UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

Departamento de Física Teórica II

TESIS DOCTORAL

Polinomios biortogonales y sus generalizaciones: una perspectiva desde los sistemas integrables

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Gerardo Ariznabarreta García de Cortázar

Directores

Manuel Mañas Baena
Piergiulio Tempesta

Madrid, 2018

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS FÍSICAS

Departamento de Física teórica II

TESIS DOCTORAL:

Polinomios biortogonales y sus generalizaciones. Una perspectiva desde los sistemas integrables.

Memoria para optar al grado de doctor presentada por:
Gerardo Ariznabarreta García de Cortázar

Directores:
Manuel Mañas Baena
Piergiulio Tempesta

Como un ganso desplumado y escuálido,
me preguntaba a mi mismo con voz
indecisa si de todo lo que estaba leyendo
haría el menor uso alguna vez en la vida.
James Clerk Maxwell

Agradecimientos

Siete letras tiene mi nombre y siete letras tiene la palabra *gracias*. Tan sólo siete letras con las que recordar a todos esos que de un modo u otro han contribuido a tejer semejante destino, que parte de lo que está aquí faltaría de no haber sido por ellos. Con esto quiero decir que más de uno no será (muy inmerecidamente) aquí nombrado, pero puedo asegurar que cada capítulo, cada sección, cada frase, se podría lentamente destilar hasta encontrar la esencia y los matices heredados de aquellos que me han influenciado tanto como científico como persona. Desde luego siete letras parecen pocas, siquiera setenta veces siete bastarían, pero si te sintieras mínimamente responsable de que hoy me encuentre aquí, déjame darte las gracias de todo corazón.

De repente la gratitud me embiraga y siento que debiera dar las gracias a profesores de la carrera; como Miguel Escobedo Martínez (profesor atípico donde los haya, cuyas referencias a matemáticos rusos y clases semi-caóticas siempre recordaré cuando me falte la motivación). ¡Pero quizás debiera retroceder más en el tiempo para satisfacer esta necesidad que ahora tengo de agradecer! Entonces haría bien deteniéndome en el bachiller para recordar la voz firme, segura y rigurosa de Jose Ramón Díaz de Argote (el primer profesor a quien realmente admiré y cuyos conocimientos seguramente a día de hoy tenga que seguir envidiando). Me doy cuenta de que no puedo seguir este camino reverso del agradecimiento (olvidándome además injustamente de muchos) ya que viajando en el tiempo terminaría dando gracias incluso a la cultura helenística e islámica por sus primeros avances científicos. Por ello detengo aquí esta dinámica para centrarme únicamente en la no tan pequeña distancia cronológica marcada por la duración de mi periodo como doctorando.

Quiero dar las gracias de un modo un tanto gráfico: Tomemos el plano complejo (no es coincidencia) $\mathbb{C}$. Sobre el eje real situaré el grado de compromiso con el producto, esto es, la tesis. El eje imaginario medirá la implicación con la parte humana, es decir, con mi persona. Puesto que se trata de la sección de agradecimientos sólo me interesarán puntos del primer cuadrante. De este modo, a cada persona a quien deba mis agradecimientos le corresponderá un punto cuyo módulo pretende dar cuenta de la intensidad de su ayuda, esfuerzo y tiempo a mi otorgados y por lo tanto, proporcional a este ha de ser la profundidad de mi gratitud. Tratándonos todos de seres finitos acotaré dicho módulo por la unidad. El argumento del punto en cuestión da idea del interés que ha movido al agradecido, bien la persona, o bien el producto. En resumidas cuentas quiero dar un punto $z = re^{i\theta}$ con $0 \leq r \leq 1$ y $0 \leq \theta \leq \frac{\pi}{2}$ a los que me habéis acompañado a lo largo de este proceso:

- $z = 0$: Esto no sería propiamente un agradecimineto, tendría que hacer referencia a algo a lo que no le importe ni yo ni mi tesis pero que sin embargo haya estado ahí. Así que casi a modo ilustrativo dedicaré este punto al infatigable Seat Arosa con el que he repetido el trayecto Vitoria-Madrid-Vitoria hasta sumar unos bien estimados 100,000 km.

- $z = 1$: Este punto tiene que corresponder a un organismo con quien no haya interactuado como persona, sino únicamente como investigador. Así que he aquí mi agradecimiento a la Universidad Complutense de Madrid por haberme concedido una beca predoctoral.

A partir de aquí es donde se ponen los agradecimientos realmente interesantes:

- $z = ae^{i\theta}$ con $0 < a \leq 1$ y $0 \leq \theta \leq \frac{\pi}{4}$: Esta región va dedicada a aquellos directamente relacionados con la investigación, esos que han vivido o vivirán esta experiencia del doctorado y que me han mostrado su

empatía de un modo especial. Gracias a mis compañeros de despacho y de departamento, especialmente a Federico Finkel Morgenstern, por darme la oportunidad de probar el sabor de la docencia. Gracias también a Francisco Marcellán, Juan Carlos y Carlos Álvarez por todas esas tardes de investigación que hemos compartido. No puedo dejarme tampoco a Ángela, nuestra querida secretaria del departamento, gracias por haberme hecho reir cada vez que he tenido que hacer alguna fotocopia.

- $z = ae^{i\theta}$ con $0 < a \leq 1$ y $\frac{\pi}{4} \leq \theta \leq \frac{\pi}{2}$: Estos puntos son para esa gente cercana, amigos y conocidos que me han oido tantas veces hablar, bien con odio o bien con amor, de eso a lo que dedico gran parte de mi tiempo y que ellos llaman *hacer cuentas*. Gracias a los Templarios junto a los que, siempre entre magnesio y a base de arqueos, he conseguido desconectar de mis labores como investigador. Gracias también a la gente de Biciobsesión por las pedaladas de desfogue con las que tantas veces he conseguido dejar atrás cualquier forma de estrés.

- $z = e^{i\frac{\pi}{4}}$: Este lugar se lo han ganado indiscutiblemente mis directores, Manuel Mañas y Piergiulio Tempesta. Nadie como ellos se ha volcado tanto en mi experiencia personal sin dejar jamás de lado la calidad del producto que íbamos construyendo juntos. Desde luego ha sido un privilegio trabajar con vosotros así como dejarme asesorar por vuestra experiencia en tantas ocasiones. Recuerdo con claridad la charla que mantuve con Piergiulio y que hizo de detonante a la hora animarme a hacer el doctorado, muchas gracias por la motivación que me supiste transmitir. Manuel, a ti te debo tantas cosas...me quedo con todas las horas de sueño que te he robado, y también con la honestidad científica que siempre has demostrado.

- $z = 2i$: Sé que había acotado el módulo por la unidad, pero no hay cotación que baste para agradecer lo que mi familia merece. No importa si es práctico o no, lógico o absurdo, siempre están ahí y aunque no entiendan la razón de mis decisiones o acciones terminan por apoyarme. La posición del plano complejo en la que he situado este punto pone de manifiesto precisamente esto. Gracias a mis padres, mi hermano, mi chica, mis tíos y primos por no tener ni idea de qué es de lo que tratarán los capítulos que siguen, ni importaros en absoluto la calidad o precisión de los mismos. Gracias por preocuparos simplemente por mi.

Abstract

The existing connection between the theory of orthogonal polynomials and other branches of mathematics, physics and engineering is truly astonishing. There is no better proof of the usefulness of the theory than the recognition of its constant development and the wide generalizations that the original meaning of orthogonal polynomial has experienced since the dawn of the theory. The original concepts were generalized at the same time as the techniques for their study. Many of these new techniques were suggested by the new connections that kept appearing with different branches of mathematics. The approach that this thesis presents towards the study of the orthogonal polynomials is an example of such an interrelationship among disciplines, sharing tools and ideas with the theory of integrable systems.

A privileged role throughout this thesis will be played by the notion of semi infinite Gram matrices. These will be associated to a sesquilinear form suited to the kind of orthogonality under study. Additionally, some conditions will be imposed on the Gram matrix with the aim of guaranteeing the existence and uniqueness of the associated biorthogonal sequences. The following step consists of searching for any symmetry that the Gram matrix may have. There are two main reasons why such a task is worth the effort. In the first place, each found symmetry can be translated into a property of the biorthogonal sequences, for example: The Hankel structure of the matrix is equivalent to the well known three term recurrence relation satisfied by the standard orthogonal polynomials; the symmetry that the classical (Hermite, Laguerre, Jacobi) matrices possess induces the existence of the second order linear differential operator of which the classical orthogonal polynomials are solutions; etc. In the second place, the matrices that codify these kind of symmetries also help to surmise possible deformations of the problem, this is, they suggest *wise* perturbations of the Gram matrix. Whenever these deformations preserve the initial conditions that where imposed on the original Gram matrix, new biorthogonal sequences associated to the deformed case and related to the original ones will arise. If these perturbations are allowed to be modeled by a set of parameters, the resulting coefficients of the deformed biorthogonal polynomials will accordingly inherit this parametric or time evolution. It turns out that these time dependent coefficients are the solutions of differential equations that are well known in the theory of infinite dimensional integrable systems. This fact evidences the profound relationship between the theory of orthogonal polynomials and that of integrable systems and at the same time motivates and justifies the approach of this thesis.

The techniques that have been developed and that flesh out our results allow for the construction of adapted Gram matrices for each kind of biorthogonality and try to shed light both on their common properties and special symmetries. From this information the properties of its associated biorthogonal sequences are easy to extract. Additionally, the construction is ready for the corresponding deformations that allow to obtain new sequences from known ones. The method has been successfully applied to the following kinds of biorthogonality.

- In the real line: standard, matrix, multivariate, multiple and Sobolev biorthogonalities.
- In the unit circle: standard, matrix and multivariate biorthogonalities.

As can be observed there is an asymmetry in the number of cases that we have considered in each context. A study for the multiple and Sobolev cases in the unit circle is still missing. These two cases are certainly worth studying with the techniques mentioned above; we plan to devote special attention to them in the near future.

Resumen

La conexión existente entre los polinomios ortogonales y otras ramas de la matemática, la física o la ingeniería es verdaderamente asombrosa. Además, no hay mejor prueba de la utilidad de estos que el propio florecimiento, avance perpetuo y generalización en diversas direcciones de lo que se entendía por polinomio ortogonal en los albores de la teoría. Conforme el concepto se fue generalizando, también fueron evolucionando las técnicas para su estudio, algunas de estas claramente influenciadas por aquellas disciplinas matemáticas con las que iban surgiendo conexiones. La perspectiva que esta tesis adopta frente a los polinomios ortogonales es un ejemplo de este tipo de influencias, compartiendo herramientas y entrelazándose con la teoría de los sistemas integrables.

Una posición privilegiada en esta tesis la ocuparán las matrices de Gram semi infinitas; cada cual asociada a una forma sesquilineal adaptada al tipo de biortogonalidad en cuestión. A estas matrices se les impondrán una serie de condiciones cuyo objeto será el de garantizar la existencia y unicidad de las secuencias biortogonales asociadas a las mismas. El siguiente paso consistirá en buscar simetrías de estas matrices de Gram. Existen dos razones por las que este esfuerzo resulta ventajoso. En primer lugar, cada simetría encontrada podrá traducirse en propiedades de las secuencias biortogonales, por ejemplo: una estructura Hankel de la matriz es equivalente a gozar de la recurrencia a tres términos de los polinomios ortogonales; la simetría propia de las matrices asociadas a pesos clásicos (Hermite, Laguerre, Jacobi) implica la existencia del operador diferencial lineal de segundo orden de que los polinomios clásicos son solución; etc. En segundo lugar, las matrices que codifican este tipo de simetrías también sugerirán posibles deformaciones del problema, es decir, permitirán plantear perturbaciones bastante *sensatas* de la matriz de Gram. Cuando estas deformaciones preserven las condiciones inicialmente impuestas a la matriz de Gram de partida, será posible construir secuencias biortogonales desde el caso deformado y relacionar estas últimas con las originales. En caso de que dichas perturbaciones vengan modeladas por parámetros, se obtendrán secuencias biortogonales con coeficientes teniendo su correspondiente dependencia paramétrica. Resulta que dichos coeficientes son solución de ecuaciones diferenciales propias de la teoría de los sistemas integrables, quedando así patente el entrelazamiento entre las dos disciplinas matemáticas que como decíamos motiva el enfoque de esta tesis.

