Sección de Matemáticas
Universidad de La Laguna

Teresa Padilla Miranda

Aproximación Padé y Hermite-Padé

Padé & Hermite-Padé Approximation

Trabajo Fin de Grado
Grado en Matemáticas
La Laguna, Junio de 2022

Dirigido por
Carlos Javier Díaz Mendoza

Carlos Javier Díaz Mendoza
Departamento de Análisis
Matemático
Universidad de La Laguna
38200 La Laguna, Tenerife

Agradecimientos

Me gustaría aprovechar la ocasión para darle las gracias a Carlos por su gran implicación, por la pasión con la que me ha transmitido sus conocimientos, proporcionándome una visión que me ha facilitado entender este TFG y disfrutarlo, y por sacar el máximo de mí, además de despertar en mí una motivación e inquietudes desconocidas por esta rama de las matemáticas. ¡Mil gracias!

Por otra parte, evidentemente me gustaría agradecer a mi familia el apoyo dado, en especial a mi tía por apoyarme en cualquier decisión y por haber sido una “madre” para mí y mis hermanos durante nuestra infancia y adolescencia. Y, sin duda, a quien más le debo agredecer es a mi abuela por acogerme desde pequeña y también cuidarme como si fuese una hija más.

Asimismo, agradecer a toda esa gente maravillosa con la que me he cruzado a lo largo de esta carrera, fundamentalmente a quienes han estado siempre a mi lado durante esta etapa.

Finalmente, debo agradecer también a los que me han ayudado a disfrutar del bodyboard, mi pasión, mientras realizaba mis “deberes”favoritos.

Teresa Padilla Miranda
La Laguna, 7 de julio de 2022

Resumen · Abstract

Resumen

Es bien sabido que mediante los polinomios de Taylor no es posible emular funciones meromorfas en todo su dominio, debido a la localización de sus respectivos polos. Esto deja de ser una barrera cuando se toma como aproximante una adecuada función racional. Así, en la búsqueda de la mejor aproximación local mediante funciones racionales surgen los protagonistas de esta memoria, los aproximantes Padé. Si ampliamos nuestras aspiraciones a la aproximación simultánea de un vector de funciones aparecen de manera natural los aproximantes de Hermite-Padé. En ambos casos, obtener una caracterización del denominador será esencial para obtener resultados importantes de convergencia. Para ello, restringiremos nuestro estudio a funciones de Markov, y en el caso múltiple además a un sistema Angelesco. Asimismo, esta pondrá de relieve la estrecha relación entre la aproximación Padé, la teoría de polinomios ortogonales y las fórmulas de cuadratura; y paralelamente lo hará la correspondiente a la aproximación Hermite-Padé con las pertinentes teorías múltiples.

Palabras clave: *Aproximantes Padé – Aproximación Hermite-Padé – Funciones de Markov – Polinomios ortogonales – Fórmulas de cuadratura ....*

Abstract

It is well known that through Taylor polynomials it is not possible to emulate meromorphic functions in their entire domain, due to the location of their respective poles.This is no longer a barrier when a suitable rational function is taken as approximant. Thus, in the search for the best local approximation through rational functions the protagonists of this report, Padé approximants, emerge. If we broaden our aspirations to the simultaneous approximation of a vector of functions, Hermite-Padé approximants appear naturally. In both cases, knowing a characterization of the denominator will be essential to obtain important convergence results. Therefore, our study will be reduced to Markov functions, and in the multiple case in addition to an Angelesco system. Also, it will highlight the close relationship among Padé approximation, orthogonal polynomials theory and quadrature formulas; and in parallel, that of the Hermite-Padé approximation with the multiples theories pertinent.

Keywords: *Padé Approximants – Hermite-Padé Approximation – Markov functions – Quadrature formulas–Angelesco system ....*

Contenido

Introducción

El problema de aproximar una cantidad dada es uno de los desafíos más antiguos a los que se han enfrentado los matemáticos. Probablemente, el primer matemático que afrontó este problema fue Euler en 1777, cuando estudió la mejor aproximación posible a la relación entre latitudes y altitudes considerando todos los puntos de un meridiano entre latitudes conocidas, es decir, sobre un intervalo. Este tipo de problemas despertó interés en otros reconocidos matemáticos como Bernstein, Chebyshev, Hermite, Kolmogorov, Lagrange, Markov, los cuales asentarían lo que hoy en día se conoce por *Teoría de Aproximación*. Esta teoría sería establecida formalmente con la publicación *Journal of Approximation Theory in 1968*. Aunque bien es cierto que a mediados del siglo XX se obtuvieron numerosos resultados trascendentales, no se conocería realmente la productividad de ellos hasta la consolidación de los ordenadores y el gran desarrollo de la computación digital a finales del siglo pasado.

Uno de los principales retos de la teoría de aproximación es emular todo lo posible el comportamiento de funciones "complicadas"(ya sea porque no se conoce una expresión explícita o porque resulten difícilmente manejables) mediante funciones mucho más sencillas, las cuales reciben el nombre de *Aproximantes*. Comúnmente los polinomios de Taylor, uno de los grandes protagonistas de este Grado, han ofrecido unas características idóneas para realizar el papel de aproximante. Sin embargo, como se verá en el primer capítulo, tienen ciertas limitaciones, las cuales nos obligan a buscar otras alternativas. Por ejemplo, si la función es meromorfa, caso que será considerado en esta memoria, ningún polinomio podrá reproducir los polos de la función, por presentar exclusivamente polos en el infinito. De esta forma, la siguiente opción que surge de manera natural es considerar como aproximante una función racional. Así, si en vez de plantearnos el *problema de interpolación de Taylor* para la familia de polinomios lo hacemos para la clase de las funciones racionales, obtenemos una única fracción de polinomios conocida como *Aproximante de Padé*. El responsable del nombre de estas funciones fue Henri Padé, el primer matemático en estudiarlos sitemáticamente y disponerlos en una tabla de doble entrada conocida como

Tabla de Padé, en la que destacan dos importantes sucesiones: las filas y las diagonales. La convergencia por filas no será un desafío dado que el número de polos estará fijado, mientras que, por el contrario, en las diagonales sí lo será al crecer estos indefinidamente a medida que se avanza por la diagonal. Por ello, buscaremos una caracterización del denominador que nos dé información relevante, principalmente sobre la colocación y el comportamiento de sus ceros. Al final del primer capítulo, nos percataremos de que necesitaremos considerar el problema de interpolación en el infinito, así como restringir la familia de funciones a aproximar para obtener resultados generales sobre la convergencia. Esto dará lugar al segundo capítulo.

En el capítulo 2, veremos que cuando restringimos la clase de funciones a la formada por las funciones de Markov, el denominador es el polinomio ortogonal con respecto a una medida μ. Esto invita a dedicar un tiempo a estudiar la teoría de polinomios ortogonales. Una sección estará así destinada a estudiar, principalmente, aquellos resultados fundamentales para demostrar la convergencia tanto cualitativa como cuantitativa de los aproximantes Padé. Tras ello tendremos a nuestro alcance todas las herramientas necesarias para obtener dichos resultados de convergencia. Finalmente, como cabría esperar, el hecho de que el denominador Padé sea justamente el polinomio ortogonal, dará lugar a *Fórmulas de Cuadratura de Gauss-Christoffel*.

Finalmente, en el tercer y último capítulo de esta memoria, abordaremos el caso múltiple, la *Aproximación Hermite-Padé*. En este capítulo, en resumidas cuentas, trataremos de adaptar los resultados obtenidos en el capítulo anterior a este nuevo contexto donde se busca aproximar simultáneamente un vector de r funciones. Veremos que algunos de ellos se podrán obtener en esta nueva situación, limitándonos nuevamente a funciones de *Markov*, a las cuales se exigirá que los soportes de sus respectivas medidas asociadas mantengan a lo sumo contacto en los extremos, es decir, exigiremos que el vector de medidas sea *Angelesco*.

1

Capítulo 1: Introducción. Aproximación Padé

En este primer capítulo se introducen los aproximantes de Padé. Se ilustrará la utilidad de este tipo funciones para aproximar funciones meromorfas, señalando algunas de sus propiedades más relevantes, las cuáles justifican que los aproximantes de Padé hayan sido objeto de estudio e interés para prominentes matemáticos desde el siglo XIX.

1.1. Motivación.

En pleno siglo XXI es un hecho que el presente y futuro de la sociedad está en manos de la digitalización. Este proceso resulta inviable sin las aportaciones de la computación científica. Y, aquí, la eficacia, esto es, la optimización del tiempo, recursos y del coste computacional, es lo más perseguido. En este sentido, en el momento de aproximar una función resulta evidente (por su sencillez) que la familia más versátil que se ajusta a lo anterior es la formada por los polinomios. Veremos que, en efecto, los polinomios son excelentes aproximantes para una amplia clase de funciones, las funciones enteras.

Los polinomios con esta virtud se denominan *polinomios de Taylor*, pues satisfacen *el problema de interpolación de Taylor* definido de la siguiente forma:

Definición 1.1. *Sea f una función analítica en $z = a$ y g otra función de una clase dada. Si g verifica:*

$$f(z) - g(z) = C(z-a)^s + \textit{potencias de } z\text{-}a \textit{ de orden superior a } s, z \to a,$$

con s un entero positivo lo más grande posible, diremos que la función g resuelve el problema de interpolación de Taylor en dicha clase.

Cuando dicho problema se plantea para los polinomios de grado a lo sumo n, su existencia está garantizada por el denominado polinomio de Taylor de grado n asociado a f, siendo $s = n + 1$.

Mostraremos la cualidad de aproximación más relevante de estos polinomios.

Sea $f \in \mathcal{H}(\Omega)$ (holomorfa en Ω) con $\Omega \subset \mathbb{C}$ un dominio, abierto y conexo. Dado que toda función holomorfa es analítica tenemos que, $\forall a \in \Omega$,

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{k)}(a)}{k!}(z-a)^k, \forall z \in D(a,r),$$

siendo r la distancia de a a la frontera de Ω, esto es, $r = d(a, \partial\Omega)$. Mientras que el polinomio de Taylor de f de grado n viene dado por

$$P_n(f,a,z) = \sum_{k=0}^{n} \frac{f^{k)}(a)}{k!}(z-a)^k,$$

es decir, se trata de la suma enésima parcial de dicha serie. Consecuentemente, $\{P_n(f,a,z)\}_{n\in\mathbb{N}}$ diverge fuera del disco de convergencia de la serie.
Veamos que $P_n(f,a,z) \overset{\text{c.u.}}{\rightrightarrows} f(z)$ en $D(a,r)$, es decir, la sucesión de polinomios de Taylor de f converge uniformemente a f en subconjuntos compactos de $D(a,r)$.

Un primer paso para conseguir la convergencia es tener una expresión apropiada del error cometido. Por definición,

$$f(z) - P_n(f,a,z) = \sum_{k=n+1}^{\infty} \frac{f^{k)}(a)}{k!}(z-a)^k.$$

Esta expresión del error es poco eficiente, buscamos otra que se pueda estimar de manera más concisa.

Aplicando la fórmula integral de Cauchy para la derivada k-ésima, dado que $f \in \mathcal{H}(\Omega)$,

$$f^k(a) = \frac{k!}{2\pi i}\int_{\Gamma_\delta} \frac{f(t)}{(t-a)^{k+1}}dt,$$

siendo $\Gamma_\delta = \partial D(a,\delta)$, $\delta < r$. Teniendo en cuenta que $\forall z \in D(a,\delta)$ se verifica $|\frac{z-a}{t-a}| < 1$, $t \in \Gamma_\delta$, obtenemos que

$$f(z) - P_n(f,a,z) = \tfrac{1}{2\pi i}\int_{\Gamma_\delta} \tfrac{f(t)}{t-a} \textstyle\sum_{k=n+1}^{\infty} (\tfrac{z-a}{t-a})^k dt = \tfrac{1}{2\pi i}\int_{\Gamma_\delta} \tfrac{f(t)(z-a)^{n+1}}{(t-a)^{n+1}(t-z)}dt =$$
$$= \tfrac{(z-a)^{n+1}}{2\pi i}\int_{\Gamma_\delta} \tfrac{f(t)}{(t-a)^{n+1}(t-z)}dt.$$

Considerando la parametrización $t = a + \delta e^{i\theta}, \theta \in [0, 2\pi]$ para la curva Γ_δ y tomando módulos:

$$|f(z) - P_n(f,a,z)| \leq \frac{|z-a|^{n+1}}{2\pi}\int_0^{2\pi} \frac{|f(a+\delta e^{i\theta})|}{\delta^{n+1}|a+\delta e^{i\theta} - z|}\delta d\theta.$$

Sean $M_f := \max_{z \in \Gamma_\delta} |f(z)|$ y $d(z, \Gamma_\delta)$ la distancia de z a Γ_δ. Se sigue que:

$$|f(z) - P_n(f,a,z)| \leq \frac{M_f}{d(z,\Gamma_\delta)} \frac{|z-a|^{n+1}}{\delta^n}, \quad \forall z \in D(a,\delta),$$

si tomamos límite superior de la raíz n-ésima

$$\overline{\lim}_{n\to\infty} |f(z) - P_n(f,a,z)|^{\frac{1}{n}} \leq \frac{|z-a|}{\delta}, \quad \forall z \in D(a,\delta), \quad \forall \delta < r,$$

se concluye que

$$\overline{\lim}_{n\to\infty} |f(z) - P_n(f,a,z)|^{\frac{1}{n}} \leq \frac{|z-a|}{r}, \quad \forall z \in D(a,r).$$

Observamos que si la función es entera, esto es $\Omega = \mathbb{C}$, podemos hacer tender r a ∞, concluyéndose

$$\overline{\lim_{n\to\infty}} |f(z) - P_n(f,a,z)|^{\frac{1}{n}} = 0.$$

Por tanto, queda demostrado que los polinomios de Taylor convergen uniformemente en compactos en todo el plano complejo para las funciones enteras con velocidad más que geométrica.

Sin embargo, si f es meromorfa en $\mathbb{C}$, es decir, una función analítica cuyas singularidades son únicamente polos, además de una posible singularidad esencial en el infinito, los polinomios de Taylor no son idóneos para aproximar dichas funciones en sus respectivos dominios de meromorfía, ya que los polinomios no dejan de ser funciones enteras que poseen un polo en el infinito de orden el grado del polinomio. Por ello, únicamente son muy buenos aproximantes para las funciones enteras.
Ante la imposibilidad de reproducir polos finitos, surge la propuesta de usar como aproximantes de las funciones meromorfas las funciones racionales, cocientes de polinomios, pues retienen varias propiedades de los polinomios, continúan siendo funciones sencillas, fácilmente computables, entre otras propiedades; y, quizás, sus polos reproduzcan los de la función a aproximar.

1.2. Aproximantes de Padé

1.2.1. Interpolación racional de Taylor

A partir de ahora, sin pérdida de generalidad, tomamos, en aras de la claridad, a=0.

Nos planteamos ahora el problema de interpolación de Taylor mediante funciones racionales. Concretamente, consideramos el subconjunto de funciones racionales que incluye a los polinomios de grado a lo sumo n, es decir, a $\mathbb{P}_n$:

$$\mathcal{R}_{n,m} = \left\{ \frac{Q(z)}{P(z)} / Q \in \mathbb{P}_n \ \text{ y } \ P \in \mathbb{P}_m : P \not\equiv 0 \right\}.$$

Esto es, buscamos $r_{n,m} \in \mathcal{R}_{n,m}$ tal que:

$$f(z) - r_{n,m} = \mathcal{O}(z^s), \quad z \longrightarrow 0, \tag{1.1}$$

siendo s lo más grande posible, donde $\mathcal{O}(z^s) = Az^s +$ potencias superiores de z.

Dado que los elementos de $\mathcal{R}_{n,m}$ dependen de $n+m+1$ coeficientes, parece natural exigir $s = n+m+1$. Sin embargo, obsérvese que el problema ha dejado de ser lineal, dificultándose de esta manera el análisis de su solución. Es más, este problema de interpolación racional de Taylor no siempre tiene solución.

Por ejemplo, supongamos que queremos aproximar $f(z) = \frac{1}{1-z^2}$, $f \in \mathcal{H}(\mathbb{D})$, siendo $\mathbb{D}$ el disco unidad, por un elemento de $\mathcal{R}_{1,1}$ verificando (1.1). De existir, sabemos que debe ser de la forma $r_{1,1}(z) = \frac{az+b}{cz+d}$, $a, b, d, c \in \mathbb{C}$, tal que $(c, d) \neq (0, 0)$, y debe verificar $f(z) - r_{1,1}(z) = \mathcal{O}(z^3)$. Observamos que $r_{1,1}(z) = f(z) + \mathcal{O}(z^3) = \sum_{k=0}^{\infty} z^{2k} + \mathcal{O}(z^3) = 1 + z^2 + \mathcal{O}(z^3)$ y por tanto, $r'_{1,1}(0) = 0$, lo cual es absurdo, ya que $r_{1,1}$ es una transformación de Möbius, y por tanto, inyectiva.

1.2.2. Aproximación Padé

Surge así la siguiente cuestión, ¿será posible modificar ligeramente el problema de interpolación racional de Taylor de tal forma que se pueda garantizar solución y que coincida con el que resuelve (1.1) cuando este tenga solución? Veamos que sí. Para ello, linealizamos el problema, pasando a denominarse *problema débil de interpolación de Taylor*. Consecuentemente, el problema de interpolación de Taylor también es conocido por *el problema fuerte de interpolación de Taylor.*

Siendo ahora nuestro problema el siguiente:

$$\begin{aligned} &a)\ \exists\, Q_{n,m} \in \mathbb{P}_n \quad y \quad \exists\, P_{n,m} \in \mathbb{P}_m, \quad P_{n,m} \not\equiv 0, \\ &b)\ P_{n,m}(z)f(z) - Q_{n,m}(z) = \mathcal{O}(z^{n+m+1}), \quad z \longrightarrow 0. \end{aligned} \tag{1.2}$$

De (1.2), observamos que $Q_{n,m}$ es el polinomio de Taylor de $P_{n,m}f$ de grado n en $z = 0$. Por tanto, $Q_{n,m}$ está determinado unívocamente por $P_{n,m}$. Por otra parte, también de (1.2) observamos que los coeficientes de $P_{n,m}f$ que acompañan a $z^{n+1}, z^{n+2}, \ldots, z^{n+m}$ deben ser cero. Nos percatamos de que este análisis conduce a un sistema homogéneo de m ecuaciones con m+1 incógnitas, los coeficientes de $P_{n,m}$. De modo que, siempre hay solución distinta de la trivial, es decir, existe $P_{n,m} \in \mathbb{P}_m \quad y \quad P_{n,m} \not\equiv 0$, verificando (1.2). Determinado $P_{n,m}$ basta tomar, como hemos mencionado, como $Q_{n,m}$ el polinomio de Taylor de grado n asociado a $P_{n,m}f$.

