \sum_{k=0}^{n} 1/k!) \in \mathbb{Z}$. Además, $e - \sum_{k=0}^{n} 1/k!$ es positivo. Entonces,

$$\begin{aligned}
0 < n! \left(\frac{m}{n} - \sum_{k=0}^{n} \frac{1}{k!}\right) &= n! \left(e - \sum_{k=0}^{n} \frac{1}{k!}\right) \\
&= n! \left(\frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \frac{1}{(n+3)!} + \frac{1}{(n+4)!} + \dots\right) \\
&= \frac{n!}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \frac{1}{(n+2)(n+3)(n+4)} + \dots\right) \\
&\leq \frac{1}{n+1} \left(1 + \frac{1}{(n+2)} + \frac{1}{(n+2)^2} + \frac{1}{(n+2)^3} + \dots\right) \\
&= \frac{1}{n+1} \cdot \frac{1}{1 - 1/(n+2)} = \frac{n+2}{(n+1)^2} < 1,
\end{aligned}$$

y hemos llegado a que el entero $n!(m/n - \sum_{k=0}^{n} 1/k!)$ está entre 0 y 1, que es absurdo. El hecho de que no existen enteros entre 0 y 1 es mencionado con el nombre de *principio fundamental de la teoría de números.*

□

4.2.3. Irracionalidad de $\log 2$

En la demostración de la irracionalidad de $\log 2$, comenzaremos a ver la potencia y la utilidad de los aproximantes de Padé y Hermite-Padé para este tipo de pruebas. Vamos a construir sucesiones de enteros p_n y q_n que verifiquen el lema 4.1. Esta demostración está basada en [23].

Teorema 4.2. *El número* $\log 2$ *es irracional.*

Demostración. Consideramos la sucesión

$$a_n = \int_0^1 \frac{x^n(1-x)^n}{(1+x)^{n+1}}dx,\ n \in \mathbb{N}. \tag{4.1}$$

Primero vamos a ver que la sucesión es de la forma $a_n = A_n \log 2 - B_n$, con A_n y B_n números racionales. Daremos una expresión explícita para A_n y B_n.

Integramos n veces por partes la expresión de a_n, teniendo en cuenta que como $x = 0$ y $x = 1$ son raíces de multiplicidad n de la expresión $x^n(1-x)^n$, entonces

$$\frac{d^k}{dx^k}(x^n(1-x)^n)\Big|_{x=0} = \frac{d^k}{dx^k}(x^n(1-x)^n)\Big|_{x=1} = 0, \text{ para todo } k \in \mathbb{N}, \text{ con } 1 \leq k \leq n-1.$$

Aplicamos ahora integración por partes

$$a_n = \int_0^1 \frac{x^n(1-x)^n}{(1+x)^{n+1}}dx = \left\{ \begin{matrix} u_1 = x^n(1-x)^n & u_1' = \frac{d}{dx}(x^n(1-x)^n) \\ v_1' = (1+x)^{-n-1}dx & v_1 = -\frac{1}{n(1+x)^n} \end{matrix} \right\}$$

$$= \left(x^n(1-x)^n\left(-\frac{1}{n(1+x)^n}\right)\right)\Big|_{x=0}^{x=1} - \int_0^1 \frac{d}{dx}(x^n(1-x)^n)\left(-\frac{1}{n(1+x)^n}\right)dx$$

$$= \frac{1}{n}\int_0^1 \frac{d}{dx}(x^n(1-x)^n)\frac{1}{(1+x)^n}dx = \left\{ \begin{matrix} u_2 = \frac{d}{dx}(x^n(1-x)^n) & u_2' = \frac{d^2}{dx^2}(x^n(1-x)^n) \\ v_2' = (1+x)^{-n}dx & v_2 = \frac{1}{(-n+1)(1+x)^{n-1}} \end{matrix} \right\}$$

$$= \frac{1}{n}\left[\frac{d}{dx}(x^n(1-x)^n)\,\frac{(1+x)^{1-n}}{-n+1}\Big|_{x=0}^{x=1} - \int_0^1 \frac{d^2}{dx^2}(x^n(1-x)^n)\frac{1}{(-n+1)(1+x)^{n-1}}dx\right]$$

$$= \frac{1}{n(n-1)}\int_0^1 \frac{d^2}{dx^2}(x^n(1-x)^n)\frac{1}{(1+x)^{n-1}}dx = \ldots \text{(n veces partes en total)} \ldots$$

$$= \frac{1}{n!}\int_0^1 \frac{d^n}{dx^n}(x^n(1-x)^n)\frac{1}{1+x}dx.$$

En resumen,

$$a_n = \frac{1}{n!}\int_0^1 \frac{d^n}{dx^n}(x^n(1-x)^n)\frac{1}{1+x}dx. \tag{4.2}$$

La integral a_n contiene la expresión

$$P_n(x) = \frac{1}{n!}\frac{d^n}{dx^n}(x^n(1-x)^n).$$

Es obvio que P_n es un polinomio de grado n. Siendo más precisos, es la fórmula de Rodrigues del *polinomio de Legendre* de grado n en el intervalo $[0,1]$. Los polinomios de Legendre son los polinomios ortogonales respecto a la medida de Lebesgue. Para más información, ver [19]. Usando el binomio de Newton y derivando coeficiente a coeficiente, podemos escribirlo de la forma:

$$P_n(x) = \frac{1}{n!}\frac{d^n}{dx^n}\sum_{k=0}^{n}\binom{n}{k}(-1)^k x^{n+k} = \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}(-x)^k. \tag{4.3}$$

Ahora, podemos reescribir la ecuación (4.2) como

$$a_n = \int_0^1 \frac{P_n(x)}{1+x}dx = \int_0^1 \frac{P_n(x)-P_n(-1)}{1+x}dx + P_n(-1)\int_0^1 \frac{1}{1+x}dx. \tag{4.4}$$

Notemos que

$$\int_0^1 \frac{1}{1+x} = \big(\log(1+x)\big|_0^1 = \log 2$$

y que

$$A_n = P_n(-1) = \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}$$

es claramente un entero. Tomando ahora

$$B_n = \int_0^1 \frac{P_n(-1)-P_n(x)}{1+x}dx,$$

observamos que $P_n(-1)-P_n(x)$ es un polinomio de grado n que tiene una raíz en $x=-1$. Por ello, $(1+x)$ divide al polinomio y se tiene que lo de dentro de la integral de B_n es también un polinomio, pero, de grado $n-1$. Usando la expresión (4.3),

$$\frac{P_n(-1)-P_n(x)}{1+x} = \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\frac{1-(-x)^k}{1+x} = \sum_{k=0}^{n}\binom{n}{k}\binom{n+k}{k}\frac{1-(-x)^k}{1-(-x)};$$

y si pensamos en una suma finita de los términos de una progresión finita de razón $r=-x$

$$\sum_{j=0}^{k-1}(-x)^j = \frac{1-(-x)^k}{1-(-x)},$$

entonces

$$\frac{P_n(-1)-P_n(x)}{1+x} = \sum_{k=0}^{n}\sum_{j=0}^{k-1}\binom{n}{k}\binom{n+k}{k}(-x)^j.$$

E integrando esta última expresión entre 0 y 1, obtenemos que

$$B_n = \sum_{k=0}^{n}\sum_{j=0}^{k-1}\binom{n}{k}\binom{n+k}{k}\frac{(-1)^j}{j+1}.$$

Es trivial ahora que B_n es racional. Como los términos que hay en el denominador son enteros entre 1 y n, si multiplicamos por el mínimo común múltiplo de estos

$$d_n = mcm(1, 2, 3, \dots n),$$

entonces el resultado será un entero.

Tomemos ahora

$$q_n = d_n A_n, \qquad p_n = d_n B_n,$$

que ya son enteros y veamos que se verifican las dos condiciones del lema 4.1 para $x = \log 2$.

Para la primera, como el integrando de a_n en la expresión (4.1) es

$$\frac{x^n(1-x)^n}{(1+x)^{n+1}},$$

es positivo para $x \in (0, 1)$, entonces $a_n > 0$ y en consecuencia $d_n a_n > 0$.