Las técnicas que hemos desarrlollado y que dan cuerpo a nuestros resultados permiten construir matrices de Gram adaptadas a cada tipo de biortogonalidad y tratan de poner de manifiesto tanto sus propiedades como sus simetrías, quedando estas prácticamente listas para ser transferidas a las secuencias biortogonales y para posteriormente ser deformadas con el ánimo de construir nuevas secuencias partiendo de unas conocidas. El método se ha aplicado con éxito a los siguientes tipos de biortogonalidad:

- En la recta real: biortogonalidad estándar, matricial, multivariable, múltiple y Sobolev.
- En la circunferencia unidad: biortogonalidad estándar, matricial y multivariable.

Como puede observarse existe una disimetría en cuanto a los casos tratados en la recta real y la circunferencia que dejan en desventaja numérica a esta última. Sería una buena idea aplicar los métodos aquí propuestos a los dos casos que restan por estudiar. Esto presenta por tanto, una posible línea de investigación a futuro.

Índice general

Capítulo 1

Introducción

Polinomios ortogonales y su conexión con los sistemas integrables

Hablar de un personaje y momento de la historia precisos a los que vincular los orígenes de los polinomios ortogonales en la recta real (OPRL) no es tarea sencilla. Una serie de resultados aislados pueden encontrarse en la obra *Recherches sur l'attraction des spheroides homogénes* de A. M. Legendre (1752-1833) o bien en los trabajos de de P. S. Laplace (1749-1827) respecto a la teoría de la probabilidad. Un primer paso hacia el estudio de los OPRL lo dio J. C. F. Gauss (1777-1855) [65] en sus consideraciones acerca de lo que hoy se conoce como cuadratura de Gauss-Jacobi. A pesar de ello, muchos expertos coinciden en asociar el nacimiento de los OPRL a C. G. J. Jacobi (1804-1851) quien observó la conexión entre los OPRL y las fracciones continuas [83]. El estudio de los OPRL como materia en sí misma lo comienza por un lado P. L. Tchebychev (1821-1894) (y su estudiante A. A. Markov (1856-1922)) en sus trabajos relacionados con la probabilidad y la aproximación por el método de mínimos cuadrados, y por otro T. J. Stieltjes (1856-1894) en sus consideraciones respecto a las fracciones continuas y el problema de momentos [116]. Un tratamiento sistemático de los OPRL, similar al que tenemos hoy en día, ha de esperar hasta la primera mitad del siglo XX esta vez de la mano de N. Abramesco [2],[3], J. Shohat (1886-1944) [112] y especialmente por los influyentes y detallados dieciséis capítulos de [119] por G. Szegő (1895-1985). Fue este mismo ilustre matemático quien, en los comienzos de su carrera profesional [118] (y en las numerosas publicaciones que siguieron a esta), motivado por una conjetura propuesta por G. Polya (1887-1985), se convirtió en el alma máter de una pequeña pero importante parte de la teoría general de los polinomios ortogonales: los polinomios ortogonales en la circunferencia unidad (OPUC). A parte de la tarea de G. Szegő en este contexto se deben mencionar los esfuerzos dedicados en esta misma dirección por la escuela matemática de Cracovia, con reputados representantes como Ya. L. Geronimus (1898-1984), N. I. Akhiezer (1901-1980) o M. G. Krein (1907-1989). Lamentablemente, todo el trabajo llevado a cabo por los anteriormente citados no obtuvo la atención merecida por la comunidad matemática de la época. Un ejemplo notorio de esta situación lo vivió S. Verblusnky (1906-1996) cuyos resultados [126] [127] tuvieron que ser redescubiertos posteriormente. No fue hasta mediados del siglo XX en que los OPUC obtuvieron el reconocimiento que merecían. El uso de polinomios ortogonales de Laurent[1] resulta especialmente útil y adaptado al caso de la circunferencia unidad, siendo la base propuesta por los matemáticos M. J. Cantero, L. Moral y L. Velázquez [37] la piedra angular para establecer la conexión entre los OPUC y los polinomios de Laurent ortogonales en la circunferencia unidad (OLPUC).

Con el tiempo, las técnicas de estudio de los OPRL y los OPUC y su comprensión han ido avanzando, lo que ha permitido una larga serie de generalizaciones de los conceptos iniciales; aquellas que se mencionan a lo largo de esta tesis son las siguientes:

[1]La idea de generalizar el concepto de ortogonalidad de polinomios a ortogonalidad de polinomios de Laurent fue inicialmente considerada para un problema de momentos en el caso de la recta real [86], [85] y generalizada posteriormente aún más, dando lugar al concepto de las funciones ortogonales [35].

- Biortogonalidad: Consiste en considerar dos secuencias de polinomios biortogonales entre ellas; el caso particular en que ambas coinciden es justo el caso ortogonal.
- Biortogonalidad matricial: Supone tomar los coeficientes de los polinomios, del anillo de las matrices de tamaño $n \times n$.
- Biortogonalidad en varias variables: Implica una dependencia multivariable de los polinomios.
- Biortogonalidad múltiple: Las condiciones de ortogonalidad en lugar de venir dadas por una única forma sesquilineal vendrán dadas por un número entero (no negativo) de estas.
- Biortogonalidad tipo Sobolev: Consiste en permitir, en las relaciones de ortogonalidad, la aparición de las derivadas de los polinomios.

Tanto los casos iniciales como sus generalizaciones tienen aplicaciones en diferentes ramas de la matemática como son las funciones trigonométricas, hipergeométricas, de Bessel, elípticas y especiales; los operadores de Jacobi, los problemas de momentos, la teoría de matrices aleatorias; fracciones continuas, aproximaciones de Padé, teoría de números, interpolación y cuadratura; electrostática, mecánica cuántica, matemática estadística, ecuaciones diferenciales e integrales, etc. Como no puede ser de otro modo, la teoría de los polinomios ortogonales también se ha beneficiado de aquellas aplicaciones donde estos han sido de utilidad; así, una serie de métodos propios de otras disciplinas han demostrado su idoneidad para tratar los polinomios ortogonales. Este ha sido precisamente el caso de nuestro enfoque del tema en cuestión desde los "sistemas integrables". En lo que sigue, el término sistemas integrables no se referirá al concepto de integrabilidad clásico (teorema de Arnold–Liouville, variables acción-ángulo) que se centra en sistemas que evolucionan en un espacio de fases de dimensión finita y por lo tanto propiamente "integrables", sino a aquellos sistemas ligados a la integrabilidad de ecuaciones no lineales en derivadas parciales y las soluciones "solitónicas".
Los orígenes de los solitones los podemos encontrar ya en el siglo XIX apareciendo estos bien como soluciones de problemas no lineales de ondas (KdV, [33] [89]), o bien en geometría diferencial (Sine–Gordon [32]). El término solitón, un paquete "solitario" que se propaga sin deformarse por un medio no lineal y por lo tanto con un comportamiento que se asemeja al de una partícula, lo acuñaron N. J. Zabusky y M. D. Kruskal [128] tras estudiar numéricamente las soluciones de la ecuación de KdV. Estos, mediante el uso de estas soluciones solitónicas fueron capaces de explicar los misteriosos resultados previamente obtenidos por E. Fermi, J. Pasta y S. Ulam [61] en sus estudios computacionales relativos a una cadena unidimensional de partículas con interacciones no lineales entre próximos vecinos. El paso definitivo en cuanto a dar la solución al problema de condiciones iniciales (de Cauchy) de la ecuación de KdV vino de la mano de C. S. Gardner, J. M. Green, M. D. Kruskal y R. M. Miura [64] con el descubrimiento del método de la transformada espectral inversa. El método toma como inspiración la ecuacion de KdV modificada y la transformación de Miura, desde donde infirieron la posibilidad de asociar a la ecuación de KdV un operador diferencial de Schrödinger para el que la solución, aún desconocida, de la ecuación de KdV juega el papel de potencial cuántico y los autovalores resultan constantes del movimiento de la misma. Su procedimiento para dar con la solución final puede resumirse en tres pasos: En primer lugar, resolver la ecuación de Schrödinger con el potencial cuántico dado en $t = 0$ (condiciones iniciales de la ecuación de KdV) con idea de obtener los datos de scattering iniciales. En segundo lugar, obtener la más "simple" (dada por la ecuación de KdV) evolución temporal de estos datos. En tercer y último lugar tomar las técnicas del problema de scattering inverso (Faddeev, Marchenko [59], [97]) para la ecuación de Schrödinger que permiten expresar el potencial cuántico una vez conocidos los datos de scattering. A este notable descubrimiento le siguieron una serie de avances que permitían deducir ecuaciones no lineales en derivadas parciales resolubles mediante el comentado método de la transformada espectral inversa: los pares de Lax para la ecuación de KdV [92], la representación de curvatura nula de Zakharov–Shabat para Schrödinger no lineal [129], y las técnicas propuestas por Ablowitz, Kaup, Newel Segur para la ecuación de Sine–Gordon [1]).
La descripción de las jerarquías integrables desde la teoría de los grupos de Lie se hizo esperar hasta que

llegaron los influyentes trabajos de M. Sato [108],[109] y las subsiguientes aportaciones de la escuela de Kyoto [49],[50],[51]. Poco después M. Mulase [103] conectó por fin los problemas de factorización, la técnica del revestimiento y la integrabilidad. M. Adler y P. van Moerbeke [4], [5], [6], [7], [9] no tardaron en darse cuenta de que tanto la factorización de Gauss–Borel como la técnica del revestimiento no sólo estaban presentes en la teoría de los sistemas integrables (Toda bidimensional o jerarquía KP discreta) sino también en el estudio de los polinomios ortogonales, lo que les permitió esclarecer la estrecha relación entre estos dos temas. En pocas palabras, esta conexión se resume en que los coeficientes de los polinomios ortogonales son solución de ecuaciones no lineales en derivadas parciales presentes en la teoría de los sistemas integrables. Podemos considerar este resultado como semilla del enfoque con que esta tesis se acerca a los polinomios ortogonales.

Metodología

A continuación se enumeran los cinco pasos que están presentes en cada uno de nuestros estudios de las secuencias de polinomios biortogonales.

1. Definición de una **forma sesquilineal** sobre $\mathbf{R}[z]$ (los polinomios con coeficientes en el anillo $\mathbf{R}$ y con variables tomando puntos de $\mathbb{C}^d$) mediante un funcional bivariado u_{z_1,z_2}.