A pesar de que $P_{n,m}$ no está unívocamente determinado, consecuentemente tampoco lo está $Q_{n,m}$, puesto que el sistema de ecuaciones que satisface presenta un grado de libertad, su cociente $\frac{Q_{n,m}}{P_{n,m}}$ sí lo está. En efecto, supongamos que $R_{n,m} = \frac{Q_{n,m}}{P_{n,m}}$ y $R'_{n,m} = \frac{Q'_{n,m}}{P'_{n,m}}$ son soluciones distintas de (1.2), entonces

$$(1) \quad P_{n,m}(z)f(z) - Q_{n,m}(z) = \mathcal{O}(z^{n+m+1}), \;\; z \to 0,$$

$$(2) \quad P'_{n,m}(z)f(z) - Q'_{n,m}(z) = \mathcal{O}(z^{n+m+1}), \;\; z \to 0.$$

Multiplicamos (1) y (2) por $P'_{n,m}$ y $P_{n,m}$, respectivamente. Al restar, se sigue que $h(z) := P_{n,m}(z)Q'_{n,m}(z) - P'_{n,m}(z)Q_{n,m}(z)$, $h \in \mathbb{P}_{n+m}$, verifica:

$$h(z) = \mathcal{O}(z^{n+m+1}), \;\; z \to 0,$$

es decir, h tiene al origen como cero de multiplicidad $n + m + 1$. Dado que $h \in \mathbb{P}_{n+m}$, entonces $h \equiv 0$. Y, por tanto,

$$R_{n,m}(f) = R'_{n,m}(f).$$

Definición 1.2. *La única función racional $\frac{Q_{n,m}}{P_{n,m}} \in \mathcal{R}_{n,m}$ que satisface (1.2) recibe el nombre de Aproximante de Padé asociado a f en $z = 0$ de orden (n,m) y la denotamos por $\Pi_{n,m}(f)$ (otra notación muy utilizada es $[n/m]_f$).*

Dentro de la familia de los polinomios tenemos que los polinomios de Taylor son la mejor aproximación local de funciones analíticas, de forma paralela, los aproximantes de Padé son la mejor aproximación racional local de funciones meromorfas en $\mathcal{R}_{n,m}$. Por tanto, los aproximantes de Padé son una extensión racional de los polinomios de Taylor. Sin embargo, a diferencia de los polinomios de Taylor los aproximantes de Padé no siempre resuelven el problema de interpolación de Taylor. Esto se debe a un posible defecto de interpolación originado por la aparición de raíces comunes en $Q_{n,m}$ y $P_{n,m}$ al linealizar el problema. Una de estas raíces puede ser z=0 (si $P_{n,m}$ tiene un cero en $z = 0$ de orden d_n, entonces $Q_{n,m}$ también lo posee con al menos ese mismo orden). En estas circunstancias, como consecuencia de la definición dada y la unicidad obtenemos:

$$f(z) - \Pi_{n,m}(f) = \mathcal{O}(z^{n+m+1-d_n}), \quad z \to 0,$$

de modo que perdemos d_n condiciones de interpolación. En efecto, si $P_{n,m} = z^{d_n}p_{n,m}$, $p_{n,m}(0) \neq 0$, entonces $Q_{n,m} = z^{d_n}q_{n,m}$. Luego,

$$z^{d_n}p_{n,m}(z)f(z) - z^{d_n}q_{n,m}(z) = \mathcal{O}(z^{n+m+1}), \quad z \to 0,$$

equivalentemente,

$$p_{n,m}(z)f(z) - q_{n,m}(z) = \mathcal{O}(z^{n+m+1-d_n}), \quad z \to 0,$$

y, por tanto,

$$f(z) - \frac{q_{n,m}(z)}{p_{n,m}(z)} = \mathcal{O}(z^{n+m+1-d_n}), \quad z \to 0.$$

Concluimos que el elemento $\Pi_{n,m} \in \mathcal{R}_{n,m}$ solo interpola con multiplicidad $n + m + 1 - d_n$.

Ahora bien, si $Q_{n,m}(0) \neq 0$, entonces el aproximante de Padé coincide con la función racional que resuelve el problema de interpolación racional de Taylor para funciones racionales.

1.2.3. Pinceladas históricas.

La primera aparición testificada de los aproximantes de Padé fue en 1731 por el matemático George Anderson quien envió una carta a William Jones donde aparecían tales funciones. Más tarde, Joseph Louis Lagrange, en 1777, mostró un método para dar la solución de una ecuación diferencial ordinaria en forma de fracción continua. Lagrange redujo los sucesivos convergentes de dicha fracción continua (poco manejables) a fracciones ordinarias, cuyos desarrollos en serie tenían la propiedad de coincidir con el de la serie de los respectivos convergentes hasta el término cuya potencia era la suma de los grados del denominador y numerador. Este realmente fue el nacimiento certificado de los aproximantes Padé. En 1845, Carl Gustav Jacob Jacobi dio una fórmula para el cálculo de estos aproximantes como cociente de dos determinantes.

No obstante, el nombre de estos aproximantes se debe a Henri Padé (1863-1953). Padé llevó a cabo un estudio sistemático de ellos en su tesis (1892), bajo la dirección de Charles Hermite. En ella dispuso los aproximantes en una tabla de doble entrada, ahora conocida como *Tabla de Padé*, obtuvo propiedades algebraicas de la misma y estudió con detalle la asociada a la función exponencial (ver [1] y referencias contenidas en él para más detalles).

A finales del siglo XIX y principios del XX, se obtuvieron numerosas propiedades algebraicas relacionadas con la tabla, así como algunos resultados cualitativos de convergencia. Pero no hubo un desarrollo significativo de la teoría hasta los años 60 del siglo pasado cuando dos hechos trascendentales lo precipitaron:

1. Surgen numerosas aplicaciones en Física, en Química, en Análisis Numérico, etc.
2. Se consolida la llegada de los ordenadores digitales, los cuales permitían liberar al científico de realizar cálculos tediosos y costosos.

La sinergia de estos hechos despierta en los matemáticos de la época un renovado interés, por su utilidad, y se da un “gran salto” en su consolidación, convirtiéndose en una herramienta matemática con entidad propia. Dando lugar, en función de las aplicaciones y los recursos disponibles, al desarrollo de las diferentes extensiones existentes y a otras nuevas: Aproximación tipo-Padé (fijando polos de antemano), Aproximación Hermite-Padé (interpolación simultánea de varias funciones) y Aproximación Padé multipuntual (interpolación en más de un punto); cada una de ellas acompañada de sus respectivos resultados de convergencia cuantitativos. Cabe destacar que estas generalizaciones, desde sus orígenes, dieron muestra de su aplicabilidad.

Entre sus primeras aplicaciones conviene destacar su contribución, a la teoría de números, en particular, a la demostración de Charles Hermite de la trascendencia del número e (1873) y a la de Ferdinand von Lindemann de π (1882), ambos problemas fueron resueltos mediante unas funciones, bautizadas posteriormente como aproximantes de Hermite-Padé; y en Física, a la detección de algunas propiedades magnéticas del modelo Ising cuyo autor fue George Alan Backer, Jr, y en el estudio de la resolvente y del espectro de un operador autoadjunto, de gran importancia en el desarrollo de la teoría mecánica cuántica y campos cuánticos (ver [2], [3]).

1.2.4. Tabla de Padé

Como hemos señalado, Padé dispuso los aproximantes de Padé de la siguiente forma:

$$\begin{array}{cccc} \Pi_{0,0} & \Pi_{1,0} & \Pi_{2,0} & \cdots \\ \Pi_{0,1} & \Pi_{1,1} & \Pi_{2,1} & \cdots \\ \Pi_{0,2} & \Pi_{1,2} & \Pi_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

dónde las filas varían acorde al índice m, número de polos del aproximante, y las columnas según n, número de ceros del aproximante. Sean $n, m \in \mathbb{N}$, destacamos las siguientes dos sucesiones de la tabla de Padé de distinta naturaleza:

a) $\{\Pi_{n,m}\}_{n\in\mathbb{N}}$, m fijo (filas).
b) $\{\Pi_{n+l,n}\}_{n\in\mathbb{N}}$, $l \in \mathbb{Z}$ (diagonales).

<u>Nota</u>. La primera fila de la tabla de Padé es la sucesión de polinomios de Taylor.

Definición 1.3. *Diremos que la tabla de Padé asociada a f es normal si ninguna entrada se repite.*

El primer ejemplo de normalidad fue dado por Padé en 1892, en su tesis doctoral probó que la tabla de la función exponencial es normal.

La tabla de Padé posee una estructura de bloques. Esta característica resulta fundamental en el estudio de la convergencia, ver [5]. Señalamos dos resultados muy significativos.

Proposición 1.4. *Sea f una función analítica, $f(0) \neq 0$. Supongamos que existe $\Pi_{n,m}(f)$ verificando:*

$$P_{n,m}(z)f(z) - Q_{n,m}(z) = Az^{n+m+k+1} + \cdots, \quad z \longrightarrow 0, \text{ para un cierto } k \in \mathbb{N}.$$

Entonces,

$$\Pi_{n+i,m+j}(f) = \Pi_{n,m}(f), \quad \forall i, j = 1, ..., k.$$

Proposición 1.5 (Criterio de Krönecker). *f es una función racional si y solo si existe un bloque de dimensión infinita.*

Debemos señalar que si nos desplazamos por una fila de la tabla de Padé obtenemos una sucesión de aproximantes de Padé con el número de polos fijos, tantos como indique la fila. De modo que para aproximar una función meromorfa con un número finito de polos basta escoger una fila con tantos o más polos que la función. En el caso de que el número de polos del aproximante sea superior, se tratará de expulsar "los sobrantes"del dominio de meromorfía o en su defecto acumularlos en la frontera. Así, con este tipo de desplazamiento conseguimos que los denominadores pertenezcan a un espacio vectorial finito dimensional, lo cual permite obtener resultados de convergencia de cierta generalidad.

En 1902, R. F Bernard publica el primer resultado general sobre la convergencia de los aproximantes de Padé, conocido como "Teorema de Montessus de Ballore", ver [4].

Teorema 1.6. *Sea f una función analítica en un entorno de $z = 0$ y meromorfa con exactamente m polos (contando las multiplicidades) en el disco $D_r = \{z \in \mathbb{C} : |z| < r\}$. Entonces, se cumple:*

I) Para todo número natural n suficientemente grande $\Pi_{n,m}(f)$ tiene m polos.
II) Los polos de $\Pi_{n,m}(f)$ tienden, cuando n tiende a infinito,a los polos de f en D_r, de acuerdo con sus multiplicidades.
III) $\Pi_{n,m}(f)(z) \rightrightarrows f(z)$ en $D_r \backslash \{polos\ de f\}$ (convergencia uniforme en compactos).

1.3. Diagonal de la tabla de Padé.

Supongamos que f, la función a aproximar, es meromorfa y posee infinitos polos. En este caso, resulta necesario acudir a las diagonales de la tabla de Padé, pues las sucesiones de filas tienen un número finito de polos. El principal inconveniente de trabajar con las diagonales reside en que el número de polos no está acotado, crece con n, perdiendo de esta forma que el denominador pertenezca a un espacio finito dimensional $\mathbb{P}_m$. Este hecho impide usar las mismas técnicas utilizadas al estudiarlos por filas, dificultándose así enormemente el análisis de emular las singularidades de la función a aproximar.

Sin pérdida de generalidad, por propiedades algebraicas, ver [2], podemos centrarnos en la diagonal principal de la tabla de Padé, es decir, en la sucesión $\{\Pi_{n,n}\}_{n\in\mathbb{N}}$.

Sea f una función holomorfa en un dominio Ω tal que $0 \in \Omega$. Planteamos el problema que nos ocupará el resto del capítulo:

a) $\exists Q_n, P_n \in \mathbb{P}_n, \quad P_n \not\equiv 0.$
b) $P_n(z)f(z) - Q_n(z) = \mathcal{O}(z^{2n+1}), \quad z \longrightarrow 0.$

Como hemos señalado en la primera sección todo pasa por determinar P_n. Por ello, necesitamos profundizar más para obtener alguna caracterización suya que nos permita deducir propiedades de sus ceros, ya que el comportamiento de estos será esencial para el éxito en la aproximación.

Con el fin de lograr nuestro objetivo “explotaremos” las condiciones de interpolación mediante la teoría de variable compleja.

Observamos que la condición b) es equivalente a:

$$\frac{P_n(z)f(z) - Q_n(z)}{z^{2n+1}} \in H(\Omega).$$

Aplicando el teorema de Cauchy sobre una curva cerrada de Jordan Γ en Ω conteniendo al origen en su interior, se sigue que:

$$0 = \oint_\Gamma P_n(t)\frac{f(t)}{t^{2n+1}}dt - \oint_\Gamma \frac{Q_n(t)}{t^{2n+1}}dt = \oint_\Gamma P_n(t)\frac{f(t)}{t^{2n+1}}dt,$$

ya que la segunda integral se anula, puesto que $Q_n^{2n)}(0) = 0$. Nos percatamos de que esta misma cuenta es válida si la desarrollamos para $P\frac{P_n(z)f(z)-Q_n(z)}{z^{2n+1}}$, $\forall P \in \mathbb{P}_{n-1}$. Esto es, se verifica

$$\oint_\Gamma P(t)P_n(t)\frac{f(t)}{t^{2n+1}}dt = 0, \quad \forall P \in \mathbb{P}_{n-1}.$$

Por tanto, P_n es ortogonal a $\mathbb{P}_{n-1}$ respecto de la función peso $\frac{f(t)}{t^{2n+1}}$. No obstante, la ortogonalidad obtenida para P_n no es de gran utilidad, debido a que la función peso es compleja, variante (depende de n) y la curva de integración contenida en Ω es arbitraria, es más, la ortogonalidad no es hermitiana. Dicho de otra forma, no podemos ubicar dicha ortogonalidad en un contexto pre-Hilbert que sin lugar a duda nos ayudaría.

Podemos mejorar esta situación mediante el cambio de variable $t = \frac{1}{x}$, obteniendo

$$\oint_{\Gamma^{-1}} P(x)\hat{Q}_n(x)g(x)dx = 0, \quad \forall P \in \mathbb{P}_{n-1}, \tag{1.3}$$

donde $\hat{P}_n(x) = x^n P_n(\frac{1}{x})$ y $g(x) = f(\frac{1}{x})$. Pero, tan solo hemos logrado que la función peso g deje de ser variante. Por tanto, aún estamos muy lejos de obtener una caracterización que nos abra el camino de obtener propiedades, con cierta generalidad, para P_n. Sin embargo, si reflexionamos sobre esta última observación, parece que podremos tener algo más de éxito si en lugar de interpolar en $z = 0$, lo hacemos en $z = \infty$, dado que realmente lo que hemos hecho de manera implícita es la siguiente cuenta.

$$P_n(z)f(z) - Q_n(z) = \mathcal{O}(z^{2n+1}), \quad z \longrightarrow 0,$$

invirtiendo,

$$P_n(\frac{1}{z})f(\frac{1}{z}) - Q_n(\frac{1}{z}) = \mathcal{O}(\frac{1}{z^{2n+1}}), \quad z \longrightarrow \infty.$$

Multiplicando por z^n, se sigue

$$\hat{P}_n(z)g(z) - \hat{Q}_n(z) = \mathcal{O}(\frac{1}{z^{n+1}}), \quad z \longrightarrow \infty. \tag{1.4}$$

Finalmente, si $\partial\,(P_n) = n$, entonces

$$g(z) - \frac{\hat{Q}_n(z)}{\hat{P}_n(z)} = \mathcal{O}(\frac{1}{z^{2n+1}}), \quad z \longrightarrow \infty.$$

Esto es, $\frac{\hat{Q}_n}{\hat{P}_n}$ resuelve el problema de interpolación racional de Taylor en el infinito. Dicha función es conocida como *Aproximante de Padé de g de orden n en el infinito.*

<u>Nota</u>. Si la función a aproximar verificase $f(\infty) = 0$, tendríamos que $\partial(\hat{Q}_n) = n - 1$; si $f(\infty) = f'(\infty) = 0$, entonces $\partial(\hat{Q}_n) = n - 2$; y así sucesivamente.

Llegados a este punto, dos caminos se abren: debilitar la calidad de la convergencia o restringir la clase de funciones a aproximar. En el próximo capítulo, siguiendo la segunda vía, veremos que si consideramos el aproximante de Padé en el infinito de un determinado tipo de funciones, denominadas *funciones de Markov*, las condiciones de ortogonalidad serán hermitianas y, consecuentemente, podremos servirnos de las técnicas propias de un contexto pre-Hilbert. Esto nos permitirá obtener resultados de convergencia bastante generales.

2

Aproximación Padé y ortogonalidad. Fórmulas de cuadratura.

Tal y como evidenciamos en el capítulo anterior, al abordar la diagonal de la Tabla de Padé nos encontramos con un serio problema, el número de polos crece con n. Por ello, necesitamos algún tipo de caracterización de los denominadores que nos permita obtener información sobre sus polos, especialmente de su localización y comportamiento. La última caracterización obtenida en el capítulo anterior nos sugiere que, si queremos tener ciertas garantías de éxito, debemos considerar la interpolación en el infinito. No obstante, esta continúa sin ser lo suficientemente bondadosa, pues aún es bastante imprecisa para funciones arbitarias. Así, parece razonable pensar que una vía para llegar a buen puerto es reducir la clase de funciones a aproximar. Efectivamente, veremos que si particularizamos el problema a una familia de funciones, muy importante en la Física-Matemática, la caracterización nos conduce a los anhelados resultados de convergencia.

Finalmente, veremos que la aproximación Padé tiene una vinculación estrecha con las fórmulas de cuadratura.