Comprobemos ahora la segunda condición. Antes, notemos que la función

$$g(x) = \frac{x(1-x)}{1+x}$$

es una función continua en el intervalo $[0, 1]$, que es un compacto. Por el teorema de Weierstrass alcanza sus extremos absolutos. En particular, alcanza su máximo para $x_1 = -1 + \sqrt{2}$. Y como $g(0) = g(1) = 0$, entonces

$$g(x) = \frac{x(1-x)}{1+x} \leq g(x_1) = g\left(-1 + \sqrt{2}\right) = \left(\sqrt{2} - 1\right)^2, \quad \text{para todo } x \in [0, 1].$$

Y como elevar a una potencia natural mantiene la monotonía de la desigualdad, conseguimos la siguiente cota

$$\frac{x^n(1-x)^n}{(1+x)^n} \leq \left(\sqrt{2} - 1\right)^{2n}. \tag{4.5}$$

Usando (4.5) y recuperando la fórmula inicial para a_n,

$$a_n = \int_0^1 \frac{x^n(1-x)^n}{(1+x)^n} \frac{1}{1+x} dx \leq \int_0^1 \max_{0 \leq x \leq 1} \left\{ \frac{x^n(1-x)^n}{(1+x)^n} \right\} \frac{1}{1+x} dx = \left(\sqrt{2} - 1\right)^{2n} \log 2.$$

Para probar que $d_n a_n = q_n \log 2 - p_n$ converge a cero, es suficiente probar que

$$\limsup_{n \to \infty} (a_n d_n)^{1/n} < 1.$$

Por una parte tenemos que

$$\limsup_{n \to \infty} a_n^{1/n} \leq \limsup_{n \to \infty} \left(\left(\sqrt{2} - 1\right)^{2n} \log 2 \right)^{1/n} = \left(\sqrt{2} - 1\right)^2. \tag{4.6}$$

Y por la otra, gracias al lema 4.2

$$\limsup_{n\to\infty} d_n^{1/n} \leq e. \tag{4.7}$$

Juntando (4.6) y (4.7), podemos afirmar que

$$\limsup_{n\to\infty} (a_n d_n)^{1/n} \leq e\left(\sqrt{2}-1\right)^2 = 0{,}466383428\ldots < 1,$$

por lo que también se verifica la segunda condición del lema 4.1 y se concluye la prueba de que $\log 2$ es un número irracional.

□

La clave de esta demostración, está camuflada. Pero, se trata del aproximante de Padé de la función de Markov

$$f(z) = \int_0^1 \frac{1}{z-x}dx$$

en el punto $z = -1$, ya que $f(-1) = -\log 2$.

Los polinomios P_n de Legendre son los denominadores en los aproximantes de Padé; es decir, los polinomios ortogonales. Y cumplen que

$$\int_0^1 P_n(x)x^k = 0, \quad \text{para todo } k \in \mathbb{N}, \text{ con } 1 \leq k \leq n-1.$$

4.2.4. Irracionalidad de π

Arquímedes, en el siglo III a. C. ya conjeturó que π era irracional. Pero, hubo que esperar alrededor de 2000 años para que, Lambert, en 1761 lo probara por medio de la fracción continua:

$$\operatorname{tg}(x) = \cfrac{x}{1-\cfrac{x^2}{3-\cfrac{x^2}{5-\cdots}}}.$$

Ivan Niven, en 1947, lo probó usando tan solo cálculo elemental.

Probando que π^2 es irracional, también tendríamos que lo es π. Es así como se explica a continuación, haciendo uso de los aproximantes de Hermite-Padé, como podemos ver en [20].

Teorema 4.3. *El número $\pi^2/6$ es irracional.*

Demostración. Consideramos las dos funciones de Markov

$$f_1(z) = \int_0^1 \frac{1}{z-x}dx, \qquad f_2(z) = -\int_0^1 \frac{\log x}{z-x}dx,$$

cuyos momentos para esas medidas son

$$\int_0^1 x^k dx = \frac{1}{k+1}, \qquad -\int_0^1 x^k \log x \, dx = \frac{1}{(k+1)^2}, \qquad k = 0, 1, 2, \ldots$$

por lo que tenemos las siguientes series en el infinito para $f_1(z)$ y $f_2(z)$

$$f_1(z) = \sum_{k=0}^{\infty} \frac{1}{k+1} \frac{1}{z^{k+1}}, \qquad f_2(z) = \sum_{k=0}^{\infty} \frac{1}{(k+1)^2} \frac{1}{z^{k+1}}.$$

Aproximaremos estas dos funciones de forma simultánea buscando polinomios A_n, B_n y C_n de grado a lo más n tales que

$$A_n(z) - B_n(z) \log z = \mathcal{O}((1-z)^{n+1}), \qquad z \to 1, \tag{4.8}$$

$$A_n(z) f_1(z) + B_n(z) f_2(z) - C_n(z) = \mathcal{O}(\frac{1}{z^{n+1}}), \qquad z \to \infty. \tag{4.9}$$

Notemos que (4.9) es una aproximación de Hermite-Padé tipo I para las dos funciones (f_1, f_2), que además forman un AT sistema. Ya que son funciones de Markov y las correspondientes densidades para las medidas $u_1(x) = 1$ y $u_2(x) = -\log x$, forman un sistema de Chebyshev (se trata del ejemplo 3.1 para $n = 0$ y $-\log x$ en vez de $\log x$, que generan el mismo espacio de funciones).

Observando (4.8), vemos que se trata de un problema de un aproximante de Padé para la función logaritmo. Las dos ecuaciones juntas, como requieren de polinomios comunes A_n y B_n, forman un problema de aproximación de Hermite-Padé tipo II. Se trata de una combinación de un problema de Hermite-Padé tipo I y uno de tipo II.

Vamos a considerar la función

$$F_n(x) = A_n(x) - B_n(x) \log x.$$

Sea $\mathcal{S} = \{1, x, x^2, \ldots, x^n, \log x, x \log x, x^2 \log x, \ldots, x^n \log x\}$.

La función F_n pertenece al espacio generado por $\mathcal{S}$, que sabemos por el ejemplo 3.1 que es un sistema de Chebyshev en el intervalo $[0, 1]$. Por lo tanto, F_n tiene como mucho $2n+1$ ceros en $[0, 1]$. La condición (4.8) nos dice que $F_n(x)$ tiene un cero de multiplicidad al menos $n + 1$ en el punto $x = 1$, por lo que $F_n^{(k)}(1) = 0$ para $k = 0, 1, \ldots, n$.

Si nos fijamos en la integral

$$\frac{(-1)^{n+1}}{n!} \int_x^1 (t-x)^n F_n^{(n+1)}(t) dt,$$

teniendo en cuenta la multiplicidad del cero $x = 1$ para $F_n(x)$ y aplicando n veces partes a esta integral, llegamos a la expresión de $F_n(x)$ (similar a como se ha procedido en la demostración del teorema 4.2),

$$\frac{(-1)^{n+1}}{n!}\int_x^1 (t-x)^n F_n^{(n+1)}(t)dt = \left\{\begin{matrix} u_1=(t-x)^n & u_1'=n(t-x)^{n-1}dt \\ v_1'=F_n^{(n+1)}(t)dt & v_1=F_n^{(n)}(t) \end{matrix}\right\}$$

$$= \frac{(-1)^{n+1}}{n!}\left[(t-x)^n F_n(t)\Big|_{t=x}^{t=1} - \int_x^1 n(t-x)^{n-1}F_n^{(n)}(t)dt\right]$$

$$= \frac{(-1)^n}{(n-1)!}\int_x^1 (t-x)^{n-1}F_n^{(n)}(t)dt = \ldots (\text{n veces partes en total}) \ldots = F_n(x).$$

Por lo tanto

$$F_n(x) = \frac{(-1)^{n+1}}{n!}\int_x^1 (t-x)^n F_n^{(n+1)}(t)dt.$$

Para $F_n^{(n+1)}$, el polinomio A_n desaparece (es de grado menor o igual que n). Y, por la regla de Leibniz

$$F_n^{(n+1)}(x) = \sum_{k=0}^{n}\binom{n+1}{k}B_n^{(k)}(x)\frac{(-1)^{n-k}(n-k)!}{x^{n+1-k}}.$$

Por lo que podemos expresar $F_n^{(n+1)}(x) = \frac{1}{x}\widetilde{D}_n(\frac{1}{x})$, con $\widetilde{D}_n$ un polinomio de grado como mucho n. Para facilitar las cuentas, si hacemos $D_n(\frac{1}{x}) = \frac{(-1)^{n+1}}{n!}\widetilde{D}_n(\frac{1}{x})$, D_n sigue siendo un polinomio de grado a lo sumo n y resulta la siguiente solución general para (4.8)

$$F_n(x) = \int_x^1 (t-x)^n D_n(1/t)\frac{1}{t}dt.$$

Y si ahora denotamos por $E_n(x) = x^n D_n(1/x)$, la anterior expresión se convierte en

$$F_n(x) = \int_x^1 (1-x/t)^n E_n(t)\frac{1}{t}dt. \tag{4.10}$$

El siguiente paso es determinar el polinomio E_n, usando (4.9). Recordemos que se trata de un problema de tipo I para un AT sistema, por lo que tenemos las relaciones de ortogonalidad (3.5)

$$\int_0^1 (A_n(x) - B_n(x)\log x)x^k dx = 0, \qquad k = 0, 1, \ldots, n-1.$$

Lo que significa que F_n es ortogonal a todos los polinomios de grado a lo sumo $n-1$. Usando la expresión de (4.10), obtenemos

$$\int_0^1 x^k \int_x^1 (1-x/t)^n E_n(t)\frac{1}{t}dt\,dx = 0, \qquad k = 0, 1, \ldots, n-1.$$

Por el teorema de Fubini, podemos intercambiar el orden de integración y resulta

$$\int_0^1 E_n(t) \int_0^t (1-x/t)^n x^k dx \, \frac{1}{t} \, dt = 0, \qquad k = 0, 1, \ldots, n-1.$$

El cambio de variable $x = ty$ convierte la anterior expresión en

$$\int_0^1 E_n(t) t^k \, dt \int_0^1 (1-y)^n y^k \, dy = 0, \qquad k = 0, 1, \ldots, n-1.$$