   Definición 1. *Se denota por* $\langle *,*\rangle_u : \mathbf{R}[z_1]\times\mathbf{R}[z_2] \longrightarrow \mathbf{R}$ *a la forma sesquilineal asociada al funcional bivariado* u_{z_1,z_2}*.*

   La sesquilinealidad supone que, dados tres polinomios $p(z), q(z), r(z) \in \mathbf{R}[z]$ y coeficientes $A, B, \in \mathbf{R}$ se cumpla:

   $$\langle Ap(z_1) + Br(z_1), q(z_2)\rangle_u = A\langle p(z_1), q(z_2)\rangle_u + B\langle r(z_1), q(z_2)\rangle_u,$$
   $$\langle p(z_1), Aq(z_2) + Br(z_2)\rangle_u = \langle p(z_1), q(z_2)\rangle_u A^\dagger + \langle p(z_1), r(z_2)\rangle_u B^\dagger.$$

   Dado un un conjunto de medidas de Borel $\mu := \{\mu_{m,n}(z_1, z_2)\}$ cada cual con su soporte correspondiente $\Omega_{m,n}$ (conteniendo al menos uno de ellos un conjunto infinito de puntos) el tipo propuesto de forma sesquilineal va a incluir a aquellas del tipo:

   $$\langle p(z_1), q(z_2)\rangle_\mu = \sum_{0\leq n,m\leq\infty} \int_{\Omega_{m,n}} \frac{\partial^n p(z_1)}{\partial z_1^n} \mathrm{d}\mu_{m,n}(z_1, z_2) \left(\frac{\partial^n q(z_2)}{\partial z_2^n}\right)^\dagger .$$

2. Construcción de la **matriz de Gram** G. Se introduce un vector semi infinito de monomios adaptado al caso en cuestión $\chi(z)$ cuya componente j-ésima será denotada por $\chi_j(z) \in \mathbf{R}[z]$ y mediante el cual cualquier polinomio $p(z) \in \mathbf{R}[z]$ de grado k podrá escribirse del siguiente modo: $p(z) = \sum_{j=0}^k p_j\chi_j(z)$, con $p_j \in \mathbf{R}$. La relevancia de este vector de monomios reside en su utilidad a la hora de construir la matriz de Gram.

   Definición 2. *A continuación se define la matriz semi infinita de Gram* G *con entradas* $G_{i,j} \in \mathbf{R}$*:*

   $$G := \langle \chi(z_1), \chi(z_2)\rangle_u, \qquad G_{i,j} := \langle \chi_i(z_1), \chi_j(z_2)\rangle_u.$$

   Nótese cómo esta definición junto con la sesquilinealidad permiten, dados dos polinomios $p(z), q(z) \in \mathbf{R}[z]$ de grados k, l respectivamente, escribir $\langle p, q\rangle_u = \sum_{i=0}^k \sum_{j=0}^l p_i G_{i,j} q_j^\dagger \in \mathbf{R}$.

3. **Factorización LU** de G. Serán de relevancia únicamente aquellas formas sesquilineales cuyas matrices de Gram asociadas satisfagan unas condiciones concretas relativas a sus menores. Estas condiciones tienen por objeto garantizar que la matriz de Gram admita una única factorización LU (o de Gauss–Borel) *generalizada* desde donde construir la secuencia de polinomios biortogonales mónicos. Por factorización LU generalizada de la matriz de Gram entendemos una factorizacion $G := S_1^{-1} H S_2^{-\dagger}$ en la que H

es una matriz diagonal por bloques cuyos tamaños quedan fijados una vez impuestas las condiciones precisas anteriormente mencionadas relativas a los menores de G. Por su parte, S_1, S_2, son matrices unitriangulares inferiores con matrices identidad de los mismos tamaños que los de H a lo largo de su diagonal principal.

Definición 3. *Los vectores semi infinitos de polinomios mónicos se dan en términos de las matrices de la factorización:*

$$P_\alpha(z) := S_\alpha \chi(z) = \begin{pmatrix} P_{\alpha,0}(z) \\ P_{\alpha,1}(z) \\ \vdots \\ P_{\alpha,k}(z) \\ \vdots \end{pmatrix}, \qquad P_{\alpha,k}(z) := \sum_{k=0}^{j} (S_\alpha)_{k,j} \chi_j(z), \qquad \alpha = 1, 2.$$

Sus entradas son los polinomios mónicos biortogonales associados a la forma sesquilineal, es decir:

$$\langle P_{1,i}, P_{2,j} \rangle_u = \sum_{k=0}^{i} \sum_{r=0}^{j} (S_1)_{i,k} \langle \chi_k(z_1), \chi_r(z_2) \rangle_u (S_2^\dagger)_{r,j} = \left(S_1 G S_2^\dagger \right)_{i,j} = H_{i,j}.$$

4. La búsqueda de **simetrías de** G. La propia definición de G en términos del vector de monomios adaptado así como las características de la misma forma sesquilineal van a inducir algún tipo de simetría u orden en las entradas de G. Un ejemplo notorio de este tipo de simetrías son los casos de matrices de Gram de tipo Hankel o Toeplitz. Estas estructuras de G pueden ser traducidas a simetrías de la forma sesquilineal, y lo que es más importante, dotan a los polinomios biortogonales asociados de ciertas propiedades. Por poner un ejemplo, el caso Hankel al que antes nos referíamos dota a los polinomios asociados de su conocida ley de recurrencia a tres términos.

5. **Deformaciones** continuas y discretas de G. Existen una serie de transformaciones concretas y modeladas por parámetros que permiten deformar la matriz de Gram. Una manera de entender estas G deformadas es como matrices de Gram con una evolución dependiente de los parámetros involucrados en la transformación. Estas deformaciones están íntimamente relacionadas con las simetrías de G y son las responsables de la conexión existente entre la secuencia de polinomios biortogonales y las jerarquías integrables. Las deformaciones continuas enlazan con las ecuaciones de tipo Toda mientras que las discretas lo hacen con las transformaciones de Darboux.

Conclusiones y resultados

Como ocurre en las distintas disciplinas de la matemática, los polinomios ortogonales admiten una gran variedad de enfoques diferentes desde los que estudiarlos, cada cual con sus ventajas e inconvenientes. El punto de partida elegido para el desarrollo de esta tesis es la matriz de Gram, su factorizabilidad LU, sus simetrías y sus deformaciones. La motivación para elegir este enfoque la encontramos en la fuerte conexión existente entre las secuencias de polinomios biortogonales y los sistemas integrables de tipo Toda. A pesar de no ser la perspectiva más popular en la literatura al respecto, queda patente que las ideas que se derivan de la misma son lo suficientemente generales y potentes para ser aplicadas en un buen rango de ortogonalidades: estándar, matricial, múltiple, multivariada y Sobolev. Como es lógico, cada caso particular precisa de su necesaria serie de cuidados, propios de la idiosincrasia de cada situación, pero en rasgos generales, las ideas subyacentes son las mismas en cada una de las ortogonalidades sometidas a estudio. Tomando prestadas las siguientes palabras de G. H. Hardy [77]; "Los hombres humildes no hacen buenos trabajos. Es una de las principales tareas de un investigador exagerar un poco respecto a la importancia de su tema de estudio así como en cuanto a la relevancia de sus propias contribuciones al mismo", presentamos la siguiente lista de publicaciones que bien resumen nuestros esfuerzos:

1. G. Ariznabarreta and M. Mañas, Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems, *Adv. Math.* **264**, 396-463, (2014).
2. G. Ariznabarreta and M. Mañas, A Jacobi type Christoffel-Darboux formula for multiple orthogonal polynomials of mixed type *Linear Algebra and its Applicatios* **468**, 154-170, (2014).
3. G. Ariznabarreta and M. Mañas, Multivariate orthogonal polynomials and integrable systems, *Adv. Math.* **302**,628-739, (2016).
4. C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas and F. Marcellán, Christoffel Transformations for Matrix Orthogonal Polynomials in the Real line and the non-Abelian 2D Toda Lattice Hierarchy. *International Mathematics Research Notices.* **5**, 1285-1341, (2017).
5. G. Ariznabarreta, M. Mañas and P. Tempesta, Generalized Sobolev orthogonal polynomials, matrix moment problems and integrable systems, arXiv:1612.07229.
6. G. Ariznabarreta and M. Mañas, Darboux transformations for multivariate orthogonal polynomials, arXiv:1503.04786.
7. G. Ariznabarreta and M. Mañas, Linear spectral transformations for multivariate orthogonal polynomials and multispectral Toda Hierarchies, arXiv:1511.09129.
8. G. Ariznabarreta and M. Mañas, Multivariate orthogonal Laurent polynomials and integrable systems, arXiv:1506.08708.
9. C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas and F. Marcellán, Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line (with applications to the non-Abelian Toda lattice and nonconmmutative KP hierarchies), arXiv:1605.04617.
10. G. Ariznabarreta, M. Mañas and A. Toledano, CMV biorthogonal Laurent polynomials: Christoffel formulas for Christoffel and Geronimus perturbations, arXiv:1610.02008.
11. G. Ariznabarreta, M. Mañas and A. Toledano, CMV biorthogonal Laurent polynomials II: Christoffel formulas for Geronimus–Uvarov transformations arXiv:1611.03547.

Los artículos que se incluyen en esta tesis son los cinco que encabezan la lista. Los cuatro primeros forman ya parte de revistas indexadas mientras que el quinto aún espera su publicación. A continuación me dispongo a resumir las principales contribuciones que los mismos han aportado al campo así como su relación con el resto de artículos que completan la lista:

El objeto de la publicación número 1, es la ortogonalidad matricial en la circunferencia unidad empleando la base CMV (versus Szegő). Con el propósito de recuperar la recurrencia de Szegő desde el enfoque CMV introducimos la matriz de entrelazamiento; esta nos permite encontrar nuevas relaciones entre los núcleos de Christoffel–Darboux. Extendemos la versión no abeliana de la red de Toeplitz presentada en [36] mediante la incorporación de unos nuevos flujos parciales. Inspirados por las técnicas empleadas en el caso escalar consideramos deformaciones que llamamos discretas o elementales (de primer grado). Es necesario señalar que, dado que los polinomios matriciales no factorizan necesariamente como producto de factores lineales, estas transformaciones de primer grado no merecen el calificativo de elementales puesto que iteraciones de las mismas no nos permitirían construir cualquier polinomio. Esta observación pone de manifiesto que las deformaciones discretas matriciales en la circunferencia unidad y de grado arbitrario, aún esperan su estudio. A pesar de ello, puedo afirmar que tal análisis no debiera ser excesivamente complicado; esta afirmación la baso en los artículos 10 y 11 de la lista, en los que este tipo de deformaciones en la circunferencia unidad, pero en el caso escalar, son examinadas y en los artículos 4 y 9, en los que tratamos deformaciones matriciales pero en el caso real.

En el artículo número 2, en el contexto de la ortogonalidad múltiple de tipo mixto, aportamos, basándonos en la recurrencia (o en la correspondiente matriz de Jacobi) una fórmula de Christoffel-Darboux alternativa a la expuesta en [11] o [44] alcanzada desde el teorema ABC. La ventaja que tiene nuestra fórmula sobre la anterior es que únicamente involucra a la secuencia inicial de polinomios ortogonales. El inconveniente que presenta es que la fórmula contiene un número mayor de polinomios así como precisa del conocimiento previo de los coeficientes de la recurrencia.

En la contribución número 3, comenzamos reobteniendo una pequeña parte de los resultados contenidos en [53] desde la perspectiva de una factorización LU de la matriz de Gram que involucra bloques de tamaño creciente. Cabe citar nuestra propuesta para las funciones de segunda especie definidas como transformadas de Cauchy multivariables de los polinomios. Las deformaciones discretas y continuas que presentamos nos permiten escribir ecuaciones no lineales en derivadas parciales y diferencias cuya solución son matrices de tamaño creciente. Las deformaciones discretas en primer lugar motivan nuestra definición de las *cuasifunciones* tau, en términos de las cuales reescribimos tanto los polinomios como las funciones de segunda especie; y en segundo lugar, proporcionan una expresión en términos de *cuasideterminantes* para los polinomios transformados en función de los no transformados evaluados en un conjunto *equilibrado*[2] de nodos. Mientras que las ecuaciones de tipo Toda nos relacionan tres posiciones contiguas en la red, el uso de congruencias en el espacio de las matrices semi infinitas nos permite conseguir ecuaciones de tipo Kadomstev–Petviashvili que exclusivamente involucran una única posición en la red. La iteración de las transformaciones discretas de primer orden consideradas en 3 (donde las llamamos transformaciones elementales de Darboux y transformaciones adjuntas de Darboux) suponen una primera aproximación al problema general de las llamadas transformaciones espectrales lineales de la medida en las que se pretende hallar fórmulas que relacionen los elementos transformados con los originales. Dado que en el caso multivariable, los polinomios irreducibles pueden tomar cualquier grado (longitud para ser precisos), era necesaria una discusión más completa que la expuesta en 3; por este preciso motivo, escribimos los artículos 6 y 7 de la lista, en los que expresiones *á la Chrisfoffel* en términos de cuasideterminantes para los polinomios transformados espectral linealmente son finalmente formuladas generalizando así las expresiones correspondientes al caso unidimiensional. En el artículo 8 de la lista realizamos un examen similar del caso multivariable pero esta vez en el toro multidimensional.