2.1. Aproximación Padé. Funciones de Markov

Sea f una función que admite el desarrollo

$$f(z) = \sum_{k=0}^{\infty} f_k \frac{1}{z^{k+1}}, \quad z \longrightarrow \infty.$$

Hemos supuesto $f(\infty) = 0$ por comodidad. Sabemos que fijado $n \in \mathbb{N}$ existe una única función racional $\frac{Q_n}{P_n}$ que satisface las siguientes condiciones:

$$\begin{aligned} &i)\; Q_n \in \mathbb{P}_{n-1} \text{ y } P_n \in \mathbb{P}_n\,,\;\; P_n \not\equiv 0. \\ &ii)\; P_n(z)f(z) - Q_n(z) = \mathcal{O}(\frac{1}{z^{n+1}}), \quad z \longrightarrow \infty, \end{aligned} \tag{2.1}$$

denominada n-ésimo aproximante de Padé asociado a f en el infinito. La denotaremos por $\Pi_n(f)$.

Como vimos en el capítulo anterior, de (2.1) necesariamente los coeficientes de $P_n f$ que acompañan a las potencias $\frac{1}{z}, \frac{1}{z^2}, \frac{1}{z^3}, \cdots, \frac{1}{z^n}$ se deben anular, dando lugar a un sistema lineal homogéneo de n ecuaciones y $n+1$ incógnitas, los coeficientes de P_n. Si expresamos P_n como $P_n(z) = b_0 + b_1 z + \cdots + b_n z^n$, el sistema de ecuaciones a resolver para hallar sus coeficientes, matricialmente, es:

$$\begin{pmatrix} f_0 & f_1 & \cdots & f_n \\ f_1 & f_2 & \cdots & f_{n+1} \\ \vdots & \vdots & & \vdots \\ f_{n-1} & f_n & \cdots & f_{2n-1} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.2}$$

Al tratarse de un sistema homogéneo con un parámetro de libertad, hay soluciones no nulas, por tanto, existe $P_n \not\equiv 0$. Por otro lado, también de ii), Q_n es la parte principal del desarrollo en serie de Laurent del producto $P_n f$.

Para tener éxito en nuestro avance debemos considerar, como hemos mencionado, un determinado tipo de funciones, ya que, en general, tanto (2.2) como (1.3) resultan poco significativos. La familia a la que nos restringiremos es la formada por las *funciones de Markov*. Esta comprende muchas funciones elementales y juega un rol central en la teoría espectral de operadores autoadjuntos. La estrecha vinculación entre la aproximación Padé y las funciones de Markov, como veremos, está "apuntalada"por la teoría de polinomios ortogonales.

Definición 2.1. *Sea μ una medida de Borel positiva finita con soporte el intervalo cerrado y acotado $[a,b]$. Se denomina función de Markov asociada a μ a la transformada de Cauchy $\widehat{\mu}$ de la medida μ, es decir, a:*

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

Es inmediato verificar que $\widehat{\mu} \in \mathcal{H}(\widehat{\mathbb{C}} \setminus [a,b])$. Es más, gracias a la convergencia absoluta y uniforme de la serie geomética $\frac{1}{1-\frac{x}{z}}$ cuando $|\frac{x}{z}| < 1$, $\widehat{\mu}$ admite el siguiente desarrollo en el infinito

$$\widehat{\mu}(z) = \int_a^b \frac{d\mu(x)}{z-x} = \frac{1}{z}\int_a^b \frac{d\mu(x)}{1-\frac{x}{z}} = \sum_{k=0}^{\infty} \frac{c_k}{z^{k+1}},$$
$$c_k = \int_a^b x^k d\mu(x), \quad \forall k \in \mathbb{N},$$

donde a la sucesión $\{c_k\}_{k\in\mathbb{N}}$ se le denomina momentos de la medida μ.

Veamos que, a través de los siguientes ejemplos, algunas funciones de Markov tienen, en contra de lo esperado, expresiones explícitas relativamente sencillas (ver [3]).

1. Sea $d\mu(x) = dx$ la medida de Lebesgue en $[-1,1]$, entonces
$$\widehat{\mu}(z) = \int_{-1}^{1} \frac{dx}{z-x} = \log\left(\frac{z-1}{z+1}\right),$$
donde consideramos la rama principal del logaritmo, $\log 1 = 0$.
2. Sea $d\mu(x) = dx/\pi\sqrt{1-x^2}$ la medida de Chebyshev con soporte en $[-1,1]$, entonces
$$\widehat{\mu}(z) = \frac{1}{\pi}\int_{-1}^{1} \frac{dx}{(z-x)\sqrt{1-x^2}} = \frac{1}{\sqrt{z^2-1}},$$
donde $\sqrt{z^2-1} > 0$ para $z > 1$.
3. Sea $d\mu(x) = \operatorname{sen}(\pi\alpha)[(1+x)/(1-x)]^{\alpha}dx, \alpha \in (-1,1)$ con soporte [-1,1]
$$\widehat{\mu}(z) = \operatorname{sen}(\pi\alpha)\int_{-1}^{1} \left(\frac{1+x}{1-x}\right)^{\alpha} \frac{dx}{z-x} = \left(\frac{z+1}{z-1}\right)^{\alpha} - 1,$$
donde $[(z+1)/(z-1)]^{\alpha} > 0$ si $z > 1$.

Nota. Nótese que si el soporte de μ está formado por k puntos, la correspondiente función de Markov es una función racional, entonces su aproximación carece de sentido, pues es evidente que $\Pi_n = \widehat{\mu},\ n \geq k$.

Por otra parte, el siguiente resultado nos advierte que en $[a,b]$ no solo se localizan las singularidades de las funciones de Markov, sino que también se ubican sus posibles ceros.

Proposición 2.2. *La función de Markov $\widehat{\mu}$ tiene sus ceros (si los tiene) en $[a,b]$.*

Demostración.
$$\widehat{\mu}(z_0) = 0 \Longleftrightarrow \int_a^b \frac{d\mu(x)}{x-z_0} = 0.$$
Identificando la parte real y la parte imaginaria de dicha ecuación
$$0 = \int_a^b \frac{d\mu(x)}{x-z_0} = \int_a^b \frac{x-\bar{z}_0}{|x-z_0|^2}d\mu(x) = \int_a^b \frac{x}{|x-z_0|^2}d\mu(x) - \bar{z}_0\int_a^b \frac{d\mu(x)}{|x-z_0|^2}$$
resulta que
$$\bar{z}_0 = \int_a^b x\frac{d\mu(x)}{|x-z_0|^2}\frac{1}{\int_a^b \frac{d\mu(t)}{|t-z_0|^2}} = \int_a^b x\,d\widehat{\mu}(x), \quad d\tilde{\mu}(x) = \frac{1}{\int_a^b \frac{d\mu(t)}{|t-z_0|^2}}\frac{1}{|x-z_0|^2}d\mu(x),$$
siendo $d\tilde{\mu}(x)$ una medida de probabilidad, consecuentemente $\bar{z}_0 = x_0 \in [a,b]$ y, por tanto, $z_0 = \bar{z}_0 \in [a,b]$. □

Si particularizamos (2.2) para una función de Markov $\widehat{\mu}$ el sistema resultante es

$$\begin{pmatrix} c_0 & c_1 & \cdots & c_n \\ c_1 & c_2 & \cdots & c_{n+1} \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2n-1} \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{2.3}$$

En él, observamos que

$$H_n = \begin{pmatrix} c_0 & c_1 & \cdots & c_{n-1} \\ c_1 & c_2 & \cdots & c_n \\ \vdots & \vdots & & \vdots \\ c_{n-1} & c_n & \cdots & c_{2(n-1)} \end{pmatrix}$$

es una matriz conocida como matriz de *Hankel* de orden n. Es más, en esta situación H_n es, en particular, una matriz de Gram asociada al siguiente producto interior

$$\langle p, q \rangle = \int_a^b p(x)q(x)d\mu(x), \quad \forall p, q \in \mathbb{P},$$

ya que podemos reinterpretar los momentos como sigue

$$c_n = \int_a^b x^n d\mu(x) = \int_a^b x^i x^j d\mu(x) =< x^i, x^j >, \;\; i + j = n.$$

Por tanto, la matriz H_n es definida positiva y de (2.3)

$$H_n \begin{pmatrix} b_0 \\ \vdots \\ b_{n-1} \end{pmatrix} = -b_n \begin{pmatrix} c_n \\ \vdots \\ c_{2n-1} \end{pmatrix}, \quad b_n \neq 0, \tag{2.4}$$

queda garantizado que P_n es único, salvo constante multiplicativa, y, además, es de grado n. Por consiguiente,

Teorema 2.3. *La tabla de Padé asociada a cualquier función de Markov es normal.*

Obsérvese que las ecuaciones del sistema (2.4)

$$\sum_{j=0}^{n} b_j c_{k+j} = 0, \quad k = 0, 1, \ldots, n-1,$$

gracias a la linealidad del producto interior, pueden ser entendidas de la siguiente manera

$$\langle P_n, x^k \rangle = 0, \quad k = 0, 1, \ldots, n-1.$$

Concluyéndose así que el denominador del aproximante de Padé es ortogonal a $\mathbb{P}_{n-1}$. Si prestamos atención, realmente observamos que es una caracterización, para ello basta seguir el desarrollo anterior hacia atrás. Consecuentemente, tenemos la siguiente proposición.

Proposición 2.4. *P_n es el denominador del aproximante de Padé de orden n asociado a la función de Markov $\widehat{\mu}(z) = \int_a^b \frac{d\mu(x)}{z-x}$ si y solo si es el enésimo polinomio ortogonal con respecto a la medida μ.*

Hemos obtenido una caracterización de P_n, resta ver si es la deseada.

Observación 2.5. Señalar que otra vía para llegar a ella hubiese sido particularizar (1.3) para una función de Markov y realizar un uso adecuado de la teoría de variable compleja. Efectivamente, partiendo de

$$\int_{\Gamma^{-1}} P(x)P_n(x)\widehat{\mu}(x)dx = 0, \quad \forall P \in \mathbb{P}_{n-1},$$

donde Γ^{-1} rodea a [a,b], dado que la función $g(x,t) = \frac{1}{x-t}$ es continua en Γ^{-1}x$[a,b]$, aplicando Fubini, y luego Cauchy, teniendo en cuenta que PP_n es una función entera, obtenemos que

$$0 = \int_{\Gamma^{-1}} P(x)P_n(x)\left(\int_a^b \frac{d\mu(t)}{x-t}\right)dx = \int_a^b P(t)P_n(t)d\mu(t), \quad \forall P \in \mathbb{P}_{n-1}.$$

2.2. Polinomios ortogonales

Los polinomios ortogonales están presentes en varios tópicos del Análisis Matemático. Intervienen en la resolución de problemas de Sturm-Liouville, en la teoría de Fourier, en teoría espectral de operadores, en teoría de señales, etcétera ([6]). Siendo especialmente relevantes en la aproximación Padé de funciones tipo Markov como ilustraremos en esta memoria, ya que aliviarán en gran medida las dificultades en la aproximación Padé.

Sea μ una medida de Borel positiva finita cuyo soporte, $supp(\mu)$, está formado por una cantidad infinita de puntos y contenido en un subconjunto cerrado $E \subseteq \mathbb{R}$. Si E es un conjunto no acotado asumiremos que $x^n \in \mathcal{L}_1(\mu), n \in \mathbb{N}$. Es decir, $\mathbb{P} \subset \mathcal{L}_1(\mu)$. Denotaremos por $\mathcal{M}(E)$ al conjunto de todas las medidas soportadas en E con las citadas propiedades.

Para todo elemento $\mu \in \mathcal{M}(E)$, se define un espacio pre-Hilbert

$$\mathcal{L}_2(\mu) = \{f : E \to \mathbb{R} / \int_E |f(x)|^2 d\mu(x) < \infty\},$$

dotado del producto interior $\langle \cdot, \cdot \rangle$

$$\langle f, g\rangle = \int_E f(x)g(x)d\mu(x), \quad \forall f, g \in \mathcal{L}_2(\mu), \tag{2.5}$$

cuya norma asociada es $\|f\|^2 = \langle f, f\rangle$, y, con una propiedad que lo caracteriza

$$< xf, g >=< f, xg >, \quad \forall f, g \in \mathcal{L}_2(\mu).$$

Sin pérdida de generalidad se consideran tres situaciones: $E = [-1, 1]$, $E = [0, \infty) = \mathbb{R}^+$ ó $E = \mathbb{R}$, puesto que cualquier otro conjunto E se reduce a uno de los casos anteriores (mediante una transformación lineal u otra relación). Señalar que cuando $E = [-1, 1]$, $\mathbb{P} \subset \mathcal{L}_2(\mu)$ y, además, es denso en $\mathcal{L}_2(\mu)$, como consecuencia del teorema de Weierstrass. Por consiguiente, la familia de los polinomios constituyen una base de este espacio. Tenemos así que $\mathcal{L}_2(\mu)$ es un espacio Hilbert.

Definición 2.6. *Se dice que $\{p_n\}_{n=0}^{\infty}$ es una sucesión de polinomios ortogonales con respecto a la medida $\mu \in \mathcal{M}(E)$ si, y solo si:*

1. $\partial\,(p_n) = n, \forall n \in \mathbb{N}$.
2.
$$\langle p_n, p_m\rangle = \int_E p_n(x)p_m(x)d\mu(x) = \begin{cases} \|p_n\|^2 & si\ \ n = m, \\ 0 & si\ \ n \neq m. \end{cases}$$

Si además se tiene que $\|p_n\| = 1$, diremos que la sucesión $\{p_n\}_{n=0}^{\infty}$ es ortonormal.

En el caso de que el coeficiente director de p_n sea 1 para todo n, diremos que $\{p_n\}_{n=0}^{\infty}$ es la sucesión de polinomios ortogonales mónicos; lo diferenciaremos denotándolo por P_n. Es bien conocido que su existencia está garantizada, además de por (2.4), por el proceso de ortogonalización de Gram-Schmidt aplicado a la base canónica $\{1, x, \ldots, x^n\}$ de $\mathbb{P}_n, \forall n \in \mathbb{N}$.

A continuación, proporcionamos un resultado fundamental de la teoría de polinomios ortogonales de gran interés teórico-práctico.

Teorema 2.7 (Fórmula de recurrencia a tres términos). *La sucesión de polinomios ortonormales $\{p_n\}_{n=0}^{\infty}$ con respecto a μ verifica:*

$$zp_n(z) = a_{n+1}p_{n+1}(z) + b_np_n(z) + a_np_{n-1}(z), \;\; n \geq 1, \tag{2.6}$$

$$p_{-1}(x) \equiv 0, \;\; p_0(x) \equiv \frac{1}{\sqrt{c_0}},$$

donde

$$a_n = \frac{\|P_n\|}{\|P_{n-1}\|} > 0, \;\; b_n = \int_E xp_n^2(x)d\mu(x), \;\; n \geq 1.$$

Para su demostración ver [6].

En la siguiente proposición denotamos por $\mathcal{C}_0(\text{supp}(\mu))$ al menor intervalo que contiene a $\text{supp}(\mu)$, es decir, a la envolvente convexa del soporte de la medida μ.

Proposición 2.8. *Sea p_n el n-ésimo polinomio ortogonal respecto a μ, entonces p_n tiene todos sus ceros en el interior de $\mathcal{C}_0(\mathrm{supp}(\mu))$, $\mathring{\mathcal{C}}_0$, y además son simples.*

Demostración. Sabemos que $\int_E p_n(x)d\mu(x) = 0$. Como p_n tiene coeficientes reales necesariamente p_n tiene un cambio de signo en $\mathring{\mathcal{C}}_0$. Sean $x_1, x_2, \ldots, x_m$ los cambios de signo p_n en $\mathring{\mathcal{C}}_0$. Supongamos $m < n$. Consideremos el polinomio $Q(x) := (x - x_1)(x - x_2) \cdots (x - x_m)$. Se sigue que el polinomio $p_n(x)Q(x)$ no cambia de signo en $\mathring{\mathcal{C}}_0$. Luego,

$$\int_E p_n(x)Q(x)d\mu(x) \neq 0,$$

lo cual es absurdo ya que $p_n \perp \mathbb{P}_{n-1}$ y $Q \in \mathbb{P}_{n-1}$. Por tanto $m = n$. □

La siguiente proposición asegura que los ceros de polinomios ortogonales consecutivos se entrelazan, ofreciendo la posibilidad de que terminen llenando el soporte de la medida.

Proposición 2.9. *Los ceros de dos polinomios ortogonales de grados consecutivos, p_n y p_{n+1}, se entrelazan. En concreto, si etiquetamos los ceros de la siguiente forma: $x_{1,n} < x_{2,n} < \ldots < x_{n,n}$, $\forall n \in \mathbb{N}$, entonces*

$$x_{k,n+1} < x_{k,n} < x_{k+1,n+1}, \quad k = 1, \ldots, n. \tag{2.7}$$

Demostración. Sin pérdida de generalidad consideraremos la sucesión de polinomios ortonormales. En virtud del teorema de Bolzano, (2.7) es equivalente a

$$sig(p_{n-1}(x_{k,n})) = (-1)^{n-k}, \quad k = 1, \ldots n, \quad \forall n \in \mathbb{N}, \tag{2.8}$$

siendo $sig(x)$ la función signo, analíticamente

$$sig(x) = \begin{cases} \frac{x}{|x|} & , \text{si } x \neq 0 \\ 0 & , \text{si } x = 0. \end{cases}$$

Procederemos por inducción sobre n. Para $n = 1$, p_0 es una constante positiva, entonces (2.8) se cumple trivialmente, $p_0(x_{1,1}) > 0$. Supongamos que se verifica para n. Demostremos que el entrelazamiento de ceros es cierto para $n + 1$.