Y como $(1-y)^n y^k > 0$ para todo $k = 0, 1, \ldots, n-1$, con $y \in (0,1)$, entonces

$$\int_0^1 E_n(t) t^k \, dt = 0, \qquad k = 0, 1, \ldots, n-1.$$

Esto último muestra que E_n es ortogonal a todos los polinomios de grado a lo sumo $n-1$ en el intervalo $[0,1]$ respecto a la medida de Lebesgue; por lo que E_n es el polinomio de Legendre ya descrito anteriormente en (4.3)

$$E_n(x) = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} (-1)^k x^k. \tag{4.11}$$

Una vez que sabemos cómo es E_n, la ecuación (4.10) queda así

$$F_n(x) = \sum_{k=0}^{n} \binom{n}{k} \binom{n+k}{k} (-1)^k \int_x^1 (1-x/t)^n t^{k-1} dt.$$

Aplicando el teorema del binomio a $(1-x/t)^n$, conseguimos

$$F_n(x) = \sum_{k=0}^{n} \sum_{j=0}^{n} \binom{n}{k} \binom{n+k}{k} \binom{n}{j} (-1)^{k+j} x^j \int_x^1 t^{k-j-1} dt. \tag{4.12}$$

Es sencillo comprobar que

$$\int_x^1 t^{k-j-1} dt = \begin{cases} -\log x & \text{si } k = j, \\ \frac{1-x^{k-j}}{k-j} & \text{si } k \neq j, \end{cases}$$

por lo que, como $B_n(x)$ es el polinomio que multiplica a $\log x$, teniendo en cuenta (4.12), será de la forma

$$B_n(x) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k} x^k, \tag{4.13}$$

y A_n será de la forma

$$A_n(x) = \sum_{k=0}^{n} \sum_{\substack{j=0 \\ j \neq k}}^{n} \binom{n}{k} \binom{n+k}{k} \binom{n}{j} (-1)^{k+j} \frac{x^j - x^k}{k-j}. \tag{4.14}$$

Nuestro interés reside en

$$f_2(1) = -\int_0^1 \frac{\log x}{1-x}dx = \sum_{k=0}^{\infty} = \zeta(2) = \frac{\pi^2}{6}.$$

Queremos evaluar para $z = 1$ en (4.9), pero $f_1(z)$ diverge para $z = 1$. Sin embargo, $A_n(1) = 0$, lo que significa que podemos expresar $A_n(z) = (z-1)\hat{A}_{n-1}(z)$, con $\hat{A}_{n-1}(z)$ un polinomio de grado una unidad menor que el grado de A_n; y esto implica que

$$\lim_{z\to 1} A_n(z)f_1(z) = \hat{A}_{n-1}(1)\lim_{z\to 1}(z-1)\log\left(\frac{z}{z-1}\right) = \hat{A}_{n-1}(1)\lim_{z\to 1}(1-z)\log\left(1-\frac{1}{z}\right) = 0,$$

por lo que la primera función de Markov no causa problemas. Y el resto en la aproximación (4.9) de tipo I, como en (3.7), teniendo en cuenta que las correspondientes densidades para las medidas son $u_1(x) = 1$ y $u_2(x) = -\log x$, es

$$A_n(z)f_1(z) + B_n(z)f_2(z) - C_n(z) = \int_0^1 \frac{A_n(x) - B_n(x)\log x}{z-x}dx, \tag{4.15}$$

y el polinomio C_n es

$$C_n(z) = \int_0^1 \left(\frac{A_n(z) - A_n(x)}{z-x} - \frac{B_n(z) - B_n(x)}{z-x}\log x\right) dx.$$

Evaluando en la ecuación (4.15) para $z = 1$ conseguimos

$$B_n(1)\frac{\pi^2}{6} - C_n(1) = \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x}dx.$$

Observando la expresión (4.13) de B_n, resulta obvio que

$$B_n(1) = \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}$$

es un entero. Para calcular $C_n(1)$, calculamos previamente dos cantidades. La primera

$$\begin{aligned}\int_0^1 \frac{B_n(1) - B_n(x)}{1-x}\log x\, dx &= \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}\int_0^1 \frac{1-x^k}{1-x}\log x\, dx \\ &= -\sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}\sum_{j=0}^{k-1}\frac{1}{(j+1)^2},\end{aligned}$$

y la segunda

$$\int_0^1 \frac{A_n(x)}{1-x}dx = \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n}\binom{n}{k}\binom{n+k}{k}\binom{n}{j}(-1)^{k+j}\frac{1}{k-j}\int_0^1 \frac{x^j - x^k}{1-x}dx.$$

Tenemos que

$$\int_0^1 \frac{x^j - x^k}{1-x} dx = H_k - H_j,$$

donde $H_n = \sum_{j=1}^n 1/j$ son los números armónicos y por lo tanto

$$C_n(1) = \sum_{k=0}^n \binom{n}{k}^2 \binom{n+k}{k} \sum_{j=0}^{k-1} \frac{1}{(j+1)^2} - \sum_{k=0}^n \sum_{\substack{j=0 \\ j\neq k}}^n \binom{n}{k}\binom{n+k}{k}\binom{n}{j}(-1)^{k+j}\frac{H_k - H_j}{k-j}.$$

Podemos ver que $C_n(1)$ no es necesariamente un entero, pero si multiplicamos $C_n(1)$ por d_n^2, donde d_n es el mínimo común múltiplo de $\{1, 2, \ldots, n\}$, entonces $d_n^2 C_n(1)$ es un entero. Por lo tanto, si elegimos $q_n = d_n^2 B_n(1)$ y $p_n = d_n^2 C_n(1)$, entonces

$$q_n \frac{\pi^2}{6} - p_n = d_n^2 \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} dx.$$

Vamos a ver que se verifican las dos condiciones del lema 4.1. Antes, vamos a trabajar un poco la parte derecha de la anterior expresión usando la expresión (4.10) y la fórmula de Rodrigues para los polinomios de Legendre

$$E_n(x) = \frac{1}{n!}\left(\frac{d}{dx}\right)^n x^n(1-x)^n. \tag{4.16}$$

Usando (4.10),

$$\int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} dx = \int_0^1 \frac{F_n(x)}{1-x} dx = \int_0^1 \frac{1}{1-x}\int_x^1 (1-x/t)^n E_n(t)\frac{1}{t} dt\, dx.$$

Ahora, como se trata de una función integrable, podemos intercambiar el orden de integración (teorema de Fubini)

$$\int_0^1 \frac{1}{1-x}\int_x^1 (1-x/t)^n E_n(t)\frac{1}{t}\, dt\, dx = \int_0^1 E_n(t)\int_0^t (1-x/t)^n \frac{1}{1-x}\, dx\, \frac{1}{t}\, dt.$$

Realizando el cambio de variable $x = ty$

$$\int_0^1 E_n(t)\int_0^t (1-x/t)^n \frac{1}{1-x}\, dx\, \frac{1}{t}\, dt = \int_0^1 E_n(t)\int_0^1 (1-y)^n \frac{1}{1-yt}\, dy\, dt.$$

Por comodidad, denotemos por I la anterior expresión; es decir,

$$I = \int_0^1 E_n(t)\int_0^1 (1-y)^n \frac{1}{1-yt}\, dy\, dt.$$

Aplicando el teorema de Fubini y denotando por $I_0(y)$ a la siguiente expresión

$$I_0(y) = \int_0^1 (1-y)^n \frac{1}{1-yt}\, dt,$$

obtenemos que

$$I = \int_0^1 (1-y)^n \, I_0(y) \, dy. \tag{4.17}$$

Ahora nos centramos en $I_0(y)$. Primero sustituimos $E_n(x)$ con la fórmula (4.16); después, como en anteriores demostraciones en este trabajo, aplicamos n veces la fórmula de integración por partes y nos queda

$$I_0(y) = (-1)^n \int_0^1 \frac{y^n}{(1-yt)^{n+1}} t^n (1-t)^n \, dt.$$

Por lo tanto

$$I = \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} dx = (-1)^n \int_0^1 \frac{y^n t^n (1-y)^n (1-t)^n}{(1-yt)^{n+1}} \, dt.$$

Para comprobar la primera condición del lema 4.1, como el integrando de la parte derecha de la anterior expresión es positivo para $y \in (0,1)$ y $t \in (0,1)$, entonces

$$q_n \frac{\pi^2}{6} - p_n = d_n^2 \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} dx \neq 0.$$

Comprobemos ahora la segunda condición. Antes, notemos que la función

$$f(y,t) = \frac{yt(1-y)(1-t)}{1-yt}$$

es una función continua para $(y,t) \in [0,1] \times [0,1]$, que es un compacto. Por el teorema de Weierstrass, $f(y,t)$ alcanza sus extremos absolutos.

Para $y = 0, 1$ o para $t = 0, 1$ la función $f(y,t)$ no alcanza su máximo en $[0,1] \times [0,1]$, ya que vale 0. Derivando respecto de una de las dos variables (es indiferente sobre cuál, por la simetría de la función $f(y,t)$), buscamos los máximos en el interior del cuadrado $[0,1] \times [0,1]$ y obtenemos $y = \frac{2t-1}{t^2}$.