En el escrito número 4 estudiamos las transformaciones matriciales de tipo Christoffel, es decir, partiendo de una medida matricial multiplicamos esta, bien por su derecha o bien por su izquierda, por un polinomio matricial de grado arbitrario y nos preocupamos por obtener una expresión que nos permita dar los polinomios asociados a la medida transformada en términos de aquellos asociados a la original. Este resultado que buscábamos lo conseguimos dar en términos de cuasidetermintantes en caso de que el polinomio perturbador tenga, por coeficiente director, una matriz no singular (únicamente en tal caso tenemos a nuestra disposición la información completa dada por las herramientas de la teoría espectral de los polinomios matriciales). También consideramos, desde otra perspectiva, algunos casos (relevantes en la literatura) con coeficiente director singular. Puesto que no imponemos ninguna restricción sobre el grado del polinomio perturbador, conseguimos extender resultados existentes (en los que sólo se tenían en cuenta iteraciones de perturbaciones de grado uno) en cuanto a la conexión con las ecuaciones propias de la jerarquía no abeliana de Toda bidimensional. Un artículo que complementa a la vez que extiende este último es el número 9 de la lista, en el que se considera la teoría de la transformación de las formas sesquilineales matriciales con total generalidad.

[2]Del inglés *poised.*

Finalmente, en el artículo número 5 se puede encontrar un estudio de la biortogonalidad de tipo Sobolev realizada, como el resto de nuestras publicaciones, desde el prisma de la factorización LU. Con tal propósito se introduce un nuevo objeto matemático que llamamos matriz de medidas con el que construir una forma bilineal de Sobolev bastante general que incluye el caso diagonal habitualmente estudiado. Interesados por las deformaciones de esta matriz (y por supuesto, por las secuencias biortogonales asociadas) comenzamos considerando sus perturbaciones aditivas. La técnica desarrollada nos permite no solo recuperar el fructífero concepto de los pares coherentes de medidas sino incluso generalizarlos de un modo que resulta natural desde la perspectiva propuesta. En segundo lugar, poniendo de manifiesto la libertad que da la integración por partes, introducimos la idea de considerar dentro de ciertas clases de equivalencia a las matrices de medidas. Esta idea es particularmente útil cuando las matrices de medidas están formadas a base de medidas clásicas, puesto que en tal caso, dentro de la clase de equivalencia de dicha matriz de medidas asociada a una forma bilineal de tipo Sobolev, uno puede encontrar, bajo ciertas condiciones, otra matriz de medidas cuya forma bilineal asociada no sea de tipo Sobolev, lo que permite una conexión directa entre secuencias ortogonales de tipo Sobolev y de tipo estándar. Posteriormente obtenemos la generalización de las transformaciones espectrales lineales en este contexto (concordando con los resultados generales que desde otra perspectiva obtuvimos en el artículo número 4), es más, introducimos un tipo nuevo de transformación involucrando ya no solo polinomios, sino también operadores diferenciales. Bien es cierto que únicamente son los casos mas sencillos de estas transformaciones los que estudiamos, dejando el caso general (que involucraría ecuaciones diferenciales para distribuciones) como tarea pendiente. Por último se presenta una breve conexión con los sistemas integrables.

Estructura y organización de la tesis

Esta tesis comienza con dos partes bien diferenciadas: la primera dedicada a la biortogonalidad ligada a la recta real y la segunda al mismo concepto pero en el contexto de la circunferencia unidad. Ambas partes tienen una estructura semejante que se divide en dos capítulos. En el primero contienen una parte introductoria que trata de motivar, a base de recuperar resultados clásicos de la teoría de los polinomios ortogonales, el uso las técnicas que hemos perfeccionado y de las que se nutren los resultados principales que componen la tesis. En el segundo incluyen una serie de pequeños resúmenes en cuanto a posibles generalizaciones de los resultados clásicos mencionados y en los que las técnicas anteriormente motivadas siguen siendo aplicables. Dichos resúmenes pretenden por un lado ser nexo de unión entre las dos partes introductorias y los artículos que se incluyen en la tercera y última parte de la tesis y por otro, animar a la lectura de aquellos artículos que aún no han sido publicados y no han sido aquí incluidos.

Parte I

Biortogonalidad en la recta real

1 Polinomios ortogonales en recta real

Como dijo Nicolás Copérnico: "Las matemáticas están escritas para matemáticos" [43], aun así entiendo que una exposición de notaciones y metodología claras en un caso sencillo o conocido debiera ser bien acogida por un lector dispuesto a considerar los casos más elaborados que se pretenden exponer como resultados de esta tesis. Por este motivo, en lugar de introducir la notación y herramientas principales en los casos que hemos estudiado, incluyo el presente capítulo introducctorio en el que se repasan resultados bien conocidos ([119] [63], [40], [104], [66], [82], [28]) pero siempre desde el enfoque de la factorización LU. Comenzaré por explicar el caso más sencillo que se da al trabajar con los polinomios $\mathbb{R}[x]$ y suponer la reducción al caso univariado del funcional. Concretamente nos interesaremos por el caso de funcionales lineales definidos positivos $\langle p(x_1), q(x_2)\rangle_u = L_u\left[p(x)q(x)\right]$. El teorema de la representación de Riesz nos permite hablar indistintamente, en este caso, de medidas definidas positivas con soporte contenido en la recta real. Me dispongo, por tanto, a dar los cinco pasos enumerados en el apartado *Metodología* del resumen introductorio de esta tesis. Definida la **forma sesquilineal** (en este caso funcional lineal) construiré la **matriz de Gram** (en este caso matriz de momentos) asociada; su **factorización LU** (en este caso Cholesky) me permitirá obtener la secuencia de polinomios biortogonales (en este caso ortogonales); las **simetrías** de G darán pie a formular propiedades generales de los OPRL (por ejemplo: la recurrencia a tres términos). Puesto que con estos ejemplos pretendo conseguir que el lector se sienta cómodo con nuestra notación, me detendré para considerar muy brevemente las tan conocidas e indudablemente importantes medidas clásicas. En quinto lugar nos preguntaremos acerca de las posibles **deformaciones** de G que nos pudieran relacionar nuevas secuencias de polinomios asociadas al caso deformado con las secuencias originales. El primer caso que nos ocupará consiste en perturbar el funcional lineal mediante una función racional. El segundo tipo de deformaciones vendrá dado por un conjunto infinito de parámetros, que podemos considerar como tiempos, respecto de los cuales dependerá la matriz de momentos deformada. Consideraremos esta dependencia de los parámetros como una evolución temporal de G con condición inicial dada por el caso no deformado. Serán este tipo de deformaciones las responsables de la conexión entre los OPRL y las ecuaciones de tipo Toda mediante un par de Lax.

1.1 Funcional definido positivo, matriz de momentos y factorización LU

Como adelantaba anteriormente, se puede introducir la estructura matemática sobre la que se sostendrá nuestro estudio bien mediante una medida de Borel definida positiva $\mathrm{d}\mu(x)$ con soporte $x \in \Omega \subseteq \mathbb{R}$ o equivalentemente mediante un funcional lineal continuo definido positivo $L_\mu : V \to \mathbb{R}$ sobre un espacio de funciones V. La equivalencia de estos dos enfoques se sigue del teorema de representación de Riesz–Markov–Kakutani, [107],[98],[87] que asegura la existencia de una representación integral involucrando una medida de Borel para cualquier funcional lineal continuo, es decir: $L_\mu[f] = \int_\Omega f(x)\mathrm{d}\mu(x)$. La idea es construir el espacio de funciones V como aquel formado por las funciones[1] reales de norma finita.

[1] Para ser precisos clases de equivalencia de funciones: $[f] = \{f_i \text{ con } ||f_i|| = ||f||\}$.

Definición 4. *La norma de una función se define como sigue,*

$$||f||^2 := L_\mu[f^2] = \int_\Omega f^2(x)\mathrm{d}\mu(x) < \infty.$$

Nótese cómo el carácter definido positivo del funcional lineal $L_\mu[f^2] > 0$ es imprescindible para que $||f||$ sea verdaderamente una norma. Dicha norma hace de V un espacio de Banach $L^2[\Omega]$. Como es habitual, la norma en $L^2[\Omega]$, induce un producto interno que nos permite calificar a nuestro espacio vectorial de espacio de Hilbert.

Definición 5. *El producto interno entre dos funciones $f, g \in L^2[\Omega]$ se escribe como sigue,*

$$\langle f, g\rangle_\mu := L_\mu[fg] = \int_\Omega f(x)g(x)\mathrm{d}\mu(x).$$

La desigualdad de Cauchy-Schwarz $|\langle f, g\rangle_\mu| \leqslant ||f||||g||$ asegura que este producto interno es siempre finito.

Previo a la construción de la matriz de momentos es imprescindible la siguiente definición:

Definición 6. *Sean el vector de monomios $\chi(x)$ y el vector auxiliar $\chi^*(x)$ los siguientes vectores semi infinitos:*

$$\chi(x) :=(1, x, x^2, x^3, \dots)^\top, \qquad \chi^*(x) :=\frac{1}{x}\chi\left(\frac{1}{x}\right).$$

Existe un resultado interesante cuando se multiplican ambos:

$$\chi(y)^\top \chi^*(x) = \chi^*(x)\chi(y) = \frac{1}{x-y}, \qquad \forall \ \ |y| < |x|.$$

Una vez hechas estas presentaciones se define por fin el objeto central de nuestro estudio, la matriz de Gram, que en este caso no es sino una matriz de momentos.

Definición 7. *La matriz de momentos es la matriz que resulta de la organización Hankel de los momentos del funcional lineal,*

$$G := L_\mu[\chi\chi^\top] = \langle \chi, \chi^\top\rangle_\mu = \int_\Omega \chi(x)d\mu\chi(x)^\top, \qquad G_{i,j} := L_\mu[x^{i+j}] = \langle x^i, x^j\rangle_\mu = \int_\Omega x^{i+j}d\mu.$$

Un problema clásico relacionado con esta última definición se encarga de la situación inversa, esto es, dada una secuencia infinita de números $\{m_k\}_{k=0}^\infty$, determinar en qué casos estos pueden considerarse como los momentos asociados a una medida definida positiva. No es mi intención detenerme en este tema, me conformaré con mencionar que dicho problema recibe diferentes nombres en función del soporte del funcional involucrado; el problema de momentos de Stieltjes, considerado inicialmente por este mismo en el marco de las funciones continuas [116] y también por Chebyshev [39], estudia el caso con soporte en la semirecta positiva. Por otro lado, si es la recta real completa la que hace de soporte, el problema se asocia al nombre de Hamburger [76]. Por último, el caso soportado sobre un intervalo cerrado de la recta real se denota por problema de momentos de Hausdorff [79].