De la relación de recurrencia (2.6), al sustituir $x = x_{k,n}$, obtenemos

$$sig(p_{n+1}(x_{k,n})) = -sig(p_{n-1}(x_{k,n})), \quad k = 1, 2, \ldots, n.$$

Por tanto, de la hipótesis de inducción, $sig(p_{n+1}(x_{k,n})) = (-1)^{n+1-k}, k = 1, 2, \ldots, n$. Además, por haber tomado la sucesión de polinomios ortonormales, sus coeficientes directores son positivos, se sigue que

$$sig(p_{n+1}(+\infty)) = 1, \quad sig(p_{n+1}(-\infty)) = (-1)^{n+1}.$$

Luego, por el teorema de Bolzano, p_{n+1} tiene al menos un cero en cada uno de los siguientes intervalos

$$(-\infty, x_{1,n})(x_{1,n}, x_{2,n}), \ldots, (x_{n,n}, +\infty),$$

y, por tanto,

$$sig(p_n(x_{k,n+1})) = (-1)^{n+1-k}, \quad k = 1, \ldots, n+1,$$

como queríamos ver. □

Recordemos que el motivo de nuestro interés por los aproximantes de Padé surge de buscar una clase de funciones más versátil capaz de reproducir las singularidades de la función a aproximar. Y, más concretamente, lo que nos trajo hasta este capítulo, fue valorar la convergencia de funciones con infinitas singularidades, pues si el número de polos fuese finito ya el teorema de Montessus Ballore (teorema 1.6) nos garantizaba la idoneidad de los aproximantes de Padé en la emulación de las singularidades de la función. Nótese así, la relevancia de la proposición 2.8, la cual nos asegura que los polos del aproximante Padé van a encontrarse en el intervalo donde se hallan las singularidades de la función de Markov considerada.

Lo siguiente que nos planteamos de manera espontánea, de la proposición 2.2, es ver si los ceros del aproximante Padé se encuentran donde pueden estar los ceros de la función de Markov. Para ello, será imprescindible determinar una expresión de Q_n, con ese objetivo es esencial conocer la "mano derecha"de todo proceso de aproximación, la expresión del error.

Si particularizamos (2.1) para una función de Markov obtenemos que, para todo $P \in \mathbb{P}_n$,

$$P(z)(P_n(z)\widehat{\mu}(z) - Q_n(z)) = \mathcal{O}(\frac{1}{z}), \quad z \to \infty.$$

Por el Teorema de Cauchy para dominios no acotados, $P(z)(P_n(z)\widehat{\mu}(z) - Q_n(z)) \in \mathcal{H}(\widehat{\mathbb{C}} \setminus [a, b])$ y se anula en el infinito,

$$P(z)(P_n(z)\widehat{\mu}(z) - Q_n(z)) = \frac{1}{2\pi i} \int_{\Gamma} \frac{P(t)(P_n(t)\widehat{\mu}(t) - Q_n(t))}{t - z} dt, z \in \mathrm{Ext}(\Gamma),$$

donde Γ es una curva de Jordan que rodea a [a,b] con orientación negativa, es decir, se recorre en sentido horario.

Dado que PQ_n es una función entera y $z \in \mathrm{Ext}(\Gamma)$, $\frac{P(t)Q_n(t)}{t-z}$ como función de t es holomorfa en en el interior de Γ, entonces

$$\int_{\Gamma} \frac{P(t)Q_n(t)}{t-z} dt = 0.$$

Por lo tanto,

$$\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)} = \frac{1}{2\pi i P_n(z)P(z)} \int_{\Gamma} \frac{P(t)P_n(t)\widehat{\mu}(t)}{t-z} dt =$$

$$= \frac{1}{2\pi i P_n(z)P(z)} \int_{\Gamma} \Big(\int_a^b \frac{P(t)P_n(t)}{(t-z)(t-x)} d\mu(x) \Big) dt.$$

Por el Teorema de Fubini y aplicando la fórmula integral de Cauchy, cambiando la orientación de Γ, a la función $\frac{P(t)P_n(t)}{t-z}$ en el punto x, concluimos que

$$\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)} = \frac{1}{P_n(z)P(z)} \int_a^b \frac{P(x)P_n(x)}{z-x} d\mu(x). \tag{2.9}$$

Observación 2.10. En (2.9) llama la atención que obtenemos por cada elección del polinomio P en el espacio $\mathbb{P}_n$ distintas expresiones del mismo error. Obviamente están encriptadas las condiciones de ortogonalidad que hemos analizado. En efecto, sean P_1, P_2 dos polinomios cualesquiera de grado a lo sumo n. Observamos que

$$\frac{1}{P_n(z)P_1(z)} \int_a^b \frac{P_1(x)P_n(x)}{z-x} d\mu(x) - \frac{1}{P_n(z)P_2(z)} \int_a^b \frac{P_2(x)P_n(x)}{z-x} d\mu(x) =$$
$$= \frac{1}{P_n(z)} \Big[\frac{1}{P_1(z)P_2(z)} \int_a^b P_n(x) \frac{P_1(x)P_2(z) - P_1(z)P_2(x)}{z-x} d\mu(x) \Big] = 0$$

La última integral se anula ya que $\frac{P_1(x)P_2(z)-P_1(z)P_2(x)}{z-x} \in \mathbb{P}_{n-1}$ en la variable x y $P_n \perp \mathbb{P}_{n-1}$.

Los dos siguientes casos serán de interés en nuestro progreso.

1. Si escogemos $P(z) = P_n(z)$ en (2.9), obtenemos la siguiente expresión del error:
$$\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)} = \frac{1}{P_n^2(z)} \int_a^b \frac{P_n^2(x)}{z-x} d\mu(x). \tag{2.10}$$
2. Si consideramos $P(z) = 1$ en (2.9) resulta:
$$\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)} = \frac{1}{P_n(z)} \int_a^b \frac{P_n(x)}{z-x} d\mu(x).$$

Identidad que nos desvela una expresión integral del numerador Padé, ya que, equivalentemente, tenemos,

$$P_n(z)\widehat{\mu}(z) - Q_n(z) = \int_a^b \frac{P_n(x)}{z-x} d\mu(x), \tag{2.11}$$

despejando Q_n,

$$Q_n(z) = \int_a^b \frac{P_n(z) - P_n(x)}{z-x} d\mu(x) \in \mathbb{P}_{n-1}.$$

Como veremos, por méritos propios, este polinomio se ha ganado un nombre.

Definición 2.11 (Polinomios de segunda especie). *Se denomina enésimo polinomio de segunda especie asociado a μ al polinomio*

$$Q_n(x) = \int_E \frac{P_n(x) - P_n(t)}{x-t} d\mu(t),$$

donde P_n es el enésimo polinomio ortogonal mónico respecto de μ.

De la anterior expresión de Q_n vamos a poder conseguir de regalo el resultado perseguido, ya que de ella, obtendremos que los ceros de Q_n se van a entrelazar con los de P_n. Por lo tanto, vamos realmente a conseguir mayor información de lo que sucede en $[a, b]$, pues tendremos que los aproximantes de Padé tratan de revelarnos el extraño comportamiento de las singularidades de una función de Markov, generando indeterminaciones en $[a, b]$ a medida que aumenta n al colocar sus ceros y polos entrelazadamente en dicho intervalo.

Para ello, vamos a demostrar que Q_n también satisface la relación de recurrencia a tres términos que verifican los polinomios ortogonales.

De (2.6), se obtiene inmediatamente la relación de recurrencia a tres términos que satisfacen los polinomios ortogonales mónicos,

$$P_{n+1}(x) = (x + a_n)P_n(x) + b_n P_{n-1}(x), \quad a_n \in \mathbb{R}, \quad b_m \in \mathbb{R}^+, \quad \forall n \geq 1. \tag{2.12}$$

En primer lugar, comprobaremos que el "error Padé" (2.11)

$$E_n(z) = \int_a^b \frac{P_n(t)}{z-t} d\mu(t),$$

también la satisface. De (2.12), se tiene que

$$E_{n+1}(z) = \int_a^b \frac{P_{n+1}(t)}{z-t} d\mu(t) = \int_a^b \frac{(t+a_n)P_n(t) + b_n P_{n-1}(t)}{z-t} d\mu(t).$$

De la igualdad $t + a_n = t - z + z + a_n$ y de la linealidad de la integral, obtenemos

$$E_{n+1}(z) = -\int_a^b P_n(t) d\mu(t) + \int_a^b \frac{(z+a_n)P_n(t)}{z-t} d\mu(t) + \int_a^b \frac{b_n P_{n-1}(t)}{z-t} d\mu(t).$$

La primera integral es nula, puesto que P_n es ortogonal a las constantes. Consecuentemente, $\{E_n\}_{n\in\mathbb{N}}$ satisface la misma relación de recurrencia

$$E_{n+1}(z) = (z+a_n)E_n(z)+b_nE_{n-1}(z), \quad E_0(z) = \widehat{\mu}(z), \quad E_1(z) = \int_a^b \frac{P_1(t)}{z-t}d\mu(t).$$

De (2.11) se deduce que $\{Q_n\}_{n\in\mathbb{N}}$ también satisface dicha relación de recurrencia, pues es una combinación lineal de P_n y E_n; y $\{P_n, E_n\}$ forman un sistema fundamental de (2.12) (veáse el apéndice), ya que ambas sucesiones verifican la relación de recurrencia y son linealmente independientes, debido a que su correspondiente determinante de Casorati,

$$\begin{vmatrix} P_n(z) & E_n(z) \\ P_{n+1}(z) & E_{n+1}(z) \end{vmatrix} = \int_a^b \frac{P_n(z)P_{n+1}(t) - P_{n+1}(z)P_n(t)}{z-t} d\mu(t)$$

$$= (\prod_{j=1}^{n} b_j) \begin{vmatrix} 1 & \widehat{\mu}(z) \\ P_1(z) & \int_a^b \frac{P_1(t)}{z-t} d\mu(t) \end{vmatrix} = (\prod_{j=1}^{n} b_j)(-Q_1(z)) = (\prod_{j=1}^{n} b_j)(-c_0) \not\equiv 0,$$

es no nulo. Luego, $\{Q_n\}_{n\geq 1}$ satisface (2.12) siendo $Q_0 \equiv 0, Q_1 = c_0$.

Por el mismo argumento $\{P_n, Q_n\}$ es un sistema fundamental de (2.12). En efecto,

$$\begin{vmatrix} P_n(x) & Q_n(x) \\ P_{n+1}(x) & Q_{n+1}(x) \end{vmatrix} = \prod_{j=1}^{n}(b_j) \begin{vmatrix} 1 & 0 \\ x - \frac{c_1}{c_0} & c_0 \end{vmatrix} = (\prod_{j=1}^{n} b_j)c_0 = \Delta_n > 0. \tag{2.13}$$

De aquí

$$P_n(x)Q_{n+1}(x) - Q_n(x)P_{n+1}(x) = \Delta_n > 0, \;\; \forall x \in \mathbb{R}.$$

Si particularizamos en $x = x_{k,n}$, $x_{k,n}$ es el k-ésimo cero de P_n, $k = 1, \ldots, n$, obtenemos:

$$Q_n(x_{k,n})P_{n+1}(x_{k,n}) = -\Delta_n < 0, \;\; k = 1, \ldots, n.$$

De la proposición 2.9, esto es, del entrelazamiento de ceros de polinomios ortogonales consecutivos, sabemos que

$$sig(P_{n+1}(x_{k,n})) = (-1)^{n+1-k}, \;\; k = 1, \ldots n,$$

entonces,

$$sig(Q_n(x_{k,n})) = (-1)^{n-k}, \;\; k = 1, \ldots n,$$

Por consiguiente, entre dos ceros de P_n hay uno de Q_n y como este tiene grado $n-1$, se concluye que todos sus ceros son simples y se entrelazan con los de P_n. Queda así probado el siguiente teorema.

Teorema 2.12. *Los ceros del numerador y denominador Padé se entrelazan.*

Finalmente, obtenemos el siguiente corolario.

Corolario 2.13. *Los residuos del aproximante de Padé son positivos.*

Demostración.
$$R_n(z) = \frac{Q_n(z)}{P_n(z)} = \sum_{j=1}^{n} \frac{\lambda_{j,n}}{z - x_{j,n}},$$

donde $\lambda_{j,n} = \frac{Q_n(x_{j,n})}{P_n'(x_{j,n})} > 0$, ya que $sig(P_n'(x_{j,n})) = sig(Q_n(x_{j,n})),\ \ j = 1, 2, \ldots, n,$ por el entrelazamiento de sus ceros. □

2.3. Convergencia

Después de verificar que el aproximante de Padé coloca sus ceros y polos en el lugar correcto y de forma apropiada, estamos en condiciones de abordar el primer teorema de convergencia destacado de esta teoría, obtenido por Andréi A. Markov en 1895 (ver [3], [7]).

Teorema 2.14. *Sea μ una medida con soporte $[a, b]$ y $\widehat{\mu}$ la función de Markov asociada a μ. Entonces,*

$$\Pi_n(\widehat{\mu}) \overset{c.u.}{\rightrightarrows} \widehat{\mu} \qquad en \quad \widehat{\mathbb{C}}\backslash[a, b],$$

$\{\Pi_n\}_{n\in\mathbb{N}}$ converge uniformemente en cerrados de $\widehat{\mathbb{C}}\backslash[a, b]$.

Demostración. Vamos a demostrar que $R_n(z) = \Pi_n(\widehat{\mu})(z)$ está uniformemente acotada en cerrados de $\widehat{\mathbb{C}}\backslash[a, b]$, esto es, para cada cerrado $K \subset \widehat{\mathbb{C}}\backslash[a, b]$, existe $M_k > 0$ tal que $|R_n(z)| \leq M_k$, $\forall z \in K$, $\forall n \in \mathbb{N}$.

Sean $K \subset \widehat{\mathbb{C}}\backslash[a, b]$ un cerrado de $\widehat{\mathbb{C}}\backslash[a, b]$ y $\eta = d(K, [a, b])$, la distancia entre K y $[a, b]$. Sabemos que

$$R_n(z) = \sum_{j=1}^{n} \frac{\lambda_{j,n}}{z - x_j}, \quad \text{con} \sum_{j=1}^{n} \lambda_{j,n} = c_0, \quad \lambda_{j,n} > 0, \quad j = 1, 2, \ldots, n,$$

dado que $\lim_{z\to\infty} zR_n(z) = c_0$. Se sigue que,

$$|R_n(z)| \leq \sum_{j=1}^{n} \frac{\lambda_{j,n}}{|z - x_j|} \leq \frac{\sum_{j=1}^{n} \lambda_{j,n}}{\eta} = \frac{c_0}{\eta}, \quad \forall z \in K, \ \ n \in \mathbb{N}.$$

Por tanto, $\{R_n\}_{n\in\mathbb{N}}$ está uniformemente acotada en $\widehat{\mathbb{C}}\backslash[a, b]$. Por el teorema de Montel,ver [8], $\{R_n\}_{n\in\mathbb{N}}$ es relativamente compacta en la topología uniforme en cerrados de $\widehat{\mathbb{C}}\backslash[a, b]$. Consecuentemente, bastará con verificar que todas las subsucesiones convergentes lo hacen a la misma función, $\widehat{\mu}$.

Sea $\{R_{n_k}\}_{k\in\mathbb{N}}$ cualquier subsucesión convergente en cerrados de $\widehat{\mathbb{C}}\backslash[a,b]$, denotamos a su límite por g. Para todo $j\in\mathbb{N}$, a partir de un $n_k(j)$, el coeficiente de $\frac{1}{z^{j+1}}$ asociado a R_{n_k} es c_j, por ser un aproximante de Padé. Por lo tanto, los coeficientes del desarrolllo en el infinito de la función límite, g, coinciden con los de $\widehat{\mu}$ y, por unicidad del desarrollo de funciones analíticas, deben ser iguales, esto es, $g=\widehat{\mu}$. $\square$

Existe otra demostración del teorema de Markov que proporciona la velocidad de convergencia. En ella, son fundamentales las propiedades de la transformación conforme de $\widehat{\mathbb{C}}\backslash[a,b]$ en $\widehat{\mathbb{C}}\backslash\mathbb{D}$, donde $\mathbb{D}$ denota el disco unidad, cuya expresión es

$$\phi(z)=\frac{z-\frac{a+b}{2}+\sqrt{(z-a)(z-b)}}{\frac{b-a}{2}}\in\mathcal{H}(\widehat{\mathbb{C}}\backslash[a,b]).$$

Esta transformación, como es evidente de su expresión, verifica $\phi(\infty)=\infty$ y $\phi(z)=\frac{4}{b-a}z+f_0+f_1\frac{1}{z}+\mathcal{O}(\frac{1}{z^2})$. Además, entre sus propiedades destaca, por su relevancia geométrica y topológica, la transformación de elipses en círculos.

Denotamos por

$$C_\rho=\{z\in\widehat{\mathbb{C}}\backslash[a,b]/\ |\phi(z)|=\rho\},\quad \rho>1,$$

a la elipse cuya imagen por ϕ es el círculo de radio $\rho>1$.

Por otro lado, sabemos que

$$\widehat{\mu}(z)-\frac{Q_n(z)}{P_n(z)}=\mathcal{O}(\frac{1}{z^{2n+1}}),\quad z\longrightarrow\infty, \tag{2.14}$$

donde $\widehat{\mu}(z)-R_n(z)\in\mathcal{H}(\widehat{\mathbb{C}}\backslash[a,b])$. Dado que, $\widehat{\mu}(z)-R_n(z)$ está uniformemente acotada en cerrados de $\widehat{\mathbb{C}}\backslash[a,b]$, entonces, $\exists M_\rho\in\mathbb{R}^+$, tal que

$$\max_{z\in C_p}\left|\widehat{\mu}(z)-\frac{Q_n(z)}{P_n(z)}\right|\le M_\rho,\ \forall n\in\mathbb{N},\ \rho>1.$$

Luego, si multiplicamos por $\phi(z)^{2n+1}$, tenemos que $\phi(z)^{2n+1}(\widehat{\mu}(z)-\mathit{\Pi}_n(z))\in\mathcal{H}(\widehat{\mathbb{C}}\backslash[a,b])$. Por el teorema del módulo máximo generalizado, nótese que de (2.14) $\phi(z)^{2n+1}(\widehat{\mu}(z)-R_n(z))$ está uniformemente acotada en cerrados de $\widehat{\mathbb{C}}\backslash[a,b]$, se sigue que

$$|\phi(z)|^{2n+1}\left|\widehat{\mu}(z)-\frac{Q_n(z)}{P_n(z)}\right|\le M_\rho\rho^{2n+1},\quad \forall z\in\overline{Ext}(C_\rho),\quad \rho>1,$$

equivalentemente,

$$\left|\widehat{\mu}(z)-\frac{Q_n(z)}{P_n(z)}\right|\le\frac{M_\rho\rho^{2n+1}}{|\phi(z)|^{2n+1}},\quad \forall z\in\overline{Ext}(C_\rho),\quad \rho>1.$$

Tomando límite superior de la raíz 2n-ésima, resulta que

$$\overline{\lim_{n\to\infty}} \left|\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}\right|^{\frac{1}{2n}} \leq \frac{\rho}{|\phi(z)|}, \quad \forall z \in \overline{Ext}(C_\rho), \quad \rho > 1.$$

Haciendo $\rho \to 1^+$, se concluye que

$$\overline{\lim_{n\to\infty}} \left|\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}\right|^{\frac{1}{2n}} \leq \frac{1}{|\phi(z)|} < 1, \quad \forall z \in \widehat{\mathbb{C}}\backslash[a,b]. \tag{2.15}$$

Hemos obtenido de esta forma el resultado de convergencia cuantitativa ansiado. Sin embargo, cabe preguntarse si es posible afinar más o si incluso se puede conocer la velocidad exacta de convergencia. Observamos que la estimación de la convergencia está dada para cualquier medida. Esto nos invita a tratar de averiguar si existe alguna restricción sobre ella idónea para nuestro cometido. En efecto, veremos que si exigimos que μ' sea positiva en casi todo punto (c.t.p) podremos llegar a la velocidad exacta de convergencia.