Ahora, el problema se reduce a buscar el máximo de la función $g(t)$ tal que

$$g(t) = f\left(\frac{2t-1}{t^2}, t\right), \qquad \text{con } t \in [0,1].$$

Este máximo se alcanza para $t = \frac{-1+\sqrt{5}}{2}$, entonces

$$f(y,t) = \frac{yt(1-y)(1-t)}{1-yt} \leq g\left(\frac{\sqrt{5}-1}{2}\right) = \left(\frac{\sqrt{5}-1}{2}\right)^5, \text{ para todo } (y,t) \in [0,1] \times [0,1].$$

Como elevar a una potencia natural mantiene la monotonía de la desigualdad,

$$\left| \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} dx \right| \leq \left(\frac{\sqrt{5}-1}{2}\right)^{5n} \int_0^1 \int_0^1 \frac{1}{1-yt} dy \, dt.$$

Finalmente, vemos que se cumple la segunda condición del lema 4.1, ya que

$$\lim_{n\to\infty}\left|q_n\frac{\pi^2}{6}-p_n\right|^{1/n}\leq e^2\left(\frac{\sqrt{5}-1}{2}\right)^5=0{,}66627\ldots<1.$$

Por lo que

$$\lim_{n\to\infty}q_n\frac{\pi^2}{6}-p_n=0,$$

y por el lema 4.1, se concluye la prueba de que $\pi^2/6$ es un número irracional.

□

Corolario 4.1. *El número π es irracional.*

Demostración. Puesto que $\pi^2/6$ es irracional, es claro que π^2 también lo es.

Supongamos que $\pi\in\mathbb{Q}$ y por lo tanto $\pi=p/q$. Entonces $\pi^2=p^2/q^2$ es racional, cuando hemos probado en el teorema 4.3 que es irracional. □

4.2.5. Irracionalidad de $\zeta(3)$

En 1977, R. Apéry en [1], anunció la prueba de la irracionalidad de $\zeta(3)$. Más tarde, Beukers lo demostró en términos de los aproximantes de Hermite-Padé, [5], como se procede a continuación y podemos encontrar en [21].

Teorema 4.4. *El número $\zeta(3)$ es irracional.*

Demostración. Consideramos las tres funciones de Markov

$$f_1(z)=\int_0^1\frac{1}{z-x}dx,\qquad f_2(z)=-\int_0^1\frac{\log x}{z-x}dx,\qquad f_3(z)=\frac{1}{2}\int_0^1\frac{\log^2 x}{z-x}dx,$$

cuyos momentos para esas medidas son

$$\int_0^1 x^k dx=\frac{1}{k+1},\qquad -\int_0^1 x^k\log x\,dx=\frac{1}{(k+1)^2},\qquad \frac{1}{2}\int_0^1 x^k\log^2 x\,dx=\frac{1}{(k+1)^3}.$$

Por lo que, en particular, $f_3(1)=\zeta(3)$, que es el número en el que estamos interesados. Podemos observar que (f_1,f_2,f_3) es otra vez un AT sistema. Ya que son funciones de Markov y las correspondientes densidades para las medidas son $u_1(x)=1$, $u_2(x)=-\log x$ y $u_3(x)=\frac{1}{2}\log^2 x$, que forman un sistema de Chebyshev (se trata de un caso equivalente al ejemplo 3.2).

Planteamos el problema de aproximación

$$A_n(z)=\mathcal{O}(z-1),\qquad z\to 1,\tag{4.18}$$

$$A_n(z)f_1(z)+B_n(z)f_2(z)-C_n(z)=\mathcal{O}\left(\frac{1}{z^{n+1}}\right),\qquad z\to\infty,\tag{4.19}$$

$$A_n(z)f_2(z)+2B_n(z)f_3(z)-D_n(z)=\mathcal{O}\left(\frac{1}{z^{n+1}}\right),\qquad z\to\infty,\tag{4.20}$$

donde A_n, B_n, C_n y D_n son polinomios de grado a lo sumo n. Notemos que (4.18) solo significa que $A_n(1) = 0$; es decir, A_n tiene una raíz en 1. Esto es necesario para que $A_n(z)f_2(z)$ sea cero cuando $z = 1$.

Buscamos la aproximación racional de $\zeta(3)$, tomando $z = 1$ en la ecuación (4.20), la cual resulta ser $2B_n(1)\zeta(3) - D_n(1)$ en la parte de la izquierda.

Observemos que las expresiones (4.19) y (4.20), por separado, son ambas un problema de aproximación tipo I para los sistemas de funciones (f_1, f_2) y $(f_2, 2f_3)$, respectivamente. Y, en conjunto, conforman un problema de aproximación tipo II con común denominador el par (A_n, B_n).

Las condiciones de ortogonalidad de tipo I, como en (3.5), resultan ser

$$\int_0^1 (A_n(x) - B_n(x)\log x)\, x^k\, dx = 0 \qquad k = 0, 1, 2, \ldots, n-1 \tag{4.21}$$

$$\int_0^1 (A_n(x) - B_n(x)\log x)\, x^k \log x\, dx = 0 \qquad k = 0, 1, 2, \ldots, n-1. \tag{4.22}$$

Realizamos un pequeño inciso, para que la suposición que hagamos después en (4.23) no se vea tan alejada. Sea $P_n(x)$ un polinomio de grado a lo sumo n

$$P_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0.$$

Si ahora integramos entre x y 1 el siguiente producto de Cauchy para series finitas multiplicado por $1/t$

$$\int_x^1 P_n\left(\frac{x}{t}\right) P_n(t)\frac{1}{t}\, dt = \int_x^1 \sum_{k=0}^{n}\sum_{j=0}^{n} a_k\, a_j\, x^k\, t^{j-k}\, \frac{1}{t}\, dt = \sum_{k=0}^{n}\sum_{j=0}^{n} a_k\, a_j\, x^k \int_x^1 t^{j-k-1}\, dt$$

$$= \left(\sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n} a_k\, a_j\, x^k\, \frac{t^{j-k}}{j-k} + \sum_{k=0}^{n} a_k^2\, x^k\, \log t\right|_{t=x}^{t=1}$$

$$= \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n} a_k\, a_j\, \frac{x^k}{j-k} + \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n} a_k\, a_j\, \frac{x^j}{k-j} - \sum_{k=0}^{n} a_k^2\, x^k\, \log x.$$

Denotamos

$$\hat{A}_n(x) = \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n} a_k\, a_j\, \frac{x^k}{j-k} + \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n} a_k\, a_j\, \frac{x^j}{k-j}$$

y

$$\hat{B}_n(x) = \sum_{k=0}^{n} a_k^2\, x^k.$$

Entonces, es claro que $\hat{A}_n(x)$ y $\hat{B}_n(x)$ son dos polinomios de grado a lo sumo n. Así que podemos expresar

$$\int_x^1 P_n\left(\frac{x}{t}\right) P_n(t)\frac{1}{t}\, dt = \hat{A}_n(x) - \hat{B}_n(x)\log x.$$

Cerrando el inciso y recordando la demostración de la irracionalidad de $\pi^2/6 = \zeta(2)$, no es muy precipitado suponer que podemos expresar $A_n(x) - B_n(x)\log x$ como

$$A_n(x) - B_n(x)\log x = \int_x^1 E_n\left(\frac{x}{t}\right) E_n(t)\frac{1}{t}\, dt, \tag{4.23}$$

donde $E_n(x)$ es un polinomio de grado a lo sumo n.

Llamamos por comodidad $F_n(x)$ a lo siguiente

$$F_n(x) = A_n(x) - B_n(x)\log x = \int_x^1 E_n\left(\frac{x}{t}\right) E_n(t)\frac{1}{t}\, dt.$$

$F_n(x)$ es una función en el espacio generado por

$$\mathcal{S} = \{1, x, x^2, \ldots, x^n, \log x, x\log x, x^2\log x, \ldots, x^n\log x\},$$

tal que es cero cuando $x = 1$, así que se cumple la ecuación (4.18).

Ahora, necesitamos saber cómo tiene que ser el polinomio $E_n(x)$ para cumplir las relaciones de ortogonalidad (4.21) y (4.22). Resulta que si $E_n(x)$ es el polinomio de Legendre (4.3), podemos comprobar que se satisfacen dichas relaciones de ortogonalidad.