En tanto que cualquier polinomio $f(x) = f_0 + f_1x + f_2x^2 + f_3x^3 + \dots$ puede expresarse como sigue:

$$f(x) = \boldsymbol{f}^\top\chi(x) \qquad \text{o bien} \qquad f(x) = \chi(x)^\top\boldsymbol{f}, \qquad \text{donde} \qquad \boldsymbol{f} = (f_0, f_1, f_2, \dots)^\top,$$

y empleando la siguiente notación para denotar las matrices truncadas al tamaño l:

$$A = \left(\begin{array}{c|c} A^{[l]} & A^{[l,\geq l]} \\ \hline A^{[\geq l,l]} & A^{[\geq l]} \end{array}\right), \qquad A^{[l]} := \begin{pmatrix} A_{0,0} & A_{0,1} & \dots & A_{0,l-1} \\ A_{1,0} & A_{1,1} & \dots & A_{1,l-1} \\ \vdots & & \ddots & \\ A_{l-1,0} & A_{l-1,0} & \dots & A_{l-1,l-1} \end{pmatrix},$$

es posible, dados dos polinomios $f(x), h(x)$ siendo k el grado más alto de los dos, expresar matricialmente su producto interno así como sus normas como sigue:

$$\langle f, h\rangle_\mu = \boldsymbol{f}^\top G^{[k]}\boldsymbol{h}, \qquad ||f||^2 = \boldsymbol{f}^\top G^{[k]}\boldsymbol{f}, \qquad ||g||^2 = \boldsymbol{g}^\top G^{[k]}\boldsymbol{g}.$$

Es importante destacar que el carácter definido positivo del funcional lineal L_μ hace de $\langle f, f\rangle_\mu$ una forma cuadrática definida positiva, lo que asegura que todos y cada uno de los autovalores de $G^{[k]}$ serán positivos. Puesto que esta afirmación ha de ser cierta sea cual fuere el valor de k, ha de cumplirse que $\det\left(G^{[k]}\right) > 0\ \forall k$. Esta es condición suficiente para que exista una única factorización LU de G. En cuanto a la correspondiente condición necesaria y suficiente, bastaría con pedir que el funcional tuviera una matriz de momentos satisfaciendo que $\det\left(G^{[k]}\right) \neq 0$. Este será el tipo de funcionales por los que nos interesaremos en las publicaciones incluidas, pero en estas líneas introductorias seguiremos trabajando con los funcionales definidos positivos. En el caso que nos ocupa, la factorización es de tipo Cholesky.

Definición 8. *La factorización de Cholesky de la matriz de momentos es:*

$$G = S^{-1}HS^{-\top}, \qquad H := \begin{pmatrix} h_0 & 0 & 0 & \dots \\ 0 & h_1 & 0 & \dots \\ 0 & 0 & h_2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \qquad S := \begin{pmatrix} 1 & 0 & 0 & \dots \\ S_{1,0} & 1 & 0 & \dots \\ S_{2,0} & S_{2,1} & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

La razón de considerar esta factorización no es otra que la de construir mediante las matrices involucradas, la secuencia de polinomios mónicos ortogonales.

Definición 9. *La secuencia mónica de polinomios ortogonales y sus normas se definen del siguiente modo:*

$$P(x) := S\chi(x) = \begin{pmatrix} P_0(x) \\ P_1(x) \\ \vdots \\ P_k(x) \\ \vdots \end{pmatrix}, \qquad h_k := ||P_k(x)||^2.$$

Proposición 1. *Dichos polinomios satisfacen las siguientes condiciones de ortogonalidad:*

$$\langle P_k, x^j\rangle_\mu = h_k\delta_{k,j}, \qquad \forall j \leq k.$$

Demostración. Reescribiendo la factorización LU del siguiente modo, $SG = HS^{-\top}$ y usando la definición $G = \langle \chi, \chi^\top\rangle_\mu$ se tiene:

$$SG = S\langle \chi, \chi^\top\rangle_\mu = \langle S\chi, \chi^\top\rangle_\mu = \langle P, \chi^\top\rangle_\mu = HS^{-\top}.$$

Tomando de la última igualdad las primeras $k+1$ componentes de la fila $k+1$-*ésima* se puede ver que:

$$\langle P_k, \left(1, x, x^2, \dots, x^{k-1}, x^k\right)\rangle_\mu = h_k\left(0, 0, \dots, 0, 1\right).$$

□

Nótese cómo las condiciones de ortogonalidad que satisfacen los polinomios mónicos suponen un sistema lineal de ecuaciones para sus coeficientes que hace de ellos la única solución posible al mismo. Esta observación es equivalente a la unicidad de la factorización LU de G. Es más, este método de expresar los polinomios ortogonales en términos de las matrices de la factorización no es otro que el proceso de ortogonalización de Gram–Schmidt codificado.

Existen diferentes representaciones para los polinomios ortogonales. Un primer ejemplo (que se deducirá próximamente) consiste en darlos en función de un cociente de determinantes de la matriz de momentos truncada:

$$P_k(x) = \frac{1}{\det\left[G^{[k]}\right]} \cdot \det \left[\begin{array}{cccc|c} & & & & 1 \\ & & G^{[k]} & & x \\ & & & & \vdots \\ & & & & x^{k-1} \\ \hline G_{k,0} & G_{k,1} & \dots & G_{k,k-1} & x^k \end{array}\right].$$

Desde este resultado se puede deducir, sin demasiada complicación, un segundo ejemplo conocido como la representación integral de Heine:

$$P_k(x) = \frac{1}{k! \det\left[G^{[k]}\right]} \int \prod_{j=1}^{k} (x - x_j) \prod_{1 \leq j < n \leq k} (x_n - x_j)^2 \mathrm{d}\mu(x_1) \mathrm{d}\mu(x_2) \dots \mathrm{d}\mu(x_k).$$

A pesar de todo esto, en lugar de determinantes propondré aquí una expresión empleando *cuasideterminantes*[2]. Como vamos a comprobar, en este caso particular escalar y univariado, las expresiones que involucran determinantes y las correspondientes en términos de cuasideterminantes son equivalentes. La razón que justifica optar por los cuasideterminantes reside en que a la hora de generalizar a otros tipos de ortogonalidad, estas seguirán siendo válidas allí donde sus homólogas determinantales carecerían de sentido. Dada la constante aparición de dicha operación matemática a lo largo de esta tesis, merece la pena detenerse un instante para introducirlos al menos desde su conexión con el problema de la factorización LU de una matriz dada. Para una exposición más completa del tema recomiendo la lectura de [68] y de [106]. Comencemos con una matriz $M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathbb{M}_{(n+m)}$ con $A \in \mathbb{M}_n$, $\det(A) \neq 0$ y $D \in \mathbb{M}_m$. En tal caso, el cuasideterminante respecto del último bloque se define en términos del resto de bloques como sigue:

$$\Theta_*[M] = \Theta_* \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := D - CA^{-1}B.$$

De entre todos los cuasideterminantes de una matriz, este es el más simple, que además coincide con el complemento de Schur respecto del bloque A y que se suele denotar por $SC(M) = M/A := \Theta_*[M]$. Como decíamos, es interesante conectar esta operación con la factorización LU por bloques de la matriz M:

$$M = \begin{pmatrix} \mathbb{I}_n & 0 \\ CA^{-1} & \mathbb{I}_m \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & \Theta_*[M] \end{pmatrix} \begin{pmatrix} \mathbb{I}_n & AB^{-1} \\ 0 & \mathbb{I}_m \end{pmatrix}.$$

Tomando determinantes a ambos lados de la igualdad uno se encuentra con lo siguiente;

$$\det(M) = \det(A) \det(\Theta_*[M]).$$

De modo que en caso de que $m = 1$, tendremos que D no es sino un escalar d e igualmente, un escalar será el cuasideterminante $\Theta_*[M]$. Por lo que $\det(\Theta_*[M]) = \Theta_*[M]$, y en este caso, el cuasideterminante podrá expresarse, como anticipábamos, cual cociente de determinantes:

$$\Theta_* \left[\begin{array}{c|c} A & B \\ \hline C & d \end{array}\right] = \frac{\det(M)}{\det(A)}.$$

Dicho todo esto nos encontramos en posición de dar las ya tan anunciadas expresiones para los polinomios ortogonales en términos de cuasideterminantes en la siguiente proposición.

[2]Los complementos de Schur (SC), introducidos en [78] y revisados en [130] pueden considerarse como predecesores a la vez que casos particulares de los cuasideterminantes; tema especialemente influenciado por los trabajos de Gel'fand [67] [68].

Proposición 2. *La secuencia de polinomios ortogonales mónicos y el cuadrado de sus normas se escriben como cuasideterminantes:*

$$P_k(x) = \Theta_* \left[\begin{array}{cccc|c} & & & & 1 \\ & G^{[k]} & & & x \\ & & & & \vdots \\ & & & & x^{k-1} \\ \hline G_{k,0} & G_{k,1} & \dots & G_{k,k-1} & x^k \end{array}\right], \quad h_k = \Theta_* \left[\begin{array}{cccc|c} & & & & G_{0,k} \\ & G^{[k]} & & & G_{1,k} \\ & & & & \vdots \\ & & & & G_{k-1,k} \\ \hline G_{k,0} & G_{k,1} & \dots & G_{k,k-1} & G_{k,k} \end{array}\right].$$

Demostración. Desde la factorización LU no es complicado observar que,

$$\begin{pmatrix} S_{k,0} & S_{k,1} & \dots & S_{k,k-1} & S_{k,k} \end{pmatrix} \begin{pmatrix} G_{0,0} & G_{0,1} & \dots & G_{0,k-1} \\ G_{1,0} & G_{1,1} & \dots & G_{1,k-1} \\ \\ G_{k,0} & G_{k,1} & \dots & G_{k,k-1} \end{pmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

Por lo tanto, puesto que $S_{k,k} = 1$,

$$\begin{pmatrix} S_{k,0} & S_{k,1} & \dots & S_{k,k-1} \end{pmatrix} G^{[k]} = - \begin{pmatrix} G_{k,0} & G_{k,1} & \dots & G_{k,k-1} \end{pmatrix}.$$

De modo que siempre que $G^{[k]}$ sea invertible (recordemos que esta era la condición impuesta sobre los menores de G) tendremos,

$$\begin{pmatrix} S_{k,0} & S_{k,1} & \dots & S_{k,k-1} \end{pmatrix} = - \begin{pmatrix} G_{k,0} & G_{k,1} & \dots & G_{k,k-1} \end{pmatrix} \left(G^{[k]}\right)^{-1}.$$

Desde donde se deduce la expresión para los polinomios. Para las normas, volvamos una vez más a la factorización LU:

$$h_k = \begin{pmatrix} S_{k,0} & S_{k,1} & \dots & S_{k,k-1} & S_{k,k} \end{pmatrix} \begin{pmatrix} G_{0,k} \\ G_{1,k} \\ \vdots \\ G_{k-1,k} \\ G_{k,k} \end{pmatrix} = \begin{pmatrix} S_{k,0} & S_{k,1} & \dots & S_{k,k-1} \end{pmatrix} \begin{pmatrix} G_{0,k} \\ G_{1,k} \\ \vdots \\ G_{k-1,k} \end{pmatrix} + G_{k,k}.$$

Y usando la expresión previa para las filas de S en función de G la proposición queda probada. □

La decisión arbitraria de tomar la versión mónica de los polinomios no es necesariamente la más común que uno puede encontrar en la literatura, dado que en ciertos contextos puede resultar más ventajoso elegirlos ortonormales o incluso fijar su valor en cierto punto concreto. Otro ingrediente de gran utilidad en las discusiones que siguen son lo que llamaré funciones de segunda especie y que defino a continuación:

Definición 10. *Llamaremos funciones de segunda especie a las siguientes transformadas de Cauchy de los polinomios ortogonales*

$$C_l(x) := \int_\Omega \frac{P_l(y)}{x-y} d\mu(y), \qquad \forall x \notin \Omega.$$

Estas funciones admiten la siguiente representación alternativa:

Proposición 3. *Las funciones de segunda especie, en la región adecuada, pueden darse en términos de las matrices de la factorización,*

$$C(x) = H(S^{-1})^\top \chi(x)^*, \qquad |y| < |x|, \quad y \in \Omega.$$

Demostración. En el intervalo de convergencia mencionado, se verifica la siguiente cadena de igualdades:

$$\begin{aligned} H(S^{-1})^\top \chi^*(x) = Sg\chi^*(x) &= \int_\Omega [S\chi(y)]\, d\mu(y)\, (\chi(y))^\top \chi^*(x) \\ &= \int_\Omega [S\chi(y)]\, d\mu(y) \frac{1}{x} \sum_{n=0}^{\infty} \left(\frac{y}{x}\right)^n = \int_\Omega \frac{P(y)}{x-y} d\mu(y). \end{aligned}$$

□

Como se puede observar directamente desde la proposición anterior, las funciones de segunda especie $C_l(x) = h_l \sum_{j=l}^{\infty} (S^{-\top})_{l,j} x^{-(j+1)}$ no son polinomios. En lugar de estas funciones, la literatura emplea un elemento muy relacionado con el que aquí se presenta, y que si que es un polinomio de grado $(l-1)$:

$$Q_l(x) := \frac{1}{h_0} \int_\Omega \frac{P_l(x) - P_l(y)}{x-y} \mathrm{d}\mu(y) = \frac{P_l(x)}{h_0} \int_\Omega \frac{1}{x-y} \mathrm{d}\mu(y) - \frac{1}{h_0} \int_\Omega \frac{P_l(y)}{x-y} \mathrm{d}\mu(y) = \frac{P_l(x) C_0(x) - C_l(x)}{h_0},$$

Estos polinomios $Q_l(x)$ reciben el nombre de polinomios de segunda especie, numeradores mónicos (por una razón que ahora comprenderemos) o bien polinomios asociados.

Corolario 1. *Desde la proposición 3 es sencillo comprobar que:*

$$C_0(x) = \int_\Omega \frac{1}{(x-y)} d\mu(y) = \frac{G_{0,0}}{x} + \frac{G_{0,1}}{x^2} + \cdots + \frac{G_{0,k}}{x^{k+1}} + \cdots = \sum_{j=0}^{\infty} \frac{G_{0,k}}{x^{k+1}}.$$

Siendo por lo tanto C_0 la transformada de Stieltjes de la medida.

Resulta apropiado desde este corolario, rendir, aunque sea breve, un merecido homenaje a uno de los temas indisociable de los orígenes de los polinomios ortogonales: las fracciones continuas.

$$\cfrac{B_1}{A_1 + \cfrac{B_2}{A_2 + \cfrac{B_3}{A_3 + \cfrac{B_4}{\cdots}}}}.$$

Aproximaciones de las mismas vienen dadas por sus *n-ésimos* convergentes,

$$\frac{S_n}{R_n} := \cfrac{B_1}{A_1 + \cfrac{B_2}{A_2 + \cfrac{B_3}{\ddots + \cfrac{B_n}{A_n}}}},$$

y cuyos numeradores y denominadores satisfacen las fórmulas de Wallis (para las que se definen $R_0 = 1$, $R_{-1} = 0$ y $S_0 = 0$, $S_{-1} = 1$),

$$R_n = A_n R_{n-1} + B_n R_{n-2}, \qquad S_n = A_n S_{n-1} + B_n S_{n-2}, \qquad \forall n \geq 1.$$

Las fraciones continuas son especialmente útiles a la hora de aproximar algunas funciones. Este es el caso de la transformada de Stieltjes de la medida que se había denotado por C_0 y que admite la siguiente aproximación (teorema de Markov):

$$C_0(x) = \int_\Omega \frac{1}{(x-y)} d\mu(y) = \sum_{j=0}^{\infty} \frac{G_{0,k}}{x^{k+1}} = \cfrac{B_1}{(x-A_1) + \cfrac{B_2}{(x-A_2) + \cfrac{B_3}{(x-A_3)+\cdots}}}.$$

Sorprendentemente el numerador y denominador que aparecen en los convergentes de esta fracción continua no son otros que, ¡los numeradores mónicos (de aquí su nombre) y los polinomios ortogonales!

$$\frac{S_n(x)}{R_n(x)} = \frac{B_1}{(x-A_1)+\frac{B_2}{(x-A_2)+\dots\frac{B_n}{(x-A_n)}}} = h_0\frac{Q_n(x)}{P_n(x)}.$$

Es más, recuperando las fórmulas de Wallis en este caso particular se obtiene la afamada relación de recurrencia a tres términos para los polinomios ortogonales y los numeradores mónicos.

$$Q_n = (x-A_n)Q_{n-1} + B_nQ_{n-2}, \qquad P_n = (x-A_n)P_{n-1} + B_nP_{n-2}, \qquad \forall n \geq 1.$$

Lo cual hace de perfecto final para esta sección e introducción de la que sigue.

1.2 Simetrías de la matriz de momentos: la ley de recurrencia

A continuación introduciré una herramienta en forma de matriz semi infinita que llamaré matriz de translación y que nos será de gran utilidad en consideraciones subsiguientes. En el caso que nos ocupa, esta matriz tiene todas sus entradas nulas salvo aquellas situadas en la primera super diagonal en la que encontramos la unidad:

Definición 11. *La expresión de la matriz de translación es la siguiente:*

$$\Lambda := \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix},$$

la cual satisface las siguientes propiedades:

$$\Lambda\chi(x) = x\chi(x), \qquad \Lambda\chi^*(x) = \frac{1}{x}\chi^*(x).$$

Denotando por $(E_{i,j})_{s,t} := \delta_{s,i}\delta_{t,j}$ a la base canónica matricial, es sencillo comprobar que:

- Por un lado $\Lambda\Lambda^\top = \mathbb{I}$; mientras que por el otro, $\Lambda^\top\Lambda = \mathbb{I} - E_{0,0}$.
- $\Lambda^\top\chi(x) = \frac{1}{x}(\mathbb{I} - E_{0,0})\chi(x)$.

Proposición 4. *El carácter autoadjunto del operador multiplicación por x en las entradas del producto interno se traduce en la estructura tipo Hankel de la matriz de momentos.*

$$\langle xf, h\rangle_\mu = \langle f, xh\rangle_\mu \Rightarrow \Lambda G = G\Lambda^\top \Rightarrow G_{i,j} = G_{i+j}.$$

Empleando la factorización LU en la simetría anterior y reorganizando términos se observa lo siguiente:

$$S\Lambda S^{-1} = H(S\Lambda S^{-1})^\top H^{-1}.$$

La matriz u operador resultante (la matriz translación revestida por las matrices de la factorización) tiene especial relevancia y merece un nombre propio:

Definición 12. *Se define la matriz de Jacobi J mediante la expresión:*

$$J := S\Lambda S^{-1} = H(S\Lambda S^{-1})^\top H^{-1}. \tag{1.1}$$

Desde su definición como la matriz de translación revestida por las matrices de la factorización, las entradas de la misma pueden escribirse como sigue:

$$J := \begin{pmatrix} -S_{10} & 1 & 0 & 0 & 0 & 0 & \dots \\ h_1 h_0^{-1} & S_{10}-S_{21} & 1 & 0 & 0 & 0 & \dots \\ 0 & h_2 h_1^{-1} & S_{21}-S_{32} & 1 & 0 & 0 & \dots \\ 0 & 0 & h_3 h_2^{-1} & S_{32}-S_{43} & 1 & 0 & \dots \\ 0 & 0 & 0 & h_4 h_3^{-1} & S_{43}-S_{54} & 1 & \dots \\ 0 & 0 & 0 & 0 & \ddots & \ddots & \ddots \\ \vdots & \vdots & \vdots & \vdots & & & \end{pmatrix}.$$

Proposición 5. *Las entradas de la matriz de Jacobi dan los coeficientes de la recurrencia de los polinomios ortogonales y de las funciones de segunda especie:*

$$JP(x) = xP(x) \qquad \Longrightarrow \qquad \forall k > 0, \quad \frac{h_k}{h_{k-1}} P_{k-1} + (S_{k,k-1} - S_{k+1,k}) P_k + P_{k+1} = x P_k, \tag{1.2}$$

$$JC(x) = xC(x) - h_0 \boldsymbol{e}_0 \qquad \Longrightarrow \qquad \forall k > 0, \quad \frac{h_k}{h_{k-1}} C_{k-1} + (S_{k,k-1} - S_{k+1,k}) C_k + C_{k+1} = x C_k, \tag{1.3}$$

donde $\boldsymbol{e}_0 := (1, 0, 0, \dots, 0, \dots)^\top$.

Demostración. Se sigue directamente desde las definiciones de los elementos involucrados. □

Nótese cómo la condición impuesta sobre los menores de la matriz de momentos hace que el coeficiente que acompaña a P_{k-1}, C_{k-1} en la ley de recurrencia no pueda ser cero; es más, en el caso definido positivo que estamos considerando, ha de ser un término positivo. Es preciso mencionar que muchas de las discusiones sobre polinomios ortogonales parten desde la ley de recurrencia. Este enfoque siempre va ligado al nombre de Favard (se puede consultar [60] para justificar tal mención, así como [95] donde se repasa algo de historia y sus generalizaciones).
Aun no siendo un tema del que me vaya a ocupar en esta tesis, por la simplicidad que supone incluir algún comentario al respecto en este preciso momento, me permito enunciar la siguiente proposición en la que se ligan los ceros de los polinomios ortogonales con la recientemente introducida matriz de Jacobi.

Proposición 6. *Los ceros de* P_k *son reales, simples (multiplicidad* 1*) y coinciden con los autovalores de* $J^{[k]}$.

Demostración. Se tiene que,

$$\left(J^{[k]}\right) \begin{pmatrix} P_0(x) \\ P_1(x) \\ \vdots \\ P_{k-1}(x) \end{pmatrix} = x \begin{pmatrix} P_0(x) \\ P_1(x) \\ \vdots \\ P_{k-1}(x) \end{pmatrix} - \begin{pmatrix} 0 \\ 0 \\ \vdots \\ P_k(x) \end{pmatrix}.$$

De modo que si α es un cero de $P_k(\alpha) = 0$ la ecuación anterior queda,

$$\left(J^{[k]}\right) \begin{pmatrix} P_0(\alpha) \\ P_1(\alpha) \\ \vdots \\ P_{k-1}(\alpha) \end{pmatrix} = \alpha \begin{pmatrix} P_0(\alpha) \\ P_1(\alpha) \\ \vdots \\ P_{k-1}(\alpha) \end{pmatrix}.$$

Por lo tanto, los autovalores de $J^{[k]}$ coinciden con los ceros de P_k. Si en lugar de considerar la matriz $J^{[k]}$ se toma la matriz semejante $L := \sqrt{H^{[k]}} J^{[k]} \left(\sqrt{H^{[k]}}\right)^{-1} = L^\top$ donde se ha denotado por $\sqrt{H^{[k]}} :=$

$\operatorname{diag}\{\sqrt{h_0}, \sqrt{h_1}, \dots\}$ se puede emplear el resultado que asegura que una matriz, con entradas reales, simétrica y tridiagonal tiene todos sus autovalores reales; si adicionalmente, todas las entradas en la super y subdiagonal son no nulas (este es nuestro caso pues a lo largo de las mismas se encuentran las $\sqrt{h_j}$) se puede afirmar que todos los autovalores serán diferentes. □