Teorema 2.15. *Sea μ con soporte $[a,b]$ tal que $\mu' > 0$ c.t.p., y $\widehat{\mu}$, la función de Markov asociada a μ. Entonces,*

$$\Pi_n(\widehat{\mu}) \overset{c.u.}{\rightrightarrows} \widehat{\mu} \qquad en \quad \widehat{\mathbb{C}}\backslash[a,b].$$

Es más,

$$\lim_{n\to\infty} |\widehat{\mu}(z) - \Pi_n(z)|^{\frac{1}{2n}} = \frac{1}{|\phi(z)|}.$$

Demostración. Sea K un cerrado de $\widehat{\mathbb{C}}\backslash[a,b]$. Recordemos la expresión del error (2.10),

$$\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)} = \frac{1}{P_n^2(z)} \int_a^b \frac{P_n^2(x)}{z-x} d\mu(x), \quad z \in K.$$

Si denotamos por $d(K,[a,b])$ la distancia entre K y $[a,b]$ y por $D(K,[a,b])$ el diámetro de K y $[a,b]$, tenemos la siguiente cadena de desigualdades

$$\frac{1}{D(K,[a,b])} \leq \frac{P_n^2(z)}{||P_n||^2}\left|\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}\right| \leq \frac{1}{d(K,[a,b])}.$$

De esta forma, hemos obtenido una expresión del error en términos de P_n para cada n. Luego, por el teorema del "sandwich", aplicando límites inferiores y superiores adecuadamente, se obtiene que

$$\lim_{n\to\infty} \Big(\frac{P_n(z)}{||P_n||}\Big)^{\frac{1}{n}} \left|\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}\right|^{\frac{1}{2n}} = 1.$$

Observamos que el error cometido depende únicamente del comportamiento asintótico de $\frac{P_n}{||P_n||}$. Gracias al teorema de E.A.Rahmanov (ver [3]), $\mu' > 0$, sabemos que

$$\lim_{n\to\infty} \frac{|P_n(z)|^{\frac{1}{n}}}{||P_n||^{\frac{1}{n}}} = |\phi(z)| \text{ uniformemente en compactos de } \widehat{\mathbb{C}}\backslash[a,b],$$

y, por tanto, podemos concluir el resultado

$$\lim_{n\to\infty} \left|\widehat{\mu}(t) - \frac{Q_n(z)}{P_n(z)}\right|^{\frac{1}{2n}} = \frac{1}{|\phi(z)|}, \quad \forall z \in \widehat{\mathbb{C}}\backslash[a,b].$$

□

Finalizamos el capítulo mostrando la vinculación de la Aproximación Padé con la integración numérica, otro tópico muy importante en la Aproximación Constructiva.

2.4. Fórmulas de cuadratura de Gauss-Christoffel

Mediante un uso cauteloso de los argumentos utilizados hasta ahora, podemos seguir explotando las condiciones de interpolación (2.14). Si multiplicamos (2.14) por $P \in \mathbb{P}_{2n-1}$,

$$P(z)(\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}) = \mathcal{O}(\frac{1}{z^2}) \quad z \longrightarrow \infty, \quad \forall P \in \mathbb{P}_{2n-1},$$

por el teorema de Cauchy, al ser $P(z)(\widehat{\mu}(z) - \frac{Q_n(z)}{P_n(z)}) \in \mathcal{H}(\widehat{\mathbb{C}}\backslash[a,b])$ y anularse en el infinito, obtenemos

$$\frac{1}{2\pi i}\int_{\Gamma} P(t)\Big(\widehat{\mu}(t) - \frac{Q_n(t)}{P_n(t)}\Big)dt = 0, \quad \forall P \in \mathbb{P}_{2n-1},$$

siendo Γ una curva de Jordan con $[a,b]$ en su interior recorrida en sentido horario. Por lo tanto,

$$\frac{1}{2\pi i}\int_{\Gamma} P(t)\widehat{\mu}(t)dt = \frac{1}{2\pi i}\int_{\Gamma} P(t)\frac{Q_n(t)}{P_n(t)}dt, \quad \forall P \in \mathbb{P}_{2n-1}. \tag{2.16}$$

Si sustituimos la función de Markov por su expresión integral, aplicamos Fubini, dado que, $\frac{P(t)}{t-x} \in \mathcal{C}(\Gamma \text{x}[a,b])$ y P es entera, tras cambiar la orientación de Γ, resulta que

$$\frac{1}{2\pi i}\int_{\Gamma} P(t)\widehat{\mu}(t)dt = \frac{1}{2\pi i}\int_{\Gamma} P(t)\int_a^b \frac{d\mu(x)}{t-x}dt = \int_a^b \Big(\frac{1}{2\pi i}\int_{\Gamma}\frac{P(t)}{t-x}dt\Big)d\mu(x) =$$
$$= \int_a^b P(x)d\mu(x).$$

Por otra parte, desarrollando el miembro derecho de (2.16), por el mismo argumento, se sigue que

$$\frac{1}{2\pi i}\int_{\Gamma} P(t)\frac{Q_n(t)}{P_n(t)}dt = \sum_{j=1}^{n} \lambda_{j,n}\frac{1}{2\pi i}\int_{\Gamma}\frac{P(t)}{t-x_{j,n}}dt = \sum_{j=1}^{n}\lambda_{j,n}P(x_{j,n}).$$

Concluimos que

$$\int_a^b P(x)d\mu(x) = \sum_{j=1}^{n}\lambda_{j,n}P(x_{j,n}), \quad \forall P \in \mathbb{P}_{2n-1}.$$

Hemos conseguido una fórmula de cuadratura exacta en $\mathbb{P}_{2n-1}$, cuyos pesos son los residuos del aproximante Padé y sus nodos, los polos del aproximante de Padé, esto es, los ceros del enésimo polinomio ortogonal asociado a μ. Hemos recuperado la conocida fórmula de cuadratura de Gauss-Christoffel asociada a μ.

Dada la positividad de los pesos de las fórmulas de cuadratura de Gauss-Christoffel obtenidas vamos a lograr una estimación del error en la aproximación que dependerá únicamente del momento c_0 y de la aproximación de las funciones continuas mediante polinomios (teorema de Weierstrass).

Teorema 2.16. *Sea $f \in \mathcal{C}([a,b])$ y $\{I_n(f)\}_{n\in\mathbb{N}}$ la sucesión de fórmulas de Cuadratura de Gauss-Christoffel asociada a μ.*

$$\lim_{n\to\infty} I_n(f) = \int_a^b f(x)d\mu(x).$$

Es más,

$$|I(f) - I_n(f)| \leq 2c_0\rho_{2n-1}(f),$$

donde $\rho_n(f) = \min_{P\in\mathbb{P}_n} ||f-P||_\infty$, siendo $||f||_\infty = \max_{x\in[a,b]} |f(x)|$.

Demostración. Teniendo en cuenta que el grado de precisión de las fórmulas de cuadratura de Gauss-Christoffel es $2n-1$, $I_n(P) = I(P)$, $\forall P \in \mathbb{P}_{2n-1}$, sumando y restando esta misma cantidad,

$$\begin{aligned} I(f) - &I_n(f) = I(f) - I(P) + I_n(P) - I_n(f) = \textstyle\int_a^b (f(x)-P(x))d\mu(x) \\ +& \textstyle\sum_{j=1}^{n} \lambda_{j,n}(P(x_j) - f(x_j)), \quad \forall P \in \mathbb{P}_{2n-1}. \end{aligned}$$

Tomando valores absolutos y acotando cada término por su valor máximo, obtenemos

$$\begin{aligned} |I(f) - &I_n(f)| \leq \textstyle\int_a^b |f(x)-P(x)|d\mu(x) + \sum_{j=1}^{n}\lambda_{j,n}|f(x_j)-P(x_j)| \\ \leq& \; ||f-P||_\infty c_0 + ||f-P||_\infty c_0 = 2c_0||f-P||_\infty, \; \forall P \in \mathbb{P}_{2n-1}. \end{aligned}$$

Queda demostrada la segunda parte del teorema, la cota superior de la velocidad de convergencia. La convergencia es consecuencia del teorema de Weierstrass. Al ser f continua en $[a, b]$ se sigue que

$$\lim_{n\to\infty} \rho_n(f) = 0,$$

y, por tanto, se concluye que

$$\lim_{n\to\infty} |I_n(f) - I(f)| = 0.$$

□

El siguiente resultado muestra que, cuanto mayor sea la regularidad de f, mayor será la velocidad de convergencia de dichas fórmulas de cuadratura, ver [9].

Teorema 2.17 (Teorema de Jackson). *Si* $f \in \mathcal{C}^k([a, b])$, $k \in \mathbb{N}$, *entonces*

$$\rho_n(f) \leq \left(\frac{\pi}{2}\right)^k \frac{||f^{k)}||_\infty}{(n+1)n\cdots(n-k+2)}.$$

Corolario 2.18. *Si* $f \in \mathcal{C}^k([a, b])$, *entonces existe* $M > 0$ *tal que*

$$\lim_{n\to\infty} |I_n(f) - I(f)| \leq \frac{M}{n^k}.$$

En este capítulo hemos desvelado la estrecha conexión entre los aproximantes Padé y las fórmulas de cuadratura, emanada de la ortogonalidad de su denominador cuando la función a aproximar es de Markov. Llegados a este punto nos planteamos qué ocurre cuando se interpolan varias funciones simultáneamente y si es posible encontrar una teoría ortogonal tan bondadosa como lo ha sido la teoría de polinomios ortogonales que nos lleve a obtener resultados análogos. Su análisis nos ocupará en el próximo capítulo.

3

Aproximación Hermite-Padé y ortogonalidad múltiple. Fórmulas de cuadratura simultáneas.

Hemos analizado la eficiencia de la aproximación Padé para aproximar funciones de Markov. En el resto de la memoria vamos a ilustrar que también es bondadosa a la hora de aproximar simultáneamente varias funciones de Markov. Para ello, introduciremos la aproximación Hermite-Padé, la cuál dependiendo como sean las condiciones de interpolación, se apellidará tipo I o tipo II. Asimismo, razonando como en la aproximación Padé, destaparemos a la par la relación de los aproximantes Hermite-Padé tipo II con la ortogonalidad múltiple y las fórmulas de cuadraturas simultáneas.

3.1. Aproximación Hermite-Padé

En el capítulo 1 se evidenció que, para tener ciertas garantías de éxito en el estudio de la convergencia de la diagonal de la aproximación Padé, debíamos considerar el problema de interpolación racional en el infinito. Esta situación, como es de esperar, por ser la aproximación Hermite-Padé una extensión de la aproximación Padé, se repite. Por ello, introduciremos directamente la aproximación Hermite-Padé en el infinito.

Sea $f = (f_1, \ldots, f_r)$ un vector de r funciones holomorfas en el infinito, cuyos respectivos desarrollos en series en el infinito vienen dados por

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

Definimos el multi-índice $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$, siendo $|\vec{n}| = n_1 + n_2 + \cdots + n_r$ su tamaño. Por simplicidad, denotaremos a $|\vec{n}|$ por n. Nos planteamos dos problemas de interpolación.

En primer lugar, el problema de aproximación Hermite-Padé tipo I consiste en encontrar un vector de polinomios $(A_{\vec{n},1}, A_{\vec{n},2}, \ldots, A_{\vec{n},r})$ no simultáneamente todos nulos y un polinomio $B_{\vec{n}}$ satisfaciendo:

$$i)\ A_{\vec{n},j} \in \mathbb{P}_{n_j-1},\ j = 1,2,\ldots,r,\ \text{ y }\ B_{\vec{n}} \in \mathbb{P}_n.$$

$$ii) \sum_{j=1}^{r} A_{\vec{n},j}(z) f_j(z) - B_{\vec{n}}(z) = \mathcal{O}(\frac{1}{z^n}), \quad z \longrightarrow \infty. \tag{3.1}$$

Para hallar $A_{\vec{n},j}, j = 1,2,\ldots,r$, basta resolver el sistema lineal homogéneo de $n-1$ ecuaciones y n incógnitas, los coeficientes de $\{A_{\vec{n},j}\}_{j=1}^{r}$, resultante de igualar a cero los coeficientes de las potencias $\frac{1}{z^i}, i = 1,2,\ldots,n-1$, de $\sum_{j=1}^{r} A_{\vec{n},j} f_j$. Es evidente que existe solución distinta de la trivial. Asimismo, sabemos que $B_{\vec{n}}$ es la parte principal del desarrollo en serie de Laurent de $\sum_{j=1}^{r} A_{\vec{n},j} f_j$ en el infinito. Si denotamos $A_{\vec{n},j}(z) = a_{0,j} + a_{1,j} z + \cdots + a_{n_j-1,j} z^{n_j-1}$, $j = 1,2,\ldots,r$, el sistema a resolver es

$$(F_1^1 F_2^1 \cdots F_r^1) \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_r \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{3.2}$$

siendo

$$F_j^1 = \begin{pmatrix} f_{0,j} & f_{1,j} & \cdots & f_{n_j-1,j} \\ f_{1,j} & f_{2,j} & \cdots & f_{n_j,j} \\ \vdots & \vdots & & \vdots \\ f_{n-2,j} & f_{n-1,j} & \cdots & f_{n+n_j-3,j} \end{pmatrix}, a_j = \begin{pmatrix} a_{0,j} \\ a_{1,j} \\ \vdots \\ a_{n_j-1,j} \end{pmatrix}, \ j = 1,2,\ldots,r.$$

Óbservese que si $r = 1$ obtenemos el clásico aproximante Padé, y por tanto hay unicidad. Sin embargo, si $r \geq 2$, no hay en general unicidad.

En cambio, si exigimos que el grado de cada polinomio $A_{\vec{n},j}$, $j = 1,2,\ldots,r$, sea el menor posible, entonces hay unicidad. En efecto, si suponemos que hay al menos dos soluciones, tomando dos de ellas, necesariamente $A^1_{\vec{n},i} \neq A^2_{\vec{n},i}$ para algún $i \in \{1,2,\ldots,r\}$. Puesto que podemos asumir que dichos polinomios son mónicos, se sigue que el vector con la nueva componente $A^1_{\vec{n},i} - A^2_{\vec{n},i}$ en la coordenada i también es solución y $\partial(A^1_{\vec{n},i} - A^2_{\vec{n},i}) < \partial(A^1_{\vec{n},i}) = \partial(A^2_{\vec{n},i})$. Llegamos a un absurdo, ya que hemos encontrado una solución de menor grado.

Dado que podemos garantizar unicidad, tenemos la siguiente definición.

Definición 3.1. *Se denomina aproximante Hermite-Padé tipo I de multi-índice $\vec{n}$ para $(f_1, f_2, \ldots, f_r)$ al vector de polinomios $(A_{\vec{n},1}, A_{\vec{n},2}, \ldots, A_{\vec{n},r}, B_{\vec{n}})$ solución del problema de Aproximación Hermite-Padé tipo I.*

Nos va a interesar que los grados de los polinomios sean máximos.

Definición 3.2. *Un multi-índice $\vec{n}$ se dice normal para f respecto al problema de aproximación Hermite-Padé tipo I si $\partial(A_{\vec{n},j}) = n_j - 1$, $j = 1,2,\ldots,r$.*

Definición 3.3. *Un vector f se dice que es completo respecto al problema de aproximación Hermite-Padé I si todos los multi-índices $\vec{n}$ son normales para f.*

Por otra parte, el problema de aproximación Hermite-Padé tipo II consiste en encontrar un polinomio $P_{\vec{n}} \in \mathbb{P}_n$, $P_{\vec{n}} \not\equiv 0$, y otros r polinomios $Q_{\vec{n},j} \in \mathbb{P}_{n-1}$, $j = 1, 2, \ldots, r$, de forma que

$$P_{\vec{n}}(z) f_j(z) - Q_{\vec{n},j}(z) = \mathcal{O}(\frac{1}{z^{n_j+1}}), \;\; z \longrightarrow \infty, \;\; j = 1, 2, \ldots, r. \tag{3.3}$$

De (3.3) lo primero que observamos es que los coeficientes de $P_{\vec{n}}$ se obtienen resolviendo el sistema lineal homogéneo de n ecuaciones y $n+1$ incógnitas resultante de igualar a cero los coeficientes de las potencias $\frac{1}{z^i}, i = 1, 2, \ldots, n_j$ de $P_{\vec{n}} f_j$ para cada $j = 1, 2, \ldots, r$. Por tanto, existe solución no trivial. En segundo lugar, $Q_{\vec{n},j}$ es la parte principal del desarrollo en serie de Laurent en el infinito de $P_{\vec{n}} f_j$ para cada $j = 1, 2, \ldots, r$. Si denotamos por $P_{\vec{n}} = b_0 + b_1 z + \cdots + b_n z^n$, podemos expresar el sistema matricialmente como sigue

$$\begin{pmatrix} F_1^2 \\ F_2^2 \\ \vdots \\ F_r^2 \end{pmatrix} \begin{pmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \tag{3.4}$$

siendo

$$F_j^2 = \begin{pmatrix} f_{0,j} & f_{1,j} & \cdots & f_{n,j} \\ f_{1,j} & f_{2,j} & \cdots & f_{n+1,j} \\ \vdots & \vdots & & \vdots \\ f_{n_j-1,j} & f_{n_j,j} & \cdots & f_{n+n_j-1,j} \end{pmatrix}, \; j = 1, 2, \ldots, r.$$

Por un argumento similar al caso anterior habrá unicidad para el vector f si exigimos que $P_{\vec{n}}$ tenga el menor grado posible.