Comprobamos la primera, (4.21)

$$\int_0^1 F_n(x)\, x^k\, dx = \int_0^1 \int_x^1 E_n\left(\frac{x}{t}\right) E_n(t)\frac{1}{t}\, dt\, x^k\, dx.$$

Por el teorema de Fubini

$$\int_0^1 \int_0^t E_n\left(\frac{x}{t}\right) E_n(t)\, x^k\, dx\, \frac{1}{t}\, dt = \int_0^1 E_n(t)\, \frac{1}{t} \int_0^t E_n\left(\frac{x}{t}\right)\, x^k\, dx\, dt$$

y haciendo después el cambio de variable $x = ty$

$$\int_0^1 E_n(t)\, \frac{1}{t} \int_0^1 E_n(y)\, (ty)^k\, t\, dy\, dt = \int_0^1 E_n(t)\, t^k\, dt \int_0^1 E_n(y)\, y^k\, dy.$$

Y como el polinomio de Legendre $E_n(x)$ de grado n es ortogonal a todos los polinomios de grado a lo sumo $n-1$ en el intervalo $[0,1]$ con la medida de Lebesgue, entonces

$$\int_0^1 F_n(x)\, x^k\, dx = \int_0^1 E_n(t)\, t^k\, dt \int_0^1 E_n(y)\, y^k\, dy = 0, \quad k = 0, 1, \ldots, n-1.$$

Comprobamos la segunda, (4.22)

$$\int_0^1 F_n(x)\, x^k\, \log x\, dx = \int_0^1 \int_x^1 E_n\left(\frac{x}{t}\right) E_n(t)\frac{1}{t}\, dt\, x^k\, \log x\, dx.$$

Por el teorema de Fubini,

$$\int_0^1 \int_0^t E_n\left(\frac{x}{t}\right) E_n(t)\, x^k\, \log x\, dx\, \frac{1}{t}\, dt = \int_0^1 E_n(t)\, \frac{1}{t} \int_0^t E_n\left(\frac{x}{t}\right)\, x^k\, \log x\, dx\, dt$$

y con el cambio de variable $x = ty$, se tiene que

$$\begin{aligned}
&\int_0^1 E_n(t)\,\frac{1}{t}\int_0^1 E_n(y)\,(ty)^k\,\log(ty)\,t\,dy\,dt\\
&= \int_0^1 E_n(t)\,\frac{1}{t}\int_0^1 E_n(y)\,(ty)^k\,(\log t + \log y)\,t\,dy\,dt\\
&= \int_0^1 E_n(t)\,\frac{1}{t}\int_0^1 E_n(y)\,(ty)^k\,(\log t + \log y)\,t\,dy\,dt\\
&= \int_0^1 E_n(t)\,t^k\int_0^1 E_n(y)\,y^k\,\log t\,dy\,dt + \int_0^1 E_n(t)\,t^k\int_0^1 E_n(y)\,y^k\,\log y\,dy\,dt\\
&= 2\int_0^1 E_n(t)\,t^k\,\log t\int_0^1 E_n(y)\,y^k\,dy\,.
\end{aligned}$$

Y como el polinomio de Legendre $E_n(x)$ de grado n es ortogonal a todos los polinomios de grado a lo sumo $n-1$ en el intervalo $[0,1]$ con la medida de Lebesgue, entonces

$$\int_0^1 F_n(x)\,x^k\,\log x\,dx = 2\int_0^1 E_n(t)\,t^k\,\log t\,dt\int_0^1 E_n(y)\,y^k\,dy = 0,$$

para todo $k = 0, 1, \ldots, n-1$.

Usando la expresión explícita (4.3) del polinomio de Legendre, la función F_n puede ser escrita como

$$F_n(x) = \sum_{k=0}^{n}\sum_{j=0}^{n}\binom{n}{k}\binom{n}{j}\binom{n+k}{k}\binom{n+j}{j}(-1)^{k+j}x^k\int_x^1 t^{j-k-1}\,dt,$$

por lo que podemos identificar B_n cuando $k = j$

$$B_n(x) = \sum_{k=0}^{n}\binom{n}{k}^2\binom{n+k}{k}^2 x^k, \tag{4.24}$$

y A_n se consigue de los términos restantes

$$A_n(x) = \sum_{k=0}^{n}\sum_{\substack{j=0\\ j\neq k}}^{n}\binom{n}{k}\binom{n}{j}\binom{n+k}{k}\binom{n+j}{j}(-1)^{k+j}\,\frac{x^k - x^j}{j-k}. \tag{4.25}$$

Es claro que $B_n(1)$ es un entero. Ahora, el resto en la aproximación (4.20) de tipo I, como en (3.7), teniendo en cuenta que las correspondientes densidades para las medidas son $u_1(x) = 1$, y $u_3(x) = \frac{1}{2}\log^2 x$, es

$$A_n(z)f_2(z) + 2B_n(z)f_3(z) - D_n(z) = \int_0^1 \frac{A_n(x)(-\log x) + B_n(x)\log^2 x}{z-x}dx \tag{4.26}$$

y el polinomio D_n es

$$D_n(z) = -\int_0^1 \frac{A_n(z) - A_n(x)}{z-x}\log x\,dx + \int_0^1 \frac{B_n(z) - B_n(x)}{z-x}\log^2 x\,dx.$$

Evaluando en la ecuación (4.26) para $z = 1$ conseguimos

$$2B_n(1)\zeta(3) - D_n(1) = -\int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} \log x \, dx.$$

Para calcular $D_n(1)$, calculamos previamente dos cantidades. La primera

$$\begin{aligned}\int_0^1 \frac{B_n(1) - B_n(x)}{1-x} \log^2 x \, dx &= \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \int_0^1 \frac{1-x^k}{1-x} \log^2 x \, dx \\ &= \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{j=0}^{k-1} \frac{2}{(j+1)^3},\end{aligned}$$

y la segunda

$$\int_0^1 \frac{A_n(x)}{1-x} \log x \, dx = \sum_{k=0}^{n} \sum_{\substack{j=0 \\ j\neq k}}^{n} \binom{n}{k}\binom{n}{j}\binom{n+k}{k}\binom{n+j}{j} \frac{(-1)^{k+j}}{j-k} \int_0^1 \frac{x^k - x^j}{1-x} \log x \, dx.$$

Tenemos que

$$\int_0^1 \frac{x^k - x^j}{1-x} dx = H_k^{(2)} - H_j^{(2)},$$

donde $H_n^{(2)} = \sum_{j=1}^{n} 1/j^2$ y por lo tanto

$$\begin{aligned}D_n(1) =& \sum_{k=0}^{n} \binom{n}{k}^2 \binom{n+k}{k}^2 \sum_{j=0}^{k-1} \frac{2}{(j+1)^3} \\ &+ \sum_{k=0}^{n} \sum_{\substack{j=0 \\ j\neq k}}^{n} \binom{n}{k}\binom{n}{j}\binom{n+k}{k}\binom{n+j}{j} (-1)^{k+j} \frac{H_k^{(2)} - H_j^{(2)}}{j-k}.\end{aligned}$$

Podemos ver que $D_n(1)$ no es necesariamente un entero, pero si multiplicamos $D_n(1)$ por d_n^3, donde d_n es el mínimo común múltiplo de $\{1, 2, \ldots, n\}$, entonces $d_n^3 D_n(1)$ es un entero. Por lo tanto, si elegimos $q_n = 2d_n^3 B_n(1)$ y $p_n = d_n^3 D_n(1)$, entonces

$$q_n\zeta(3) - p_n = -d_n^3 \int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x} \log x \, dx.$$

Vamos a ver que se verifican las dos condiciones del lema 4.1. Para ello, supondremos ciertas las dos siguientes ecuaciones, por no alargar tanto esta demostración. La primera, se trata de un problema de cálculo integral, con algún que otro cambio de variable. La segunda, se trata de un sencillo (pero pesado) problema de optimización que no requiere gran dificultad.

$$\int_0^1 \frac{A_n(s) - B_n(s)\log s}{1-s} \log s \, ds = -\int_0^1\int_0^1\int_0^1 \frac{(xyz(1-x)(1-y)(1-z))^n}{(1-(1-xy)z)^{n+1}} dx\, dy\, dz,$$

$$\max_{0\leq x,y,z\leq 1}\left\{\frac{(xyz(1-x)(1-y)(1-z))^n}{(1-(1-xy)z)^{n+1}}\right\} = \left(\sqrt{2}-1\right)^4.$$

Asumiendo esto, se tiene que

$$q_n\zeta(3) - p_n = d_n^3 \int_0^1\int_0^1\int_0^1 \frac{(xyz(1-x)(1-y)(1-z))^n}{(1-(1-xy)z)^{n+1}} dx\,dy\,dz \neq 0,$$

ya que el integrando anterior es siempre positivo, para $(x, y, z) \in (0, 1) \times (0, 1) \times (0, 1)$.

Ahora, como elevar a una potencia natural mantiene la monotonía de una desigualdad

$$\left|\int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x}\log x\,dx\right| \leq (\sqrt{2}-1)^{4n}\int_0^1\int_0^1\int_0^1 \frac{1}{1-(1-xy)z}dx\,dy\,dz.$$

Finalmente, vemos que se cumple la segunda condición del lema 4.1, ya que

$$\begin{aligned}\limsup_{n\to\infty} |q_n\zeta(3) - p_n|^{1/n} &= \limsup_{n\to\infty} d_n^{3/n}\left|\int_0^1 \frac{A_n(x) - B_n(x)\log x}{1-x}\log x\,dx\right|^{1/n}\\ &\leq e^3\,(\sqrt{2}-1)^4 = 0{,}591263\ldots < 1.\end{aligned}$$

Por lo que

$$\lim_{n\to\infty} q_n\zeta(3) - p_n = 0,$$

y por el lema 4.1, se concluye la prueba de que $\zeta(3)$ es un número irracional. □

4.3. Trascendencia

Definición 4.2. Se denominan *números algebraicos* a los números que son raíz de un polinomio no nulo con coeficientes enteros.

Es destacable notar que estos coeficientes del polinomio pueden ser racionales, ya que cualquier polinomio con coeficientes racionales se puede convertir en uno con coeficientes enteros que tiene las mismas raíces, por medio del mínimo común múltiplo de todos los denominadores.

Definición 4.3. Un número es *trascendente* cuando no es algebraico.