Nos encontramos en este punto en una situación semejante a la de la sección anterior cuando se mencionaron las fracciones continuas. No sería justo avanzar sin dedicarle unas breves líneas a otro de los temas bajo el que subyacen los orígenes de los polinomios ortogonales: la cuadratura de Gauss. La primera observación reside en que siendo $\{\alpha_{k,j}\}_{j=1}^{k}$ los ceros de $P_k(x)$, la proposición anterior puede reescribirse como sigue:

$$\left(J^{[k]}\right)^j = \mathbb{P}\begin{pmatrix}\alpha_{k,1}^j & & & \\ & \alpha_{k,2}^j & & \\ & & \ddots & \\ & & & \alpha_{k,k}^j\end{pmatrix}\mathbb{P}^{-1}, \qquad \mathbb{P} := \begin{pmatrix} 1 & 1 & \dots & 1 \\ P_1(\alpha_{k,1}) & P_1(\alpha_{k,2}) & \dots & P_1(\alpha_{k,k}) \\ \vdots & & & \\ P_{k-1}(\alpha_{k,1}) & P_{k-1}(\alpha_{k,2}) & \dots & P_{k-1}(\alpha_{k,k})\end{pmatrix}.$$

Como segunda observación, esta vez desde la ley de recurrencia, se tiene que:

$$x^j P(x) = J^j P(x) \qquad \Longrightarrow \qquad x^j = (J^j)_{0,0} + (J^j)_{0,1}P_1(x) + \dots + (J^j)_{0,j}P_j(x).$$

Por lo tanto,

$$\langle x^j, 1\rangle_\mu = \int_\Omega x^j \mathrm{d}\mu(x) = (J^j)_{0,0}\langle 1,1\rangle + (J^j)_{0,1}\langle P_1, 1\rangle + \dots + (J^j)_{0,j}\langle P_j, 1\rangle = (J^j)_{0,0}h_0.$$

En tercer lugar, la estructura $(2+1)$-diagonal que presenta J garantiza la igualdad $\left(J^j\right)_{0,0} = \left(\left[J^{[k]}\right]^j\right)_{0,0}$ siempre y cuando $0 \le j \le 2k-1$. De este modo, usando la diagonalización de J^j y denotando por $\mathbb{P}^{-1}\rfloor_1$ a la primera columna de la matriz $\mathbb{P}^{-1}$, se obtiene lo siguiente:

$$\int_\Omega x^j \mathrm{d}\mu(x) = \left(\left[J^{[k]}\right]^j\right)_{0,0} h_0 = h_0\begin{pmatrix}\alpha_{k,1}^j & \alpha_{k,2}^j & \dots & \alpha_{k,k}^j\end{pmatrix}\mathbb{P}^{-1}\rfloor_1 = \sum_{l=1}^{k}\lambda_{j,l}\alpha_{k,l}^j, \qquad 0 \le j \le 2k-1.$$

Lo que implica que los ceros de P_k son los k puntos para la cuadratura de μ con exactitud[3] $(2k-1)$.

1.3 El núcleo de Christoffel–Darboux

De acuerdo al teorema de aproximación de Weierstrass, los polinomios ortogonales en un intervalo finito de la recta real Ω son un sistema denso en el espacio de funciones continuas en dicho intervalo en la norma $||f||_\infty = sup_{x\in\Omega}|f(x)|$, es decir, si $f(x)$ es una función continua en Ω, existe un polinomio $p(x)$ tal que $||f(x)-p(x)||_\infty < \varepsilon$; lo cual implica que también lo son en la norma $L^2[\Omega]$. El espacio de funciones continuas es denso a su vez en $L^2[\Omega]$, por lo tanto la familia de polinomios ortogonales va a ser un sistema ortogonal completo en $L^2[\Omega]$. Esto significa que los polinomios ortogonales permitirán aproximar los elementos de $L^2[\Omega]$ con una precisión prefijada por el número de polinomios involucrados en la aproximación. La herramienta encargada de dar cuenta de esta propiedad es el núcleo de Christoffel–Darboux, que como estamos a punto de ver, puede interpretarse como la representación integral de la proyección sobre el espacio lineal generado por la secuencia de polinomios ortogonales.

[3]Fijado el número de puntos, la exactitud m de la cuadratura mide la precisión de la misma y se define como el grado más alto de todos los polinomios cuya cuadratura es una igualdad. Cuando μ tiene soporte sobre la recta real el valor más alto que puede tomar m es $2n-1$ y dicha cota superior se obtiene precisamente cuando los puntos coinciden con los ceros de la secuencia ortogonal asociada a μ.

Definición 13. *Se define el núcleo de Christoffel–Darboux del siguiente modo:*

$$K^{[l]}(x,y) := \sum_{k=0}^{l-1} P_k(x)h_k^{-1}P_k(y) = [P(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [P(y)]^{[l]} = \left(\chi(x)^{[l]}\right)^\top \left(G^{[l]}\right)^{-1} \chi(y)^{[l]}. \tag{1.4}$$

La última de las igualdades, que puede deducirse sin más que reescribiendo los polinomios ortogonales en función de las matrices de la factorización, se conoce en la literatura como teorema ABC (Aitken, Berg, Collar, [42], [30]). Como observación cabe destacar que el super índice $[l]$ hace referencia al número de términos presentes en la suma, no así al grado más alto de los polinomios involucrados en la misma.
Puesto que $K^{[l]}(x,y)$ se trata de un proyector, no debiera sorprender la siguiente igualdad:

$$\langle K^{[l]}(x,z), K^{[l]}(z,y)\rangle_\mu = K^{[l]}(x,y),$$

tampoco sorprende que este dé la mejor de las posibles aproximaciones (minimizando la norma $||f-\Pi^{[l]}[f]||$) de una función f dada sobre el espacio lineal generado por la secuencia de polinomios ortogonales hasta el grado $(l-1)$,

$$\Pi^{[l]}[f](y) := \langle f(x), K^{[l]}(x,y)\rangle_\mu = \sum_{j=0}^{l-1} \beta_j P_j(y), \qquad \beta_j = \frac{\langle f, P_j\rangle_\mu}{h_j}.$$

En caso de que f fuera un polinomio, la proyección (al orden apropiado) será una mera igualdad que descubre otra de las propiedades inherentes a nuestra secuencia de polinomios: los polinomios mónicos ortogonales minimizan, de entre todos los polinomios mónicos, la norma $L^2[\Omega]$.
Una característica notoria de la que goza $K^{[l]}$ es que, mientras los $P_l(x)$ satisfagan una ley de recurrencia cuyo número de términos no crezca con l, admitirá una expresión alternativa en la que en lugar de estar presentes los primeros l polinomios, solo serán precisos un número menor de estos. Este es el caso que tenemos entre manos (ley de recurrencia a tres términos) lo que me permite enunciar la siguiente proposición.

Proposición 7. *El núcleo $K^{[l]}(x,y)$ puede reescribirse en términos de dos polinomios consecutivos únicamente:*

$$K^{[l]}(x,y) = \frac{1}{h_{l-1}} \frac{P_l(x)P_{l-1}(y) - P_{l-1}(x)P_l(y)}{x-y}, \tag{1.5}$$

Su expresión confluente es:

$$\sum_{k=0}^{l-1} \frac{P_k^2(x)}{h_k} = \frac{1}{h_{l-1}}\Big(P_l'(x)P_{l-1}(x) - P_{l-1}'(x)P_l(x)\Big).$$

Demostración. Para la expresión principal, desde (1.2) y (1.1) escribimos,

$$H^{-1}JP(y) = yH^{-1}P(y) \Longrightarrow \left(H^{-1}J\right)^{[l]} P(y)^{[l]} + \left(H^{-1}J\right)^{[l,\geq l]} P(y)^{[\geq l]} = y\left(H^{-1}\right)^{[l]} P(y)^{[l]},$$
$$P(x)^\top H^{-1}J = xP(x)^\top H^{-1} \Longrightarrow [P(x)^\top]^{[l]} \left(H^{-1}J\right)^{[l]} + [P(x)^\top]^{[\geq l]} \left(H^{-1}J\right)^{[\geq l,l]} = x[P(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]}.$$

Multiplíquese la primera ecuación desde la izquierda por $[P(x)^\top]^{[l]}$ y la segunda desde la derecha por $P(y)^{[l]}$. Restando ambas se obtiene,

$$[P(x)^\top]^{[l]} \left(H^{-1}J\right)^{[l,\geq l]} P(y)^{[\geq l]} - [P(x)^\top]^{[\geq l]} \left(H^{-1}J\right)^{[\geq l,l]} P(y)^{[l]} = (y-x)[P(x)^\top]^{[l]} \cdot \left(H^{-1}\right)^{[l]} P(y)^{[l]}.$$

Operando en la parte izquierda de la igualdad e identificando en la parte derecha el núcleo de Chrisfoffel–Darboux, probamos la expresión. Para probar la confluente basta con tomar en la anterior el límite $y \to x$ al que basta con añadir y sustraer una cantidad apropiada. $\square$

Definición 14. *Definimos el núcleo mixto de Christoffel–Darboux como sigue:*

$$\mathcal{K}^{[l]}(x,y) := \sum_{k=0}^{l-1} P_k(x) h_k^{-1} C_k(y) = [P(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [C(y)]^{[l]} = \left(\chi(x)^{[l]}\right)^\top \left(\ \mathbb{I}_{l\times l} \mid (S^\top)^{[l]}([S^\top]^{-1})^{[j,\geq l]} \ \right) \chi^*(y).$$

Proposición 8. *Se cumple el siguiente resultado:*

$$\mathcal{K}^{[l]}(x,y) + \frac{1}{x-y} = \frac{1}{h_{l-1}} \frac{P_l(x)C_{l-1}(y) - P_{l-1}(x)C_l(y)}{x-y}.$$

Como corolario de este resultado se cumple que:

$$h_l = P_{l+1}(x)C_l(x) - P_l(x)C_{l+1}(x).$$

Demostración. Para probar el primer resultado basta con recordar la prueba empleada en el caso del núcleo no mixto, esto es, considerar las dos maneras posibles para calcular

$$P(x)^\top \left[H^{-1}JC(y)\right], \qquad \left[P(x)^\top H^{-1} J\right] C(y).$$

El corolario se consigue al tomar el límite $x = y$. □

La suma en l del resultado de la proposición permite dar la siguiente expresión:

Proposición 9. *La suma de las normas de los polinomios ortogonales satisface lo siguiente,*

$$\sum_{j=0}^{l} h_l = P^\top(x) \left[\left(\Lambda^\top - \Lambda\right)^{[l+2]} \right] C(x).$$

Definición 15. *El núcleo de Christoffel-Darboux de las funciones de segunda especie se define del siguiente modo:*

$$Q^{[l]}(x,y) := \sum_{k=0}^{l-1} C_k(x) h_k^{-1} C_k(y) = [C(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [C(y)]^{[l]} = \left(\chi^*(x)^{[l]}\right)^\top \left(G^{[l]}\right) \chi^*(y).$$

Y satisface la siguiente relación,

$$Q^{[l]}(x,y) - \left[\frac{C_0(x) - C_0(y)}{(x-y)}\right] = \frac{1}{h_{l-1}} \frac{C_l(x)C_{l-1}(y) - C_{l-1}(x)C_l(y)}{(x-y)}. \tag{1.6}$$

La primera expresión alternativa para escribir $Q^{[l]}(x,y)$ se sigue de la propia definición de las funciones de segunda especie, mientras que la relación que esta satisface es casi razonable teniendo en cuenta que las C_l tienen "prácticamente" idéntica relación de recurrencia a la de P_l. El hecho de que sean "prácticamente" idénticas, se traduce en el término adicional que aparece entre paréntesis acompañando a $Q^{[l]}(x,y)$.