Definición 3.4. *Se denomina enésimo aproximante Hermite-Padé de tipo II asociado al vector funcional f al vector*

$$R_{\vec{n}}(f) = (R_{\vec{n},1}(f), R_{\vec{n},2}(f), \ldots, R_{\vec{n},r}(f)), \;\; \textit{siendo} \;\; R_{\vec{n},j}(f) = \frac{Q_{\vec{n},j}}{P_{\vec{n}}}, \; j = 1, 2, \ldots, r,$$

solución del problema de aproximación Hermite-Padé tipo II.

De nuevo, solo nos cautiva aquellos casos donde el grado de $P_{\vec{n}}$ sea n.

Definición 3.5. *Aquellos multi-índices donde el aproximante Hermite-Padé $R_{\vec{n}}$ asociado a f cumple $\partial(P_{\vec{n}}) = n$, se denominan normales para f respecto al problema de aproximación Hermite-Padé II.*

Definición 3.6. *Un vector f se dice que es completo respecto al problema de aproximación Hermite-Padé II si todos los multi-índices $\vec{n}$ son normales.*

En el anterior capítulo, se observó que las dificultades existentes para conseguir resultados generales de convergencia se podían superar limitando la clase de funciones a aproximar. Para el caso múltiple, no podía ser menos.

Nos centraremos en las funciones de Markov, esto es, consideraremos $\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_r)$ un vector de funciones de Markov, donde cada componente verifica $\widehat{\mu}_j \in \mathcal{H}(\widehat{\mathbb{C}} \backslash [a_j, b_j])$ y admite en el infinito el siguiente desarrollo

$$\widehat{\mu}_j(z) = \int_{a_j}^{b_j} \frac{d\mu_j(x)}{z-x} = \sum_{k=0}^{\infty} \frac{c_{k,j}}{z^{k+1}}, \ \ j = 1, 2, \ldots, r,$$

$$c_{k,j} = \int_{a_j}^{b_j} x^k d\mu_j(x), \ \ k \in \mathbb{N}.$$

Si particularizamos (3.4) a un sistema Markov, en la matriz de coeficientes podemos identificar bloques formados por matrices de Hankel. En particular, observamos que las n_j primeras columnas de cada F_j^2 forman una matriz de Gram para cada $j = 1, 2, \ldots, r$. Para ello, basta reescribir los momentos de la siguiente forma

$$c_{k,j} = \langle x_i, x_l \rangle_j = \int_{a_j}^{b_j} x^i x^l d\mu_j(x), \quad i + l = k, \quad k \in \mathbb{N}, \quad j = 1, 2, \ldots, r.$$

De modo que, reinterpretando las ecuaciones del sistema lineal homogéneo (3.4), obtenemos

$$\int_{a_j}^{b_j} Q(x) P_{\vec{n}}(x) d\mu_j(x) = 0, \ \ Q \in \mathbb{P}_{n_j - 1}, \ \ j = 1, 2, \ldots r. \tag{3.5}$$

Esto es, $P_{\vec{n}}$ satisface n condiciones de ortogonalidad repartidas de forma que n_i corresponden a la medida μ_i para cada i. $P_{\vec{n}}$ recibe el nombre de polinomio ortogonal múltiple asociado a $\vec{n}$ con respecto a μ.

De forma similar, si realizamos la misma lectura en las ecuaciones de (3.2), teniendo en cuenta la expresión dada para los momentos $c_{k,j}$, obtenemos

$$\sum_{j=1}^{r} \int_{a_j}^{b_j} x^k A_{\vec{n},j}(x) d\mu_j(x) = 0, \quad k = 0, 1 \ldots, n-2. \tag{3.6}$$

El vector formado por los polinomios que verifican estas condiciones de ortogonalidad se denomina polinomio ortogonal múltiple de tipo I.

Observación 3.7. Sea $(A_{\vec{n},1}, A_{\vec{n},2}, \ldots, A_{\vec{n},r})$ un vector de polinomios ortogonales múltiple tipo I. Si le añadimos la siguiente condición de normalización,

$$\sum_{j=1}^{r} \int_{\mathbb{R}} x^{|\vec{n}|-1} A_{\vec{n},j}(x) d\mu_j(x) = 1, \tag{3.7}$$

observamos que la matriz de coeficientes del sistema asociado (3.2) coincide con la matriz de coeficientes traspuesta del sistema correspondiente a polinomios mónicos ortogonales múltiples (3.4).

Consecuentemente, tenemos el siguiente teorema y otros importantes resultados manifestados en [12].

Teorema 3.8. *Un multi-índice $\vec{n}$ es normal respecto al tipo I si, y solo si, $\vec{n}$ es normal respecto al tipo II.*

Esta vinculación entre ambos tipos de ortogonalidad da pie a una simbiosis entre ambos, siendo su máxima expresión la obtención y el análisis de la relación de recurrencia a $r+2$ términos, que satisfacen. Tópico que se escapa de los objetivos de esta memoria, ver [11] y referencias contenidas en él. Muestra de ello son los siguientes resultados.

Corolario 3.9. *Si $\vec{n}$ es normal respecto al tipo II, entonces el polinomio $A_{\vec{n}+\vec{e}_j,j}$ tiene grado exactamente n_j, $j = 1, 2, \ldots, r$.*

Demostración. Si suponemos que existe algún j tal que $A_{\vec{n}+\vec{e}_j,j}$ tiene grado menor que n_j, entonces la matriz de coeficentes del sistema (3.2) resultante de eliminar la n_j-ésima columna de F_j^1 es regular, por ser $\vec{n}$ normal. Por lo tanto, la única solución sería $(A_{\vec{n}+\vec{e}_j,1}, \ldots, A_{\vec{n}+\vec{e}_j,r}) = (0, \ldots, 0)$, al tratarse de un sistema homogéneo. Llegamos a una contradicción, pues sabemos que hay solución no trivial. □

Teorema 3.10. *La siguiente biortogonalidad se da entre los polinomios ortogonales múltiples de ambos tipo:*

$$\sum_{j=1}^{r} \int_{a_j}^{b_j} P_{\vec{n}}(x) A_{\vec{m},j}(x) d\mu_j(x) = \begin{cases} 0, & \text{if } \vec{m} \leq \vec{n}, \\ 0, & \text{if } n \leq m-2, \\ 1, & \text{if } n = m-1, \end{cases}$$

donde $\vec{m} \leq \vec{n}$ denota $m_j \leq n_j$, $\forall j \in \{1, 2, \ldots, r\}$.

Demostración. Cuando $\vec{m} \leq \vec{n}$, de las condiciones de ortogonalidad que satisface el n-ésimo polinomio ortogonal múltiple (3.5) se sigue directamente que

$$\int_{a_j}^{b_j} P_{\vec{n}}(x) A_{\vec{m},j}(x) d\mu_j(x) = 0, \quad j = 1, 2, \ldots, r,$$

dado que $\partial(A_{\vec{m},j}) \leq n_j - 1$.

Por otra parte, de la ortogonalidad del tipo I (3.6) y su correspondiente normalización (3.7) se concluye el resultado para los casos $n \leq m-2$ y $n = m-1$, respectivamente. □

La expresión del error en toda aproximación no solo es imprescindible para su estimación, sino también para la búsqueda de resultados de convergencia. Análogamente al caso clásico, para cada j se tiene

$$P_{\vec{n}}(z)\widehat{\mu}_j(z) = \int_{a_j}^{b_j} \frac{P_{\vec{n}}(z)}{z-x} d\mu_j(x) = \int_{a_j}^{b_j} \frac{P_{\vec{n}}(z) - P_{\vec{n}}(x)}{z-x} d\mu_j(x) + \int_{a_j}^{b_j} \frac{P_{\vec{n}}(x)}{z-x} d\mu_j(x).$$

Por tanto,

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

donde

$$Q_{\vec{n},j}(z) = \int_{a_j}^{b_j} \frac{P_{\vec{n}}(z) - P_{\vec{n}}(x)}{z-x} d\mu_j(x) \in \mathbb{P}_{n-1},$$

es el enésimo polinomio de segunda especie asociado a μ_j. Como consecuencia de las condiciones de interpolación (3.3) se sigue que, para $j = 1, \ldots, r$,

$$P_{\vec{n}}(z)\widehat{\mu}_j(z) - Q_{\vec{n},j}(z) = \frac{1}{Q(z)} \int_{a_j}^{b_j} \frac{Q(x)P_{\vec{n}}(x)}{z-x} d\mu_j(x), \ \ \forall Q \in \mathbb{P}_{n_j}.$$

Así, concluimos que la expresión del error viene dada por

$$\widehat{\mu}_j(z) - \frac{Q_{\vec{n},j}(z)}{P_{\vec{n}}(z)} = \frac{1}{P_{\vec{n}}(z)Q(z)} \int_{a_j}^{b_j} \frac{Q(x)P_{\vec{n}}(x)}{z-x} d\mu_j(x), \ \ \forall Q \in \mathbb{P}_{n_j}. \tag{3.8}$$

Para lograr los ansiados resultados de convergencia resta ver si la teoría de polinomios ortogonales múltiples es tan benigna como lo fue la teoría de polinomios ortogonales para la aproximación Padé.

3.2. Polinomios ortogonales múltiples

Sean $\mu = (\mu_1, \mu_2, \ldots, \mu_r)$ un vector de r medidas tales que $\mu_j \in \mathcal{M}(\mathbb{R})$, y el multi-índice $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$, donde $|\vec{n}| = n_1 + n_2 + \cdots + n_r = n$.

Diremos que $P_{\vec{n}}$ es un polinomio ortogonal múltiple asociado a $\vec{n}$ con respecto a μ si verifica

$$\int_{\mathbb{R}} x^k P_{\vec{n}}(x) d\mu_j(x) = 0, \ \ k = 0, 1, \ldots, n_j - 1, \ \ j = 1, 2, \ldots, r. \tag{3.9}$$

Como ya hemos mencionado solo nos interesa cuando $\vec{n}$ es normal, es decir, cuando $P_{\vec{n}}$ tiene grado n. Observamos que si todas las medidas son iguales no hay unicidad y, por tanto, el grado de $P_{\vec{n}}$ se puede escoger menor que n. Así resulta evidente que si queremos obtener índices normales debemos restringir las medidas, es decir, pautar su interacción.

De este modo, nuestro punto de partida serán los sistemas Angelescos, caracterizados por el no solapamiento de los soportes de las respectivas medidas. Esto nos permite obtener resultados de cierta generalidad, por ejemplo: la localización de los ceros (simples) y con ello la posibilidad de utilizarlos como nodos en la teoría de integración numérica simultánea. Otros sistemas, como el caso de sistemas Nikishin o de sistemas algebraicos Chebyshev, han sido también ampliamente estudiados, los cuales presentan otras virtudes (ver [10]).

Definición 3.11. *Un vector Angelesco es un conjunto de medidas positivas* μ_i *con soportes* $[a_i, b_i] \subset \mathbb{R}$, $i = 1, 2, \ldots, r$ *"disjuntos" (se permite contacto en los extremos de los intervalos), es decir,* $(a_i, b_i) \cap (a_j, b_j) = \emptyset$, $i \neq j$ *.*

Sin pérdida de generalidad, podemos etiquetar los subíndices de las medidas de manera que $\mu_1, \mu_2, \ldots, \mu_r$ implique

$$-\infty < a_1 < b_1 \leq a_2 < b_2 \leq \ldots \leq a_r < b_r < \infty.$$

Diremos que un vector de Markov $\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_r)$, $\widehat{\mu}_j(z) = \int_{a_j}^{b_j} \frac{d\mu_j(x)}{z-x}$, para cada j, es de Angelesco si sus medidas forman un sistema Angelesco.

El no solapamiento de las medidas nos va a permitir contar cambios de signo como en el caso clásico, tal y como se muestra en el siguiente teorema.

Teorema 3.12. *Sea* $P_{\vec{n}}$ *el polinomio ortogonal múltiple mónico de tipo II de multi-índice* $\vec{n} = (n_1, n_2, \ldots, n_r)$ *para un sistema Angelesco* $(\mu_1, \mu_2, \ldots, \mu_r)$*, entonces* $P_{\vec{n}}$ *tiene* n_i *ceros en el intervalo* (a_i, b_i) *para cada* $i \in \{1, 2, \ldots, r\}$*.*

Demostración. En primer lugar, obsérvese que $P_{\vec{n}}$ tiene coeficientes reales. Entonces como $\int_{a_i}^{b_i} P_{\vec{n}}(t) d\mu_i(t) = 0$, $i = 1, 2, \ldots, r$, $P_{\vec{n}}$ tiene al menos un cambio de signo en $I_i = (a_i, b_i)$ para cada $i = 1, 2, \ldots, r$. Sea m_i el número de cambios de signo de $P_{\vec{n}}$ en I_i. Supongamos que $m_j < n_j$ para algún $j \in 1, 2, \ldots, r$. Definimos $Q_l(x) := (x - x_{1,j})(x - x_{2,j}) \cdots (x - x_{m_j,j}) \in \mathbb{P}_{n_j - 1}$, siendo $x_{1,j}, x_{2,j}, \ldots, x_{m_j,j}$ los cambios de signo de $P_{\vec{n}}$ en I_j, se tiene que $P_{\vec{n}}(t) Q_j(t)$ tiene signo constante en I_j, consecuentemente

$$\int_{a_j}^{b_j} P_{\vec{n}}(x) Q_j(x) d\mu_j(x) \neq 0.$$

Llegamos a una contradicción pues $P_{\vec{n}}$ es ortogonal a $\mathbb{P}_{n_j-1}$, en particular a Q_j, con respecto a μ_j. Por tanto, $m_i \geq n_i$ para cada $i = 1, 2, \ldots, r$. Además, puesto que todo cambio de signo es un cero y los intervalos $(a_1, b_1), (a_2, b_2), \ldots, (a_r, b_r)$ son disjuntos, se sigue que $P_{\vec{n}}$ tiene n_i ceros en cada I_i y, consecuentemente, $\partial(P_{\vec{n}}) = n$. □

Corolario 3.13. *Todo vector de Markov que sea de Angelesco es completo.*

Demostración. Es inmediato por la arbitrariedad de $\vec{n}$ en el teorema anterior. □

Gracias al teorema anterior, sin pérdida de generalidad, podemos factorizar $P_{\vec{n}}$ como

$$P_{\vec{n}}(x) = P_{\vec{n},1}(x)P_{\vec{n},2}(x)\cdots P_{\vec{n},r}(x),$$

donde $P_{\vec{n},j}$ es un polinomio de grado n_j cuyos ceros se encuentran en (a_j, b_j), $j = 1, 2, \ldots, r$. Por propia definición de $P_{\vec{n}}$ se sigue que

$$\int_{a_j}^{b_j} P_{\vec{n},j}(x)x^k d\mu_{j,n}^*(x) = 0, \quad k = 0, 1, \ldots, n_j - 1, \quad j = 1, 2, \ldots, r,$$

siendo $d\mu_{j,n}^*(x) = |P_{\vec{n}}^{(j)}(x)|d\mu_j(x) = \prod_{i>j}(-1)^{n_i}\prod_{i\neq j} P_{\vec{n},i}(x)d\mu_j(x)$. Dado que el producto no cambia de signo en (a_j, b_j), es evidente que $P_{\vec{n},j}$ es el polinomio ortogonal mónico de grado n_j respecto de la medida variante $\mu_{j,n}^*$.

Aplicando lo anterior, la expresión del error cobra para cada j un parecido significativo al caso clásico

$$\begin{aligned}\widehat{\mu}_j(z) - \frac{Q_{\vec{n},j}(z)}{P_{\vec{n}}(z)} &= \frac{1}{P_{\vec{n},j}^2(z)P_{\vec{n}}^{(j)}(z)}\int_{a_j}^{b_j} \frac{P_{\vec{n},j}^2(x)P_{\vec{n}}^{(j)}(x)}{z-x}d\mu_j(x) \\ &= \frac{1}{P_{\vec{n},j}^2(z)P_{\vec{n}}^{(j)}(z)}\int_{a_j}^{b_j} \frac{P_{\vec{n},j}^2(x)}{z-x}d\mu_{j,n}^*(x). \end{aligned} \tag{3.10}$$

Llegados a este punto, nos planteamos si el avance realizado es suficiente para probar el primer resultado esencial de convergencia. Echando la vista atrás, en el capítulo anterior la razón de nuestro éxito fue la positividad de los residuos del aproximante Padé. Surge así nuestro interés por conocer el signo de los residuos de los aproximantes Hermite-Padé de tipo II. Lo abordaremos haciendo uso de las condiciones de interpolación desde otro punto de vista.

Teorema 3.14. *Los residuos de los aproximantes Hermite-Padé de tipo II,*

$$R_{\vec{n},j}(z) = \sum_{k=1}^{n} \frac{\lambda_{k,n}^{(j)}}{z - x_{k,\vec{n}}},$$

tienen las siguientes propiedades:

I) $\lambda_{k,\vec{n}}^{(j)} > 0$ *cuando* $x_{k,\vec{n}} \in (a_j, b_j)$.
II) $\lambda_{k,\vec{n}}^{(j)}$ *tiene signo alterno cuando* $x_{k,\vec{n}} \in (a_i, b_i)$, $i \neq j$.