Fue Euler el primero que especuló que los números trascendentes exsitían. Y ya en 1844, Joseph Liouville logró probar la existencia de estos. Liouville construía números trascendentes concretos, como por ejemplo $\sum_{n=1}^{\infty} 1/2^{n!}$.

La gran sorpresa fue en 1874, cuando Georg Cantor dio un argumento que probaba que no solo existen los números trascendentes, sino que la mayoría de números son trascendentes.

Recordando las dos siguientes definiciones, entenderemos el argumento usado por Cantor.

Definición 4.4. Un conjunto es *numerable* cuando podemos probar que existe una biyección entre $\mathbb{N}$ y los elementos del conjunto.

Definición 4.5. Un conjunto es *no numerable* cuando podemos probar que no existe ninguna biyección entre $\mathbb{N}$ y los elementos del conjunto.

De esta forma, entre 1873 y 1874, Cantor probó que los conjuntos de los números racionales y de los algebraicos son numerables y que $\mathbb{R}$ es no numerable. Lo que implica que los números trascendentes forman un conjunto no numerable. Es decir, demostró que "casi todos" los números son trascendentes.

4.3.1. Trascendencia del número e

Charles Hermite, en 1873, mostró que e es trascendente. El método de Hermite sirvió para posteriores pruebas de trascendencia de otros números. Se apoyaba en las propiedades analíticas de la función exponencial y resultaba bastante complejo.

Antes de demostrar la trascendencia de e con la ayuda de aproximantes de Hermite-Padé, basándonos en [20], enunciamos y probamos el siguiente lema que usaremos.

Lema 4.3. *Sea x un número real. Supongamos que para cada entero positivo $m \in \mathbb{N}$ y para cualesquiera enteros $a_0, a_1, \dots, a_m \in \mathbb{Z}$ podemos encontrar $m+1$ sucesiones $p_{0,n}, p_{1,n}, \dots, p_{m,n}$ $(n \in \mathbb{N})$ tales que*

1. $\sum_{k=0}^{m} a_k p_{k,n} \neq 0$ *para todo* $n \in \mathbb{N}$,
2. $\lim_{n\to\infty}(p_{0,n}x^k - p_{k,n}) = 0$ *para* $k = 1, 2, \dots, m$.

Entonces x es trascendente.

Demostración. Sea x un número real verificando las hipótesis del lema y supongamos que x es algebraico.

Como x es algebraico, entonces $\exists\, m \in \mathbb{N}$ y $\exists\, a_0, a_1, \dots, a_m \in \mathbb{Z}$ tales que

$$\sum_{k=0}^{m} a_k x^k = 0. \tag{4.27}$$

Sean $p_{0,n}, p_{1,n}, \dots, p_{m,n}$ las $m+1$ sucesiones que verifican las hipótesis del lema para los $m+1$ enteros $a_0, a_1, \dots, a_m$. Multiplicando ahora la ecuación (4.27) por cada uno de los términos de la sucesión $p_{0,n}$

$$p_{0,n}\sum_{k=0}^{m} a_k x^k = 0, \text{ para todo } n \in \mathbb{N}.$$

Ahora, restando y sumando una misma cantidad, nos queda

$$\sum_{k=0}^{m} a_k(p_{0,n}x^k - p_{k,n}) + \sum_{k=0}^{m} a_k p_{k,n} = 0, \text{ para todo } n \in \mathbb{N}.$$

Por la primera condición del lema, $\sum_{k=0}^{m} a_k p_{k,n} \neq 0$ y es un entero, por lo que

$$\left|\sum_{k=0}^{m} a_k(p_{0,n}x^k - p_{k,n})\right| \geq 1, \text{ para todo } n \in \mathbb{N}. \tag{4.28}$$

Por otra parte, si sumamos los m límites de la segunda condición del lema multiplicados por a_k

$$\lim_{n\to\infty} \sum_{k=1}^{m} a_k(p_{0,n}x^k - p_{k,n}) = 0,$$

podemos escribir

$$\lim_{n\to\infty} \left(a_0\, p_{0,n} - a_0\, p_{0,n} + \sum_{k=1}^{m} a_k(p_{0,n}x^k - p_{k,n}) \right) = \lim_{n\to\infty} \sum_{k=0}^{m} a_k(p_{0,n}x^k - p_{k,n}) = 0.$$

Tomando valor absoluto en la anterior expresión y tomando límite cuando $n \to \infty$ en (4.28) se llega a un absurdo. Entonces x es trascendente.

$\square$

Teorema 4.5. *El número e es trascendente.*

Demostración. Vamos a buscar los aproximantes de Hermite-Padé tipo II para las r funciones $(e^{\lambda_1 x}, e^{\lambda_2 x}, \dots, e^{\lambda_r x})$ alrededor de $x = 0$. Estos aproximantes se conocen explícitamente cuando $n = (n_1, n_2, \dots, n_r)$ y $m = (m_1, m_2, \dots, m_r)$ son tales que $m_j + n_j = N + |n|$ para $1 \leq j \leq r$, donde $|n| = n_1 + n_2 + \ldots + n_r$.
Buscamos polinomios $Q_n(z)$ y $P_{m;j}(z)$ tales que para cada $j = 1, 2, \dots, r$

i) $\operatorname{grad}(Q_n) \leq |n|$, $\operatorname{grad}(P_{m;j}) \leq m_j = |n| + N - n_j$;

ii) $Q_n(z)e^{\lambda_j z} - P_{m;j}(z) = \mathcal{O}(z^{n_j+m_j+1}), \qquad z \to 0.$

Sea $T(x)$ el polinomio

$$T(x) = x^N(x-\lambda_1)^{n_1}(x-\lambda_2)^{n_2}\dots(x-\lambda_r)^{n_r}. \tag{4.29}$$

Entonces, para ciertas constantes $a_0, a_1, \dots, a_{|n|}$, el polinomio $T(x)$ se puede expresar como

$$T(x) = x^N \left(a_0 + a_1 x + \ldots + a_{|n|}x^{|n|}\right) = a_0 x^N + a_1 x^{N+1} + \ldots + a_{|n|}x^{N+|n|}. \tag{4.30}$$

Aplicando la fórmula de integración por partes k veces a la siguiente integral, para $k \in \mathbb{N}$, llegamos a

$$\int_0^\infty x^k e^{-zx}dx = \frac{k!}{z^{k+1}}. \tag{4.31}$$

Teniendo en cuenta las ecuaciones (4.30) y (4.31), resulta

$$\int_0^\infty T(x)e^{-zx}dx = a_0\frac{N!}{z^{N+1}} + a_1\frac{(N+1)!}{z^{N+2}} + \cdots + a_{|n|}\frac{(N+|n|)!}{z^{N+|n|+1}}. \tag{4.32}$$

Luego
$$z^{N+|n|+1}\int_0^{\infty} T(x)\; e^{-zx}dx$$
es un polinomio de grado $|n|$. Denotamos ahora
$$Q_n(z) = z^{N+|n|+1}\int_0^{\infty} T(x)\; e^{-zx}dx.$$

Observemos por otra parte que para $1 \leq j \leq r$,
$$T(x+\lambda_j) = (x+\lambda_j)^N(x+\lambda_j-\lambda_1)^{n_1}(x+\lambda_j-\lambda_2)^{n_2}\cdots x^{n_j}\cdots(x+\lambda_j-\lambda_r)^{n_r}.$$
Así que para ciertos $b_0, b_1, \ldots, b_{N+|n|-n_j}$, podemos expresar el polinomio $T(x+\lambda_j)$ como
$$T(x+\lambda_j) = x^{n_j}\left(b_0 + b_1 x + \cdots + b_{N+|n|-n_j}x^{N+|n|-n_j}\right). \tag{4.33}$$
Juntando las ecuaciones (4.31) y (4.33), obtenemos
$$\int_0^{\infty} T(x+\lambda_j)\; e^{-zx}dx = b_0\frac{(n_j)!}{z^{n_j+1}} + b_1\frac{(n_j+1)!}{z^{n_j+2}} + \cdots + b_{N+|n|-n_j}\frac{(N+|n|)!}{z^{N+|n|+1}}. \tag{4.34}$$
Así que denotando
$$P_{m;j}(z) = z^{N+|n|+1}\int_0^{\infty} T(x+\lambda_j)\; e^{-zx}dx,$$
se tiene que $P_{m;j}(z)$ es un polinomio de grado $N+|n|+1-n_j-1 = m_j$. Veamos ahora que los polinomios así construidos son los aproximantes de Hermite-Padé tipo II para las r funciones $(e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_r x})$. Operamos un poco los polinomios construidos
$$Q_n(z)e^{\lambda_j z} - P_{m;j}(z) = z^{N+|n|+1}\; e^{\lambda_j z}\int_0^{\infty} T(x)\; e^{-zx}dx - z^{N+|n|+1}\int_0^{\infty} T(x+\lambda_j)\; e^{-zx}dx$$
$$= z^{N+|n|+1}\; e^{\lambda_j z}\left(\int_0^{\infty} T(x)\; e^{-zx}dx - \int_0^{\infty} T(x+\lambda_j)\; e^{-z(x+\lambda_j)}dx\right).$$
Y haciendo ahora el cambio de variable $x+\lambda_j = t$,
$$Q_n(z)e^{\lambda_j z} - P_{m;j}(z) = z^{N+|n|+1}\; e^{\lambda_j z}\left(\int_0^{\infty} T(x)\; e^{-zx}dx - \int_{\lambda_j}^{\infty} T(t)\; e^{-zt}dt\right)$$
$$= z^{N+|n|+1}\; e^{\lambda_j z}\int_0^{\lambda_j} T(x)\; e^{-zx}dx.$$
Por lo que tenemos que
$$Q_n(z)e^{\lambda_j z} - P_{m;j}(z) = z^{N+|n|+1}\; e^{\lambda_j z}\int_0^{\lambda_j} T(x)\; e^{-zx}dx. \tag{4.35}$$