Demostración. Sea $j \in \{1, 2, \ldots, r\}$. Sabemos que

$$\widehat{\mu}_j(z) - R_{\vec{n},j}(z) = \mathcal{O}(\frac{1}{z^{n+n_j+1}}), \quad z \longrightarrow \infty. \tag{3.11}$$

Empleando la factorización $P_{\vec{n}} = P_{\vec{n},1} P_{\vec{n},2} \cdots P_{\vec{n},r}$, definimos

$$Q_l(x) = P_{\vec{n}}^{(j)}(x) \Big(\frac{P_{\vec{n},j}(x)}{x - x_{l,n_j}} \Big)^2, \quad l = 1, 2, \ldots, n_j.$$

siendo x_{l,n_j} un cero de $P_{\vec{n},j}$. Luego, de (3.11), se tiene

$$Q_l(z)(\widehat{\mu}_j(z) - R_{\vec{n},j}(z)) = \mathcal{O}(\frac{1}{z^3}), \quad z \longrightarrow \infty. \tag{3.12}$$

Manipulando el lado izquierdo de la anterior ecuación obtenemos por un lado,

$$Q_l(z)\widehat{\mu}_j(z) = \int_{a_j}^{b_j} \frac{Q_l(z) - Q_l(x)}{z - x} d\mu_j(x) + \int_{a_j}^{b_j} \frac{Q_l(x)}{z - x} d\mu_j(x), \tag{3.13}$$

y por otro lado,

$$Q_l(z) R_{\vec{n},j}(z) = \sum_{k=1}^{n} \frac{Q_l(z) - Q_l(x_{k,\vec{n}})}{z - x_{k,\vec{n}}} \lambda_{k,\vec{n}}^{(j)} + \sum_{k=1}^{n} \frac{Q_l(x_{k,\vec{n}})}{z - x_{k,\vec{n}}} \lambda_{k,\vec{n}}^{(j)}. \tag{3.14}$$

Nos percatamos de que, $\int_{a_j}^{b_j} \frac{Q_l(z)-Q_l(x)}{z-x} d\mu_j(x) - \sum_{k=1}^{n} \frac{Q_l(z)-Q_l(x_{k,\vec{n}})}{z-x_{k,\vec{n}}} \lambda_{k,\vec{n}}^{(j)}$ es un polinomio que en el infinito vale 0, basta con evaluar (3.12) en $z = \infty$.

Por lo tanto, dicho polinomio es idéticamente nulo. Restando (3.14)-(3.13)

$$Q_l(z)(\widehat{\mu}_j(z) - R_{\vec{n},j}(z)) - \int_{a_j}^{b_j} \frac{Q_l(x)}{z - x} d\mu_j(x) + \sum_{k=1}^{n} \frac{Q_l(x_{k,\vec{n}})}{z - x_{k,\vec{n}}} \lambda_{k,\vec{n}}^{(j)} = 0.$$

Multiplicando por z y evaluando en $z = \infty$ se sigue que

$$\int_{a_j}^{b_j} Q_l(x) d\mu_j(x) = \sum_{k=1}^{n} Q_l(x_{k,\vec{n}}) \lambda_{k,\vec{n}}^{(j)}.$$

Sustituyendo $Q_l(x)$ por $P_{\vec{n}}^{(j)}(x) \Big(\frac{P_{\vec{n},j}(x)}{x-x_{l,n_j}} \Big)^2$ y teniendo presente que Q_l se anula en todos los nodos salvo en x_{l,n_j} resulta

$$\int_{a_j}^{b_j} P_{\vec{n}}^{(j)}(x) \Big(\frac{P_{\vec{n},j}(x)}{x - x_{l,n_j}} \Big)^2 d\mu_j(x) = P_{\vec{n}}^{(j)}(x_{l,n_j}) (P'_{\vec{n},j}(x_{l,n_j}))^2 \lambda_{l,n_j}^{(j)}, \quad l = 1, 2, \ldots, n_j.$$

Entonces, dado que $P_{\vec{n}}^{(j)}(x)$ no cambia de signo en (a_j, b_j), se sigue que $\lambda_{l,n_j}^{(j)} > 0$, $l = 1, 2, \ldots, n_j$. Para estudiar el signo de $\lambda_{l,n_i}^{(j)}$, $l = 1, 2, \ldots, n_i$, $i = 1, 2, \ldots, r$, $i \neq j$, definimos

$$Q_l(x) = \prod_{s\neq i,j} P_{\vec{n},s}(x) \frac{P_{\vec{n},i}(x)}{(x - x_{l,n_i})} P^2_{\vec{n},j}(x),$$

utilizando el mismo argumento se llega, para $l = 1, 2, \ldots, n_i$, a

$$\lambda^{(j)}_{l,n_i} = \int_{a_j}^{b_j} \frac{P_{\vec{n},i}(x)}{(x - x_{l,n_i})\, P'_{\vec{n},i}\,(x_{l,n_i})} \left(\frac{P_{\vec{n},j}(x)}{P_{\vec{n},j}\,(x_{l,n_i})} \right)^2 \prod_{s\neq i,j} \frac{P_{\vec{n},s}(x)}{P_{\vec{n},s}\,(x_{l,n_i})} d\mu_j(x).$$

Dado que los factores del integrando poseen signo constante en el intervalo (a_j, b_j), independientemente de i, exceptuando $P'_{\vec{n},i}\,(x_{l,n_i})$ cuyo signo alterna según varíe el cero x_{l,n_i} tomado en (a_i, b_i), concluimos que estos pesos son distintos de cero y toman signos alternados. Además se puede verificar que los pesos correspondientes a los nodos más cercanos a cada extremo del intervalo (a_j, b_j) son positivos. □

Puesto que no todos los residuos son positivos, no podemos reproducir el resultado obtenido en el caso clásico. Sin embargo, si asumimos que el tamaño de los residuos negativos está controlado, sí podemos garantizar la convergencia de los aproximantes Hermite-Padé de tipo II a vectores Angelescos.

Teorema 3.15. *Sea* $\widehat{\mu} = (\widehat{\mu}_1, \widehat{\mu}_2, \ldots, \widehat{\mu}_r)$, $r \in \mathbb{N}$, *un vector Angelesco y* $\{R_n\}_{n\in\mathbb{N}}$ *una sucesión de aproximantes Hermite-Padé de tipo II tal que* R_n *está asociado al multi-ínidice* $\vec{n} = (n_1, n_2, \ldots, n_r) \in \mathbb{N}^r$, $n = |\vec{n}|$ *y* $\lim_{n\to\infty} n_i = \infty$. *Si* $\sum_{k=1}^{n} |\lambda^{(j)}_{k,n}| \leq M$, $\forall n \in \mathbb{N}$, $j = 1, 2, \ldots, r$, *siendo* $\{\lambda^{(j)}_{k,n}\}_{k=1}^{n}$ *los respectivos residuos de* $R_{n,j}$, *donde* $R_n = (R_{n,1}, R_{n,2}, \ldots, R_{n,r})$, *entonces*

$$R_n \overset{c.u.}{\rightrightarrows} \widehat{\mu} \qquad en \quad \widehat{\mathbb{C}}\backslash S,$$

siendo $S = \cup_{i=1}^{r} I_i$, $I_i = [a_i, b_i]$, $i = 1, 2, \ldots, r$.

Demostración. Puesto que nuestro objeto a aproximar es un vector de funciones finito dimensional, en el cual todas las normas son equivalentes, es suficiente con demostrarlo componente a componente. De este modo, veamos que para cada j

$$R_{n,j} \overset{\text{c.u.}}{\rightrightarrows} \widehat{\mu}_j \qquad \text{en} \quad \widehat{\mathbb{C}}\backslash S.$$

Dado $j \in \{1, 2, \ldots, r\}$, apreciamos que $\{R_{n,j}\}_{n\in\mathbb{N}}$ está uniformemente acotada en cerrados de $\widehat{\mathbb{C}}\backslash S$. Efectivamente, sean $K \subset \widehat{\mathbb{C}}\backslash S$ un cerrado y $\eta = d(K, S)$, la distancia de K a S, sabemos que

$$R_{n,j}(z) = \sum_{k=1}^{n} \frac{\lambda^{(j)}_{k,n}}{z - x_{k,n}},$$

entonces, por hipótesis,

$$|R_{n,j}(z)| \leq \sum_{k=1}^{n} \frac{|\lambda_{k,n}^{(j)}|}{|z - x_{k,n}|} \leq \frac{M}{\eta}, \quad \forall z \in K.$$

Luego, por el teorema de Montel es suficiente demostrar que todas las subsucesiones convergentes de $\{R_{n,j}\}_{n\in\mathbb{N}}$ tienden a una misma función. Usando los mismos argumentos que en el caso clásico, se prueba que el límite es únicamente $\widehat{\mu}_j$, ya que $\lim_{n\to\infty} n_j = \infty$. □

<u>Nota</u>. Señalar que la hipótesis implica que necesariamente la acotación uniforme de la suma de los valores absolutos de los residuos negativos está acotada, ya que $\sum_{k=1}^{n} \lambda_{k,n}^{(j)} = c_{0,j}$.

Por otro lado, es natural plantearse si también en esta situación existe una estimación de la velocidad de convergencia. La respuesta es sí. A pesar de no tener referencias hemos sugerido el siguiente resultado.

Teorema 3.16. *Bajo las mismas hipótesis del teorema anterior, se tiene que*

$$\overline{\lim_{n\to\infty}} |\mu_j(z) - R_{n,j}(z)|^{\frac{1}{n}} \leq \frac{1}{|\varphi(z)|^{\frac{1+\theta_j}{r}}},$$

siendo $\theta_j = \lim_{n\to\infty} \frac{n_j}{n}$ *y* $\varphi(z) = \prod_{k=1}^{r} \phi_k(z)$, *donde* ϕ_k *es la transformación conforme de* $\widehat{\mathbb{C}}\backslash I_k$ *en* $\widehat{\mathbb{C}}\backslash\mathbb{D}$.

Demostración. Consideramos la función $\varphi(z)$, la cual transforma $\widehat{\mathbb{C}}\backslash S$ en $\widehat{\mathbb{C}}\backslash\mathbb{D}$, definida como $\varphi(z) = \prod_{k=1}^{r} \phi_k(z)$, donde $\phi_k(z)$ es la transformación conforme de $\widehat{\mathbb{C}}\backslash I_k$ en $\widehat{\mathbb{C}}\backslash\mathbb{D}$. Definimos

$$C_\rho = \{z \in \widehat{\mathbb{C}}\backslash S / \ |\varphi(z)| = \rho\}, \quad \rho > 1.$$

Señalar que estas curvas de nivel en $\widehat{\mathbb{C}}\backslash S$ son curvas de Jordan o uniones de ellas. Concretamente, si consideramos $\rho > \rho_0$ para un cierto ρ_0 lo suficientemente grande, esta engloba a S; mientras que C_ρ rodea a todos y cada uno de los intervalos por separado cuando ρ está próximo a 1, ya que $|\phi_k(z)| > 1, \forall k = 1, 2, \ldots, r$. De hecho, cualquier compacto $K \subset \widehat{\mathbb{C}}\backslash S$ se encontrará en el exterior de una determinada C_ρ.

Sea $j \in \{1, 2, \ldots, r\}$. Dado que la familia $\{\mu_j - R_{n,j}\}_{n\in\mathbb{N}}$ está uniformemente acotada en compactos, se tiene que existe $M_\rho \in \mathbb{R}$ tal que

$$\max_{z\in C_\rho} |\widehat{\mu}_j(z) - R_{n,j}(z)| \leq M_\rho, \quad \forall n \in \mathbb{N}, \quad \forall j = 1, 2, \ldots, r.$$

Teniendo en cuenta las condiciones de interpolación

$$\widehat{\mu}_j(z) - R_{n,j}(z) = \mathcal{O}(\frac{1}{z^{n+n_j+1}}), \quad z \longrightarrow \infty,$$

y considerando la división euclidea, $\exists l_n \in \mathbb{N}$, $n + n_j + 1 = l_n r + s$, $0 \leq s < r$, $\varphi(z)^{l_n}(\widehat{\mu}_j(z) - R_{n,j}(z)) \in \mathcal{H}(\widehat{\mathbb{C}}\backslash S)$ y

$$\varphi(z)^{l_n}(\widehat{\mu}_j(z) - R_{n,j}(z)) = \mathcal{O}(\frac{1}{z^s}), \quad z \longrightarrow \infty.$$

Dado que la sucesión $\{\varphi^{l_n}(\mu_j - R_{n,j})\}_{n\in\mathbb{N}}$ está uniformemente acotada en cerrados de $\widehat{\mathbb{C}}\backslash S$, por el teorema del módulo máximo generalizado, resulta que

$$|\varphi(z)^{l_n}(\widehat{\mu}_j(z) - R_{n,j}(z))| \leq M_\rho \rho^{l_n}, \quad \rho > 1, \quad \forall z \in \overline{Ext}(C_\rho),$$

equivalentemente,

$$|\widehat{\mu}_j(z) - R_{n,j}(z)| \leq \frac{M_\rho \rho^{l_n}}{|\varphi(z)|^{l_n}}, \quad \rho > 1, \quad \forall z \in \overline{Ext}(C_\rho).$$

Tomando límite superior de la raíz n-ésima

$$\overline{\lim_{n\to\infty}} |\mu_j(z) - R_{n,j}(z)|^{\frac{1}{n}} \leq \frac{\rho^{\frac{1+\theta_j}{r}}}{|\varphi(z)|^{\frac{1+\theta_j}{r}}}, \quad \rho > 1, \quad \forall z \in \overline{Ext}(C_\rho).$$

Haciendo ρ tender a 1 por la derecha, concluimos que

$$\overline{\lim_{n\to\infty}} |\mu_j(z) - R_{n,j}(z)|^{\frac{1}{n}} \leq \frac{1}{|\varphi(z)|^{\frac{1+\theta_j}{r}}}, \quad \forall z \in \widehat{\mathbb{C}}\backslash S.$$

□

Obsérvese que si $r = 1$, obtenemos justamente la cota obtenida en el caso clásico. Para alcanzar la exactitud de la velocidad de convergencia se precisa de un uso sutil de la teoría potencial logarítmica en el caso vectorial (ver[11]), pues en el camino obtendríamos de (3.10) un “sandwich”del que no podemos extraer ninguna conclusión, ya que surge el serio “obstáculo”de que $P_{\vec{n},j}$ es un polinomio ortogonal con respecto a una medida variante $\mu^*_{j,n}$ y, por tanto, no sería válido aplicar el resultado del comportamiento asintótico de polinomios ortogonales obtenido por E.A.Rahmanov utilizado en la aproximación Padé, ya que las medidas interactuan.

3.3. Fórmulas de cuadratura simultáneas

En las aplicaciones surgen situaciones donde necesitamos aproximar simultáneamente varias integrales de una misma función, pero con respecto a diferentes medidas, por ejemplo, para calcular la intensidad de la luz de un determinado color de un haz constante policromático de espectro continuo, ver [13].

Atendiendo al capítulo anterior una posibilidad natural es utilizar las fórmulas de cuadratura de Gauss-Christoffel para aproximar cada integral. De este modo, si utilizamos n nodos para cada una de ellas necesitaríamos rn evaluaciones de la función, siendo el grado de exactitud de cada una de ellas $2n-1$. La eficiencia de este método, según la definición dada por Borges en [13], sería $2n/rn = 2/r,\ r \geq 2$. Por el contario, como veremos, usando fórmulas de cuadraturas interpolatorias con rn nodos y máximo grado de exactitud puede mejorarse su eficacia.

Surge inmediatamente de manera espontánea la siguiente pregunta, ¿puede dar lugar la aproximación Hermite-Padé tipo II asociada a un sistema Angelesco a fórmulas de cuadratura simultáneas con las propiedades deseadas? La respuesta es afirmativa, para ello basta observar la similitud entre (3.3) y (2.1). Así, resulta que multiplicando por $Q \in \mathbb{P}_{n+n_j-1}$ las condiciones de interpolación,

$$\widehat{\mu}_j(z) - \frac{Q_{\vec{n},j}(z)}{P_{\vec{n}}(z)} = \mathcal{O}(\frac{1}{z^{n+n_j+1}}), \quad j = 1, 2, \ldots, r,$$

se deduce para $j = 1, 2, \ldots, r$,

$$\int_{a_j}^{b_j} Q(x) d\mu_j(x) = \sum_{k=1}^{n} \lambda_{k,\vec{n}}^{(j)} Q_j(x_{k,\vec{n}}), \quad \forall Q \in \mathbb{P}_{n+n_j-1}, \tag{3.15}$$

donde $\lambda_{k,\vec{n}}^{(j)} = Res(R_{\vec{n},j}, x_{k,\vec{n}}) = \frac{Q_{\vec{n},j}(x_{k,\vec{n}})}{P'_{\vec{n}}(x_{k,\vec{n}})},\ k = 1, 2, \ldots, n$. Por lo tanto, tenemos las siguientes fórmulas de cuadratura simultáneas:

$$I_n^{(j)}(f) = \sum_{k=1}^{n} \lambda_{k,\vec{n}}^{(j)} f(x_{k,\vec{n}}), \quad j = 1, 2, \ldots, r. \tag{3.16}$$

Es evidente que la eficacia de las fórmulas de cuadratura ha aumentado notablemente, ya que si consideramos n nodos en cada fórmula de cuadratura (3.15) tenemos que la eficiencia es $n(r+1)/rn = r/(r+1) \approx 1$. No obstante, cabe señalarse que en el criterio definido por Borges las medidas estaban soportadas en el mismo intervalo, por tanto no es completamente correcto aplicarlo a un vector Angelesco. A pesar de ello, es obvio que hay un mejor aprovechamiento de los recursos disponibles mediante estas fórmulas de cuadratura simultáneas.

Observamos que en las fórmulas de cuadratura obtenidas aparecen ciertas “patologías” respecto a las clásicas.

Es de especial relevancia el uso de nodos fuera del intervalo de integración en las fórmulas obtenidas. Es más, si prestamos atención realmente se tiene

$$\int_{a_j}^{b_j} Q(x) d\mu_j(x) = I_{n_1}^{(j)}(Q) + \cdots + \mathbf{I_{n_j}^{(j)}}(Q) + \cdots + I_{n_r}^{(j)}(Q),$$

siendo $I_{n_s}^{(j)} = \sum_{k=1}^{n_s} \lambda_{k,n_s}^{(j)} Q(x_{k,n_s})$, $s = 1, 2, \ldots, r$, donde se tiene la misma proporción de pesos positivos y negativos, exceptuando $I_{n_j}^{(j)}$ cuyos pesos son todos positivos.

Considerando el polinomio $P_{\vec{n}}^{(j)}$ definido anteriormente, el cual posee los mismos ceros que el polinomio ortogonal exceptuando los que se encuentran en el intervalo (a_j, b_j), y $P(x) = p(x) \prod_{i>j} (-1)^{n_i} P_{\vec{n}}^{(j)}(x)$, $p \in \mathbb{P}_{2n_j-1}$, y teniendo en cuenta que $P \in \mathbb{P}_{n+n_j-1}$, se sigue que

$$\int_{a_j}^{b_j} P(x) d\mu_j(x) = \int_{a_j}^{b_j} p(x) d\mu_{j,n}^*(x) = \sum_{k=1}^{n_j} \lambda_{k,n_j}^{G_j} p(x_{k,n_j}), \quad j = 1, 2, \ldots, r,$$

siendo $\lambda_{k,n_j}^{G_j} = \lambda_{k,n_j}^{(j)} |P_{\vec{n}}^{(j)}(x_{k,n_j})|$, $k = 1, 2, \ldots, n_j$, y $d\mu_{j,n}^* = |P_{\vec{n}}^{(j)}(x)| d\mu_j$. Es decir, obtenemos la conocidas fórmulas de cuadratura de Gauss-Christoffel con respecto a medidas variantes.