Ahora, por la ecuación (4.30)

$$\int_0^{\lambda_j} T(x)\, e^{-zx} dx = \int_0^{\lambda_j} x^N \sum_{k=0}^{|n|} a_k x^k e^{-zx} dx = \sum_{k=0}^{|n|} a_k \int_0^{\lambda_j} x^{N+k} e^{-zx} dx. \tag{4.36}$$

Por un argumento de inducción y la fórmula de integración por partes, se puede probar que para todo $n \in \mathbb{N}$, se cumple que

$$\int x^n e^{-zx} dx = -\sum_{l=0}^{n} \frac{n!}{l!}\, (zx)^l\, \frac{e^{-zx}}{z^{n+1}}. \tag{4.37}$$

Juntando ahora (4.35), (4.36) y (4.37) y operando

$$\begin{aligned}
Q_n(z)e^{\lambda_j z} - P_{m;j}(z) &= z^{N+|n|+1}\, e^{\lambda_j z} \sum_{k=0}^{|n|} a_k \int_0^{\lambda_j} x^{N+k} e^{-zx} dx \\
&= z^{N+|n|+1}\, e^{\lambda_j z} \sum_{k=0}^{|n|} \left[a_k \left(- \sum_{l=0}^{N+k} \frac{(N+k)!}{l!}\, (zx)^l\, \frac{e^{-zx}}{z^{N+k+1}} \Big|_{x=0}^{x=\lambda_j} \right]\right. \\
&= \sum_{k=0}^{|n|} \left[a_k\, (N+k)!\, z^{|n|-k} \left(e^{\lambda_j z} - \sum_{l=0}^{N+k} \frac{(z\lambda_j)^l}{l!} \right) \right] \\
&= \sum_{k=0}^{|n|} \left[a_k\, (N+k)!\, z^{|n|-k} \left(\sum_{l=0}^{\infty} \frac{(z\lambda_j)^l}{l!} - \sum_{l=0}^{N+k} \frac{(z\lambda_j)^l}{l!} \right) \right] \\
&= \sum_{k=0}^{|n|} \left[a_k\, (N+k)!\, z^{|n|-k} \sum_{l=N+k+1}^{\infty} \frac{(z\lambda_j)^l}{l!} \right] \\
&= \hat{A}_1\, z^{N+|n|+1} + \hat{A}_2\, z^{N+|n|+2} + \hat{A}_3\, z^{N+|n|+3} + \ldots = \mathcal{O}(z^{m_j+n_j+1}), \qquad z \to 0,
\end{aligned}$$

donde $\hat{A}_1, \hat{A}_2, \hat{A}_3, \ldots$ son ciertas constantes. Así que hemos probado que los polinomios $Q_n(z)$ y $P_{m;j}(z)$ son los aproximantes de Hermite-Padé tipo II cerca del origen de las r funciones $(e^{\lambda_1 x}, e^{\lambda_2 x}, \ldots, e^{\lambda_r x})$.

Para probar la trascendencia de e, tomamos $\lambda_j = j$, $z = 1$ y un primo $p > r$ que no sea divisible por a_0 (de la expresión (4.32)). Además elegimos $N = p - 1$ y $n_j = p$, para $j = 1, 2, \ldots, r$. Comprobemos ahora las siguientes dos afirmaciones.

- Veamos que $p_{0,n} = Q_n(1)/(p-1)!$ es un entero no divisible por p, con la expresión (4.32), sustituyendo las nuevas condiciones

$$Q_n(z) = a_0\, (p-1)!\, z^{rp} + a_1\, p!\, z^{rp-1} + \cdots + a_{rp}\, (rp+p-1)!$$

$$p_{0,n} = \frac{Q_n(1)}{(p-1)!} = a_0 + a_1\, p + a_2\, p(p+1) + \ldots + a_{rp}\, p(p+1)\cdots(rp+p-1) \;\in\; \mathbb{Z}.$$

  Y como p no divide a a_0, entonces $p_{0,n}$ no es divisible por p.

- Notemos que $p_{j,n} = P_{m;j}(1)/(p-1)!$ $(1 \le j \le r)$ es un entero divisible por p, con la expresión (4.34), sustituyendo las nuevas modificaciones

$$P_{m;j}(z) = b_0\, p!\, z^{rp-1} + b_1\,(p+1)!\, z^{rp-2} + \cdots + b_{rp-1}\,(rp+p-1)!$$

$$p_{j,n} = \frac{P_{m;j}(1)}{(p-1)!} = b_0\, p + b_1\, p(p+1) + b_2\, p(p+1)(p+2) + \ldots + b_{rp-1}\, p(p-1)\cdots(rp+p-1).$$

Y claramente, $p_{j,n}$ es un entero divisible por p.

Como consecuencia de las anteriores afirmaciones,

$$\sum_{j=0}^{r} a_j p_{j,n} \text{ no es divisible por } p, \text{ por lo que } \sum_{j=0}^{r} a_j p_{j,n} \neq 0,$$

y se verifica la primera condición del lema 4.3. Además, por la ecuación (4.35) evaluada en $z = 1$, sustituyendo las nuevas condiciones y dividiendo entre $(p-1)!$

$$p_{0,n} e^j - p_{j,n} = \frac{e^j}{(p-1)!} \int_0^j T(x)\, e^{-x}\, dx. \tag{4.38}$$

Ahora, si $0 \le x \le r$, para todo $1 \le j \le r$, se cumple que $-j \le x - j \le r - j \le r$. Por lo que también se cumple que $|x| \le r$ y que $|x - j| \le r$. Recuperando la ecuación (4.29) y sustituyendo las nuevas condiciones, podemos afirmar que para todo $0 \le x \le r$,

$$|T(x)| = |x^{p-1}(x-1)^p(x-e)^p \cdots (x-r)^p| \le r^{p-1} r^p \cdots r^p = r^{(r+1)p-1}. \tag{4.39}$$

Como consecuencia de esta cota para $|T(x)|$ y de la expresión (4.38),

$$|p_{0,n} e^j - p_{j,n}| = \left| \frac{e^j}{(p-1)!} \int_0^j T(x)\, e^{-x}\, dx \right| \le \frac{e^j}{(p-1)!} \int_0^j |T(x)\, e^{-x}|\, dx$$

$$\le e^j\, \frac{r^{(r+1)p-1}}{(p-1)!} \int_0^j e^{-x}\, dx \le e^j\, \frac{r^{(r+1)p-1}}{(p-1)!},$$

donde en la última desigualdad hemos usado que

$$\int_0^j e^{-j} dx = \big(-e^{-x}\, \big|_0^j = -e^{-j} + 1 = 1 - e^{-j} \le 1.$$

Es decir, tenemos que

$$|p_{0,n} e^j - p_{j,n}| \le e^j\, \frac{r^{(r+1)p-1}}{(p-1)!}. \tag{4.40}$$

Como hay infinitos primos, se cumple que si $n \to \infty$, entonces $p \to \infty$. Además, por órdenes de infinitos, para $1 < s$, se cumple que $s^p \ll p!$, cuando $p \to \infty$. Así que tomando límites en (4.40)

$$\lim_{n\to\infty} |p_{0,n} e^j - p_{j,n}| \le \lim_{p\to\infty} e^j\, \frac{r^{(r+1)p-1}}{(p-1)!} = 0,$$

vemos que se cumple la segunda condición del lema (4.3) y por lo tanto e es un número trascendente.

□

4.3.2. Trascendencia de π

La trascendencia de π fue demostrada por Ferdinand von Lindemann en 1882. La trascendencia de π implica que la cuadratura del círculo no es posible. Un argumento analítico de teoría de números lograba resolver un problema geométrico de la época griega.

La demostración de la trascendencia no hace uso de los aproximantes de Padé ni Hermite-Padé, por lo que no presentamos su prueba. Sin embargo, en un trabajo que involucre la trascendencia, no podíamos dejar de mencionar el siguiente resultado.

Teorema 4.6 (Trascendencia de π)**.** *El número π es trascendente.*

4.4. Otros resultados y problemas abiertos

El campo de la irracionalidad y trascendencia es bastante desconocido. Sin embargo, se conocen más resultados que los demostrados anteriormente en este trabajo. Enunciamos a continuación los más conocidos.

Resultado 4.3 (Euler)**.** *Los valores de la función zeta de Riemann evaluada en números pares están definidos y son trascendentes (por lo que también son irracionales):*

$$\zeta(2n) = \frac{2^{2n-1}B_n}{(2n)!}\pi^{2n},$$

donde B_n son los números de Bernoulli (los cuales son racionales)

$$B_n = \begin{cases} 1, & \text{si } n = 0, \\ B_n \text{ tal que } \sum_{k=0}^{n-1}\binom{n}{k}B_k = 0, & \text{si } n \geq 1. \end{cases}$$

Demostración. La trascendencia de π implica que $\zeta(2n)$ es irracional (y trascendente) para cada $n \geq 1$.