Por tanto, está a nuestro alcance considerar simultáneamente r fórmulas de cuadratura de Gauss-Christoffel, aunque pagando el precio de la aparición de medidas variantes en ellas. Nuevamente, este es un incoveniente que probablemente pueda ser abordado mediante la teoría potencial logarítmica.

Por otra parte, como ya hemos visto en (3.14), no todos los pesos son positivos. Por ello, precisamos conocer el comportamiento de la suma de sus valores absolutos, pues esta restingirá la clase de funciones para las cuales las fórmula de cuadratura (3.16) convergen y la velocidad a la que lo hacen.

Teorema 3.17. *Sea $\{I_n^{(j)}(f)\}_{n\in\mathbb{N}}$ la sucesión de fórmulas de cuadratura definidas anteriormente (3.16). Si existe $d \in \mathbb{N}$ tal que*

$$\sum_{k=1}^{n} |\lambda_{k,\vec{n}}^{(j)}| \leq M n^d, \quad \forall n \in \mathbb{N}, \quad j = 1, 2, \ldots, r,$$

entonces

$$\lim_{n\to\infty} I_n^{(j)}(f) = \int_{a_j}^{b_j} f(x) d\mu_j(x), \quad j = 1, 2, \ldots, r,$$

para para toda función f derivable y con continuidad hasta el orden $d+1$ en S, es decir, $f \in \mathcal{C}^{d+1}(S)$. De hecho,

$$|I(f) - I_n^{(j)}(f)| \leq \rho_{n+n_j-1}(f)(c_{0,j} + M n^d), \quad j = 1, 2, \ldots, r,$$

siendo $\rho_m = \min_{P\in\mathbb{P}_m} ||f - P||_\infty$, $||f||_\infty = \max_{x\in S} |f(x)|$, $S = \cup_{i=1}^r [a_i, b_i]$.

Demostración. Trasladaremos al caso múltiple el desarrollo realizado en el clásico. Sabemos que $I_n^{(j)}$ es exacta en $\mathbb{P}_{n+n_j+1}$, entonces sumando y restando la cantidad $I_n^{(j)}(P) = I(P)$, $\forall P \in \mathbb{P}_{n+n_j-1}$, resulta que

$$I(f) - I_n^{(j)}(f) = I(f) - I(P) + I_n^{(j)}(P) - I_n^{(j)}(f) = \int_{a_j}^{b_j} (f(x) - P(x)) d\mu_j(x)$$
$$+ \sum_{k=1}^{n} \lambda_{k,\vec{n}}^{(j)} (P(x_{k,\vec{n}}) - f(x_{k,\vec{n}})), \forall P \in \mathbb{P}_{n+n_j-1} \quad j = 1, 2, \dots, r.$$

Tomando valores absolutos y acotando por los respectivos valores máximos, se sigue que

$$\begin{aligned}
|I(f) - I_n^{(j)}(f)| &\leq \int_{a_j}^{b_j} |f(x) - P(x)| d\mu_j(x) + \sum_{k=1}^{n} \lambda_{k,\vec{n}}^{(j)} |P(x_{k,\vec{n}}) - f(x_{k,\vec{n}})| \\
&\leq ||f - P||_\infty c_{0,j} + ||f - P||_\infty M n^d \leq ||f - P||_\infty (c_{0,j} + M n^d) \\
&\leq \tilde{M} n^d \rho_{n+n_j-1}(f), \quad j = 1, 2, \dots, r.
\end{aligned}$$

Como consecuencia del teorema de Jackson (2.17) se tiene la convergencia.

□

Conclusiones

Hemos realizado un estudio sistemático de la aproximación Padé para funciones de Markov. Al mismo tiempo, obtuvimos como aplicación las fórmulas de cuadratura de Gauss-Christoffel. También hemos analizado la aproximación racional de un vector de funciones de Markov, cuyas medidas forman un sistema Angelesco. En particular, hemos probado algunos resultados relevantes con el fin de obtener la convergencia e incluso una estimación de su velocidad. Finalizamos la memoria realizando un estudio de las fórmulas de cuadratura simultáneas obtenidas como aplicación de la aproximación Hermite-Padé tipo II. Como perspectiva de futuros trabajos planteamos las líneas de trabajo siguientes:

I) Fijada una clase de funciones, se plantea reducir las condiciones de interpolación a favor de un ahorro computacional. A esto se le conoce por *Aproximación tipo-Padé*.

II) En el caso múltiple, analizar la relación de recurrencia a $r + 2$ términos y su relación con las funciones de segunda especie de cada medida.

III) Establecer y estudiar la teoría potencial logarítmica vectorial para alcanzar velocidad exacta de convergencia en la aproximación Hermite-Padé.

A

Apéndice

A.1. Relaciones de recurrencia

Diremos que una sucesión $\{y_n\}_{n=0}^{\infty} \subset \mathbb{R}$ satisface una ecuación en diferencias de orden 2 o una relación de recurrencia a tres términos homogénea con datos iniciales y_0 e y_1, si y solo si,

$$y_{n+1} + a_n y_n + b_n y_{n-1} = 0, \quad \forall n \geq 1, \tag{A.1}$$

con y_0, y_1 conocidos.

Es inmediato comprobar, por construcción propia, que dados $\{a_n\}_{n\in\mathbb{N}}$ y $\{b_n\}_{n\in\mathbb{N}}$ para cualquier par $(y_0, y_1) \in \mathbb{R}^2$, va a existir una única sucesión $\{z_n\}_{n=0}^{\infty}$ que satisfaga (A.1). Así, para los datos $(1,0)$ y $(0,1)$ habrán solo dos sucesiones, $\{u_n\}_{n=0}^{\infty}$ y $\{v_n\}_{n=0}^{\infty}$ respectivamente, que cumplan la relación y cualquier otra solución de (A.1) $\{y_n\}_{n=0}^{\infty}$ verifica

$$y_n = y_0 u_n + y_1 v_n, n \geq 0$$

El conjunto $\{u_n, v_n\}$ recibe el nombre de sistema fundamental canónico asociado a (A.1).Es decir, se trata de una base del conjunto de soluciones $\{y_n\}_{n\in\mathbb{N}}$ y (y_0, y_1) son sus coordenadas.

Ahora bien, si $\{u_n, v_n\}$ es un sistema fundamental, también lo es $\{u'_n, v'_n\}$ donde:

$$u'_n = \eta u_n + \mu v_n$$
$$v'_n = \sigma u_n + \delta v_n$$

dónde η, μ, σ y δ son números reales tales que

$$\begin{vmatrix} \eta & \mu \\ \sigma & \delta \end{vmatrix} \neq 0$$

Observamos

$$\begin{pmatrix} u'_n \\ v'_n \end{pmatrix} = \begin{pmatrix} \eta & \mu \\ \sigma & \delta \end{pmatrix} \begin{pmatrix} u_n \\ v_n \end{pmatrix}.$$

entonces

$$\begin{pmatrix} u_n \\ v_n \end{pmatrix} = \begin{pmatrix} \eta & \mu \\ \sigma & \delta \end{pmatrix}^{-1} \begin{pmatrix} u'_n \\ v'_n \end{pmatrix}$$

Con respecto al mismo sistema fundamental podemos escribir

$$y_n = \begin{pmatrix} y_0 & y_1 \end{pmatrix} \begin{pmatrix} \eta & \mu_n \\ \sigma & \delta \end{pmatrix}^{-1} \begin{pmatrix} u'_n \\ v'_n \end{pmatrix} = y'_0 u'_n + y'_1 v'_n$$

siendo (y'_0, y'_1) las coordenadas respecto al nuevo sistema fundamental.

Teorema A.1. *Sea $\{u_n, v_n\}$ un sistema fundamental asociado a*

$$y_{n+1} + a_n y_n + b_n y_{n-1} = 0, \quad \forall n \geq 1.$$

Entonces,

$$W_n(u,v) = \begin{vmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{vmatrix} \textit{(determinante Casorati)} \tag{A.2}$$

satisface

$$W_n(u,v) = b_n W_{n-1}(u,v),$$

y por tanto,

$$W_n(u,v) = \prod_{j=1}^{n} b_j W_0(u,v) = \prod_{j=1}^{n} b_j \begin{vmatrix} u_0 & v_0 \\ u_1 & v_1 \end{vmatrix}$$

Demostración. Sabemos que

$$(1) \quad u_{n+1} + a_n u_n + b_n u_{n-1} = 0,$$

$$(2) \quad v_{n+1} + a_n v_n + b_n v_{n-1} = 0,$$

Si multiplicamos (1) por v_n y (2) por u_n

$$u_{n+1} v_n + a_n u_n v_n + b_n u_{n-1} v_n = 0$$

$$v_{n+1} u_n + a_n v_n u_n + b_n v_{n-1} u_n = 0$$

y restamos, concluimos que

$$W_n(u,v) - b_n W_{n-1}(u,v) = 0$$

□

Bibliografía

[1] Brezinski, Claude. *History of Continued Fractions and Pade Approximants*, Springer-Verlag, 1980. W. Rudin, Real and complex analysis, 3rd ed., McGraw-Hill, New York, 1987.

[2] C. Brezinski, *Outlines of Padé approximation*, in: H. Werner, et al. (Eds.), Computational Aspects of Complex Analysis, Reidel, Dordrecht, 1983, pp. 1–50.

[3] Nikishin, E. M. y Sorokin, V. N., *Rational approximations and orthogonality*, Transl. of Math. Monographs, A.M.S, Providence, Rhode Island, 1991.

[4] G. López Lagomasino. Constructive Theory of Approximation. *An Intoduction to Padé Approximation.* Coimbra Lecture Notes on Orthogonal Polynomials, Nova Science Pub., A. Branquinho and A. Foullquié Eds.,101-139, 2008.

[5] Baker, George A., Graves-Morris, Peter.*Padé Approximants Encyclopedia of mathematics and its applications.* Cambridge University Press, 1996.

[6] G. Szegö. *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ. Vol 23, Amer. Math. Soc. Providence, R.I. 1975

[7] G. López Lagomasino y H. Pijeira, *Polinomios Ortogonales*, XIV Escuela Venezolana de Matemáticas, Mérida, Venezuela, 2001

[8] W. Rudin, *Análisis real y complejo*, 3a ed., MacGraw-Hill, Madrid, 1988.

[9] CHENEY, E. W.. *Introduction to Approximation Theory*, McGraw-Hill, New York, 1966.

[10] U. Fidalgo, G. López Lagomasino. *Hermite Padé approximation and simultaneous quadrature formulas*, J. Approx. Theory, Vol. 126, 171-197, 2004.

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

[12] W. Van Assche, *Multiple Orthogonal Polynomials*, Mourad E. H. Ismail, Classical and Quantum Orthogonal Polynomials in One Variable (2005) , pp. 607 - 647.

[13] Borges, Carlos F., *On a Class of Gauss-Like Quadrature Rules*, Numerische Mathematik, Vol. 67 (1994), No. 3, 1994, pp 271-288.

Padé & Hermite-Padé Approximation

Teresa Padilla Miranda
Facultad de Ciencias • Sección de Matemáticas
Universidad de La Laguna
alu0101216713@ull.edu.es

Abstract

The limitations of Taylor polynomials in the approximation of meromorphic functions lead us to pose the Taylor interpolation problem for the class of rational functions. Padé approximants be the ones to solve this problem. Thus, first of all, we analyze the goodness of Padé approximation respect to Markov functions. For this, we look for a characterization of its denominator, which lead us to study the theory of orthogonal polynomials. Also, the close link between the Padé approximation and orthogonal polynomials give rise to Gauss-Christoffel quadrature formulas whose nodes and weights be extracted from the Padé approximant. Finally, considering the Hermite-Padé approximation we proceed in a similar way.

1. Introduction. Padé Approximation

We linearised Taylor interpolation problem for rational functions, i.e.,

$$a)\,\exists Q_{n,m}\in\mathbb{P}_n \quad y \quad \exists P_{n,m}\in\mathbb{P}_m,\quad P_{n,m}\not\equiv 0,$$
$$b)\,P_{n,m}(z)f(z)-Q_{n,m}(z)=\mathcal{O}(z^{n+m+1}),\quad z\longrightarrow 0.$$

Definition 1. *The rational function* $\Pi_n(f)=\frac{Q_n(f)}{P_n(f)}$ *that satisfy the above interpolation conditions is called **Padé Approximant**.*

The person responsible for naming these functions was **Henri Padé**, the first mathematician that study them systematically and arrange them in a double-entry table known as **Padé's Table**, which is shown below,

$$\begin{matrix} \Pi_{0,0} & \Pi_{1,0} & \Pi_{2,0} & \cdots \\ \Pi_{0,1} & \Pi_{1,1} & \Pi_{2,1} & \cdots \\ \Pi_{0,2} & \Pi_{1,2} & \Pi_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

In this table we highlight two kind of sequences:
i) $\{\Pi_{n,m}\}_{n\in\mathbb{N}}$, m fixed (rows).
ii) $\{\Pi_{n+l,n}\}_{n\in\mathbb{N}}$, $l\in\mathbb{Z}$ (diagonals).
If we want to approximate a meromorphic function with infinity singularities, we must move along a diagonal of Padé's table, but this has the inconvenient that poles grows with n, so we need a characterization of the denominator.
On the way to finding that characterization, we see that a first step is to consider the problem at infinity,

$$a')\,Q_n\in\mathbb{P}_{n-1} \text{ y } P_n\in\mathbb{P}_n,\ P_n\not\equiv 0.$$
$$b')\,P_n(z)f(z)-Q_n(z)=\mathcal{O}(\frac{1}{z^{n+1}}),\quad z\longrightarrow\infty.$$

Since this is not enough, we take another step to get convergences results without having to lost quality on them. So we restrict the class of functions to be approximated to the family formed by the **Markov functions**.

2. Padé Approximation and orthogonality. Quadrature formulas

Definition 2. *Let μ be a finite positive Borel measure with support $[a,b]$. Markov's function is given by:*

$$\widehat{\mu}(z)=\int_a^b\frac{d\mu(x)}{z-x}\in\mathcal{H}(\widehat{\mathbb{C}}\backslash[a,b]).$$

$\widehat{\mu}$ has a expansion in the neighbourhood of infinity,

$$\widehat{\mu}(z)=\sum_{k=0}^{\infty}\frac{c_k}{z^{k+1}},\quad c_k=\int_a^b x^k d\mu(x),\quad z\longrightarrow\infty.$$

*Reinterpreting the resulting interpolation conditions, it is obtained that $\mathbf{P_n}$ **is the** nth **orthogonal polynomial** with respect μ.*

We prove that zeros and poles of Padé approximant are interlaced in $[a,b]$. This let us get the desired convergence results.

Theorem 1. *(Markov) Let μ be a positive finite Borel measure with support $[a,b]$ and $\widehat{\mu}$ the Markov function associated. Then we have,*

$$\Pi_n(\widehat{\mu})\ \overset{\text{c.u.}}{\rightrightarrows}\ \widehat{\mu}\qquad \textit{in}\quad \widehat{\mathbb{C}}\backslash[a,b].$$

Furthermore, if we add the hypothesis $\mu'>0$ a.e., then

$$\lim_{n\to\infty}|\widehat{\mu}(z)-\Pi_n(z)|^{\frac{1}{2n}}=\frac{1}{\phi(z)},$$

where ϕ is the conformal transformation of $\widehat{\mathbb{C}}\backslash[a,b]$ on $\widehat{\mathbb{C}}\backslash\mathbb{D}$.
Finally, as application, we obtain Gauss-Christoffel Quadrature.

3. Hermite-Padé approximation

We consider simultaneous rational approximation to a vector of functions. In particular, Hermite-Padé approximation of type II. Fixed multi-index $\vec{n}=(n_1,n_2,\dots,n_r)\in\mathbb{N}^r$, $n=n_1+n_2+\dots+n_r$.
The rational vector $R_{\vec{n}}=(R_{\vec{n},1},R_{\vec{n},2},\dots,R_{\vec{n},r})$, $R_{\vec{n},j}=\frac{Q_{\vec{n},j}}{P_{\vec{n}}}, j=1,2,\dots,r$, satisfying the following interpolation conditions

$$P_{\vec{n}}(z)f_j(z)-Q_{\vec{n},j}(z)=\mathcal{O}(\frac{1}{z^{n_j+1}}),\ z\longrightarrow\infty,\ j=1,2,\dots,r,$$

is called Hermite-Padé approximant associated to $(f_1,f_2,\dots,f_r)$.
In a similar way to the classical case, we prove when vector of functions is an Angelesco system that $\mathbf{P}_{\vec{\mathbf{n}}}$ **is the multiple orthogonal polynomial**.
As a consequence of this, we have the next result.

Theorem 2. *If $\sum_{k=1}^n|\lambda_{k,n}^{(j)}|\le M, \forall n\in\mathbb{N}, j=1,2,\dots,r$, where $\{\lambda_{k,\vec{n}}^{(j)}\}_{k=1}^n$ are the respective residues of $R_{\vec{n},j}$, then*

$$R_{\vec{n}}\ \overset{\text{c.u.}}{\rightrightarrows}\ \widehat{\mu}\qquad \textit{en}\quad \widehat{\mathbb{C}}\backslash S,$$

where $S=\cup_{i=1}^r[a_i,b_i]$, $i=1,2,\dots,r$.

Finally, we obtain quadrature simultaneous formulas with maximum degree of accuracy,

$$\int_{a_j}^{b_j}Q(x)d\mu_j(x)=\sum_{k=1}^n\lambda_{k,\vec{n}}^{(j)}Q_j(x_{k,\vec{n}}),\quad \forall Q\in\mathbb{P}_{n+n_j-1},\ j=1,2,\dots,r.$$

References

[1] Nikishin, E. M. and Sorokin, V. N., *Rational approximations and orthogonality*, Transl. of Math. Monographs, A.M.S, Providence, Rhode Island, 1991.

[2] A. A. Gonchar, E. A. Rakhmanov, Lopez Guillermo, *"Some old and new results in rational approximation theory"*, Rational approximation and orthogonal polynomials, Zaragoza, 1988.