$\square$

Resultado 4.4 (Séptimo problema de Hilbert)**.** *Si $\alpha \neq 0, 1$ es un número algebraico en $\mathbb{C}$ y β es un número algebraico no racional, entonces*

$$\alpha^\beta \text{ es trascendente.}$$

Este resultado fue demostrado por A. Gelfond y por Th. Schneider de forma independiente y como consecuencias inmediatas de este resultado, citamos algunas afirmaciones:

- $2^{\sqrt{2}}$, que es el número conocido como la constante de Gelfond-Schneider, es trascendente.
- $\sqrt{2}^{\sqrt{2}}$ es trascendente.
- $e^\pi = i^{-2i}$ es trascendente.

Resultado 4.5 (Ball y Rivoal, [4]). *Infinitos valores de la forma* $\zeta(2n+1)$ *con* $n \in \mathbb{N}$ *son irracionales.*

Resultado 4.6 (Zudilin, [25]). *Al menos uno de los números* $\zeta(5), \zeta(7), \zeta(9)$ *y* $\zeta(11)$ *es irracional.*

Además, quedan muchos problemas que siguen siendo una incógnita. Mostramos tres de ellos muy populares.

- La función zeta de Riemann en los impares es un auténtico misterio. Solo sabemos que $\zeta(3)$ es irracional, pero, ni siquiera se sabemos si es o no trascendente.
- La famosa constante de Euler-Mascheroni

$$\gamma = \lim_{n\to\infty} \left(\sum_{k=1}^{n} \frac{1}{k} - \log n \right),$$

  no se sabe si es irracional y tampoco se sabe si es trascendente.
- Otra constante para la que la irracionalidad sigue siendo un problema abierto es la constante de Catalan, la cual viene dada por

$$G = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)^2}.$$

  Además, se desconoce también si es o no trascendente.

Capítulo 5

Conclusiones

Haberme enfrentado a este Trabajo de Fin de Grado ha supuesto un gran desafío. No solo a nivel matemático, en lo que se refiere a comprender la teoría y lo que se está estudiando, sino también a nivel personal, ya que ha supuesto afrontar un emocionante y nuevo reto.

A lo largo de este bonito y apasionante camino, con esta memoria como resultado, he tenido el placer (o la desdicha) de haber tenido dos incansables compañeros de viaje: los muchos momentos de desesperación y las dudas puntuales que te quitan el sueño. Estos fieles amigos me han acompañado en todo el trayecto. Pero, es cierto que también han estado las ganas de aprender, de dar lo máximo de mí mismo y, sobre todo, mi predilección por las matemáticas. Es esto último precisamente, lo que hizo que hace cuatro años decidiera ampliar mis conocimientos matemáticos y hoy pueda ratificar que estoy orgulloso de esa elección.

El tema de este trabajo se engloba dentro de la teoría de aproximación. Durante el transcurso del mismo, he podido descubrir cómo los polinomios de Taylor no se comportan tan bien fuera de su disco de convergencia, para ciertas funciones con singularidades. Es por eso que, disponer de los aproximantes de Padé y de Hermite-Padé para poder extender las aproximaciones más allá del disco de convergencia, es una gran ventaja.

Además, en el contenido de esta memoria, no solo se introducen los aproximantes, sino que, como colofón en el cuarto capítulo, se emplean para demostrar la irracionalidad y trascendencia de tan célebres números como e o π. No obstante, es importante mencionar que la mayoría de las demostraciones vistas, son "ad hoc", se construyen expresamente para esa prueba en concreto. Por desgracia, no se pueden generalizar.

Aunque también podemos ver la otra cara de la moneda. Incluso con todas las aportaciones a los campos de la irracionalidad y trascendencia de las mentes más brillantes de la aritmética como por ejemplo Charles Hermite, que dan pie a este escrito, todavía resta mucho camino por recorrer. La irracionalidad y la trascendencia son dos campos en los que queda mucho por explorar. Siguiendo un poco la línea de D. Hilbert dando su opinión en sus clases, es probable que vivamos para ver algún resultado nuevo en estos dos campos de la aritmética.

Por último, quiero comentar que aprender de forma autodidacta no es nada sencillo.

Resultando más difícil por estar inmersos en una pandemia y por haber tenido que estar confinados en nuestras casas. Sin embargo, he tenido la suerte de contar con una tutora que me ha guiado y apoyado desde el primer momento, me ha aclarado todas las dudas que he tenido y ha conseguido transmitirme sus conocimientos sobre este campo de la teoría de aproximación. Gracias, Judit, sin tu ayuda no hubiera sido posible.

Bibliografía

[1] R. Apéry, Irrationalité de $\zeta(2)$ et $\zeta(3)$, *Astérisque* **61** (1979), 11–13.

[2] A. Aptekarev, Asymptotics of polynomials of simultaneous orthogonality in the Angelescu case, *Math. USSR-Sb.* **64** (1989), 57–84.

[3] G.A. Baker y P.R. Graves-Morris, "Padé Approximants", Cambridge University Press, Cambridge, 1996.

[4] K. Ball y T. Rivoal, Irrationalité d'une infinité de valeurs de la fonction zèta aux entiers impairs, *Invent. Math.* **146** (2001), no. 1, 193–207.

[5] F. Beukers, A note on the irrationality of $\zeta(2)$ and $\zeta(3)$, *Bull. London Math. Soc.* **11** (1979), 268–272.

[6] B. de la Calle Ysern y G. López-Lagomasino, Strong asymptotics of orthogonal polynomials with varying measures and Hermite-Padé approximants, *J. Comput. Math.* **99** (1998), 91–103.

[7] B. de la Calle Ysern y G. López-Lagomasino, Weak convergence of varying measures and Hermite-Padé orthogonal approximants, *Constr. Approx.* **15** (1999), 553–575.

[8] J.I. Extremiana, L.J. Hernández y M.T. Rivas, La divina razón de la belleza, *Sigma* **27** (2005), 145–178.

[9] U. Fidalgo, S. Medina-Peralta, J. Mínguez-Ceniceros, Mixed type multiple orthogonal polynomials: perfectness and interlacing properties of zeros, *Linear Algebra Appl.* **438** (2013), no. 3, 1229–1239.

[10] A.A. Gonchar, Poles of rows of the Padé table and meromorphic continuation of functions, *Math. USSR-Sb.* **43** (1982), 527–546.

[11] A.A. Gonchar, E.A. Rakhamanov y V.N. Sorokin, Hermite-Padé approximants for systems of Markov type functions, *Sb. Math.* **188** (1997), 671–696.

[12] C. Hermite, Sur la fonction exponentielle, *C. R. Acad. Sci. Paris Sér. I Math.* **77** (1873), 18–24. 74–79, 226–233, 285–293; publicado también en sus "Oeuvres", Tome III, Gauthier-Villars, Paris, 1912, 150–181.

[13] J.M. López Fernández, El suicidio de Hipaso de Metaponto o la tensión entre la intuición matemática y la matemática formal. 10.13140/RG.2.1.1667.5688. (2019).

[14] G. López Lagomasino, Constructive Theory of Approximation. An Intoduction to Padé Approximation. Coimbra Lecture Notes on Orthogonal Polynomials, Nova Science Pub. (2008), 101–139.

[15] A.A. Markov, Deux démonstrationes de la convergence de certaines fractions continues, *Acta Math.* **19** (1895), 93–104.

[16] R. de Montessus de Ballore, Sur le fractions continues algébriques, *Bull. Soc. Math. de France* **30** (1902), 28–36.

[17] S.P. Suetin, On poles of the mth row of a Padé table, *Math. USSR-Sb.* **48** (1984), 493–497.

[18] S.P. Suetin, On an inverse probleme of the mth row of the Padé table, *Math. USSR-Sb.* **52** (1985), 231–244.

[19] G. Szegő, "Orthogonal polynomials", 4th ed. Providence (RI): American Mathematical Society; (1975). American Mathematical Society Colloquium Publications.

[20] W. van Assche, Approximation theory analytic number theory, en "Special Functions and Differential Equations" (K. Srinivasa Rao et al., eds.), Allied Publishers, New Delhi (1998), 336–355.

[21] W. Van Assche, Multiple orthogonal polynomials, irrationality and transcedence, *Contemporary Mathematics* **236** (1999), 325–342.

[22] W. Van Assche, Padé and Hermite-Padé approximation and orthogonality. *Surv. Approx. Theory* **2** (2006), 61–91.

[23] W. van Assche, Analytic number theory and approximation, Coimbra Lecture Notes on Orthogonal Polynomials, Nova Science Pub. (2007), 211–243.

[24] J.L. Varona, "Recorridos por la teoría de números" (2ª edición). Electolibris y Real Sociedad Matemática Española (2019).

[25] Z. Z. Zudilin, One of the numbers $\zeta(5)$, $\zeta(7)$, $\zeta(9)$, $\zeta(11)$ is irrational, *Uspekhi Mat. Nauk* **56** (2001), no. 4 (340) 149–150; *Russian Math. Surveys* **56** (2001), no. 4 774–776.