1.4 Polinomios ortogonales clásicos y simetrías adicionales

En la sección que comienza escogeremos, de entre todas las posibles medidas definidas positivas con soporte contenido en la recta real, las llamadas medidas clásicas: Hermite, Laguerre y Jacobi. Los polinomios ortogonales asociados a estas medidas son, sin duda alguna, los polinomios ortogonales más famosos de entre todas las secuencias ortogonales. Sería imposible sobreestimar su importancia y ubiquidad en análisis numérico, ingeniería y física matemática. Por ejemplo, los polinomios de Hermite son necesarios para describir un oscilador cuántico símple; los polinomios de Laguerre aparecen al resolver la parte radial de la ecuación de Schrödinger para el átomo de hidrógeno; los polinomios de Jacobi son imprescindibles en el estudio de

ecuaciones armónicas con simetría esférica. Incluso la localización de los ceros de estos polinomios está ligada a problemas de optimización energética dependiente de la posición relativa de un número establecido de cargas en ciertos potenciales electrostáticos. A pesar de que se podría continuar con la lista de aplicaciones de estos polinomios durante varios volúmenes, en este caso los hemos escogido porque las matrices de momentos asociadas a las medidas respecto de las que son ortogonales gozan de unas simetrías adicionales que darán cuenta de las propiedades típicas de los polinomios clásicos. Como se puede deducir observando los resultados de las secciones anteriores, basta con conocer ciertos elementos de las matrices de factorización para expresar aquellos en función de estos. Dichos elementos son las h_n (el cuadrado de las normas) y las $S_{n+1,n}$ (entradas de la subdiagonal de las matrices de la factorización) para todos los valores $n = 0, 1, 2 \dots$. De este modo, nuestro propósito será el de extraer de algún modo la información necesaria para conocer una expresión para los mismos.
En lo que resta de sección denotaremos a las medidas involucradas por $\mathrm{d}\mu(x) = u_\gamma(x)dx$; donde u_γ será cualquiera de los pesos clásicos y γ los parámetros que caracterizan estas: Hermite ($\gamma = \{\varnothing\}$), Laguerre ($\gamma = \{\alpha\}$), Jacobi ($\gamma = \{\alpha, \beta\}$). Recordemos que los polinomios de Jacobi contienen como casos particulares los polinomios de segunda especie de Chebyshev ($\alpha = \beta = \pm\frac{1}{2}$), los polinomios de Legendre ($\alpha = \beta = 0$) y los de Gegenbauer o ultraesféricos ($\alpha = \beta$).

1.4.1 Propiedades de los pesos clásicos

Los polinomios clásicos se pueden caracterizar por ejemplo como las únicas secuencias de polinomios ortogonales que son autofunciones de un operador diferencial lineal de segundo orden prefijado [31], bien como aquellos con una expresión determinada en términos de una fórmula de Rodrigues' [123], o bien como aquellos cuyas derivadas no pierden la propiedad de ortogonalidad [74] [75]. A nosotros nos interesa caracterizarlos desde las medidas respecto de las que estos son ortogonales, es decir, como aquellos asociados a medidas que satisfacen una ecuación diferencial dada (Pearson). Estas propiedades específicas de estas medidas van a inducir las simetrías adicionales de las matrices de momentos que buscamos. Nos interesamos por las dos siguientes:

1. Ecuación diferencial de tipo Pearson.

   Definición 16. *El peso u_γ se considerará clásico en caso de existir polinomios $p_2(x) = ax^2 + bx + c$ y $p_{1,\gamma}(x) = (A_\gamma - 2a)x + (B_\gamma - b)$ con $A_\gamma \neq 0$ de modo que u_γ satisfaga la siguiente ecuación:*

   $$p_2(x)\frac{\mathrm{d}}{\mathrm{d}x}u_\gamma = p_{1,\gamma}u_\gamma. \tag{1.7}$$

   Por completitud enumeraremos aquí los posibles pesos $u_\gamma(x)$, sus correspondientes $p_{1,\gamma}$, p_2 y soporte Ω.

   - Hermite $u(x) = e^{-x^2}$, $x \in \mathbb{R}$. Con $p_1 = -2x$, $p_2 = 1$.
   - Laguerre $u_\alpha(x) = x^\alpha e^{-x}$, $\alpha > -1$, $x \in \mathbb{R}_+$. Con $p_{1,\alpha} = (\alpha - x)$, $p_2 = x$.
   - Jacobi $u_{\alpha,\beta}(x) = (1-x)^\alpha(1+x)^\beta$, $\alpha, \beta > -1$, $x \in (-1, 1)$. Con $p_{1,\alpha,\beta} = -[(\alpha - \beta) + (\alpha + \beta)x]$, $p_2 = 1 - x^2$.

   Dichos pesos dependen de: cero ($\gamma = \{\varnothing\}$), uno ($\gamma = \{\alpha\}$) y dos parámetros ($\gamma = \{\alpha, \beta\}$) respectivamente.

2. Aumento en una unidad de cada uno de los parámetros.
   Denotaremos por $u_{\gamma+1}(x)$ a cualquiera de los pesos anteriores pero con la adición de la unidad a cada uno de los parámetros presentes en su definición; en el peso de Hermite no habrá cambio alguno, en el de Laguerre pasaremos de α a $\alpha + 1$ y finalmente en el de Jacobi pasaremos de α, β a $\alpha + 1, \beta + 1$. La razón de este apunte tiene que ver con la siguiente proposición.

Proposición 10. *El polinomio p_2 tiene la siguiente propiedad,*

$$p_2(x)u_\gamma(x) = u_{\gamma+1}(x). \tag{1.8}$$

Demostración. Basta con comprobarlo en las definiciones. □

1.4.2 Simetrías adicionales de las matrices de momentos clásicas

Veamos ahora que las dos propiedades de los pesos clásicos que acabamos de enumerar se traducen bien en características de los productos internos asociados o bien en simetrías de las matrices de momentos. Una integración por partes y el uso de la primera de las propiedades anteriores nos permite escribir para dos funciones f, h de nuestro espacio de Hilbert,

$$\langle p_2 f', h\rangle_{u_\gamma} = -\langle f, (p_2' + p_1)h\rangle_{u_\gamma} - \langle f, p_2 h'\rangle_{u_\gamma}. \tag{1.9}$$

Nótese cómo también usamos implícitamente la segunda de las propiedades gracias a la cual se anulan los términos de contorno,

$$p_2 u_\gamma\Big|_{\partial\Omega} = u_{\gamma+1}\Big|_{\partial\Omega} = 0.$$

Para entender todo esto como una ecuación matricial que involucre a las matrices de momentos es preciso introducir la siguiente matriz relacionada con $\Lambda^\top$.

Definición 17. *La representación matricial del operador derivada respecto de la base $\chi(x)$ es la siguiente:*

$$D := \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 2 & 0 & 0 & \dots \\ 0 & 0 & 3 & 0 & \dots \\ 0 & 0 & 0 & 4 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}, \qquad D\chi(x) = \chi(x)'.$$

Empleando esta notación, dado un polinomio cualquiera $f(x) = (f_0, f_1, \dots)\,\chi(x) = \boldsymbol{f}^\top\chi(x)$ podremos expresar su multiplicación por x o su derivada como sigue:

$$xf(x) = \boldsymbol{f}^\top\Lambda\chi(x), \qquad f(x)' = \boldsymbol{f}^\top D\chi(x).$$

De este modo, es prácticamente directo desde (1.9) comprobar que,

$$DG_{\gamma+1} = -G_\gamma\Big(p_2'(\Lambda) + p_{1,\gamma}(\Lambda) + Dp_2(\Lambda)\Big)^\top.$$

Donde $p_i(\Lambda)$ no son otros que los operadores que aparecen al sustituir en los polinomios correspondientes las x por la matriz Λ.
La factorización LU de las matrices de momentos en la relación anterior permite enunciar la siguiente proposición.

Proposición 11. *Las matrices de la factorización LU de la matriz de momentos satisfacen,*

$$S_\gamma D S_{\gamma+1}^{-1} = -H_\gamma\Big(S_{\gamma+1}\big(p_2'(\Lambda) + p_{1,\gamma}(\Lambda) + Dp_2(\Lambda)\big)S_\gamma^{-1}\Big)^\top H_{\gamma+1}^{-1}. \tag{1.10}$$

Ecuación que en componentes se reescribe como,

$$(S_\gamma)_{n+1,n} = (n+1)\frac{B+nb}{A+2na}, \tag{1.11}$$

$$(h_\gamma)_n = \frac{-n}{A_\gamma + (n-1)a}(h_{\gamma+1})_{n-1}. \tag{1.12}$$

Recordemos que estos son precisamente los coeficientes de la teoría que se trataban de encontrar.

Demostración. Mientras la parte izquierda de la primera ecuación impone al resultado tratarse de una matriz con entradas no nulas únicamente a lo largo de las diagonales m-*ésimas* siendo estas $m = -1, -2, \dots$, la parte derecha dictamina un resultado semejante pero esta vez para las diagonales j-*ésimas* con $j = -1, 0, 1, 2, \dots$. Para que ambas condiciones sean equivalentes será preciso que $j = m = -1$. Además, la parte derecha y la parte izquierda de la igualdad proporcionan dos maneras alternativas de calcular cada una de las entradas de la matriz resultante, con lo que tendremos una ecuación para cada entrada de la misma desde las que deducir el resto de relaciones.
Para $m = -2$,

$$(S_\gamma)_{n+1,n} = \frac{n+1}{n}(S_{\gamma+1})_{n,n-1}, \qquad n = 1, 2, \dots.$$

Para $j = 0$,

$$(S_\gamma)_{1,0} = \frac{B}{A}, \qquad (S_{\gamma+1})_{n,n-1}(A_\gamma + [n-1]a) + B_\gamma + nb = (S_\gamma)_{n,n+1}(A_\gamma + na).$$

Combinando ambos resultados se obtiene la expresión para las $(S_\gamma)_{n,n+1}$. □

Volviendo a la ecuación (1.9) y empleándola para el producto interno $\langle p_2 f'', h\rangle$ se deduce que,

$$\left\langle \left[p_2 \frac{\mathrm{d}^2}{\mathrm{d}x^2} + (p_2' + p_{1,\gamma})\frac{\mathrm{d}}{\mathrm{d}x}\right] f, h \right\rangle_{u_\gamma} = \left\langle f, \left[p_2 \frac{\mathrm{d}^2}{\mathrm{d}x^2} + (p_2' + p_{1,\gamma})\frac{\mathrm{d}}{\mathrm{d}x}\right] h \right\rangle_{u_\gamma}. \tag{1.13}$$

Este no es sino un operador diferencial de segundo orden con coeficientes polinómicos y autoadjunto. La versión matricial de esta propiedad la resumimos en la siguiente proposición.

Proposición 12. *Las matrices de momentos clásicas gozan de la siguiente simetría adicional dada por la representación matricial de un operador diferencial lineal de segundo orden con coeficientes polinómicos:*

$$[D^2(a\Lambda^2 + b\Lambda + c) + D(A_\gamma\Lambda + B_\gamma)]G_\gamma = G_\gamma\left[D^2(a\Lambda^2 + b\Lambda + c) + D(A_\gamma\Lambda + B_\gamma)\right]^\top.$$

Factorizando la matriz de momentos y despejando adecuadamente nos encontramos con una interesante observación.

Proposición 13. *Las matrices de la factorización diagonalizan el operador diferencial autoadjunto,*

$$N_\gamma := S_\gamma\left[D^2(a\Lambda^2 + b\Lambda + c) + D(A_\gamma\Lambda + B_\gamma)\right]S_\gamma^{-1} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & A_\gamma & 0 & 0 & 0 & \dots \\ 0 & 0 & 2(A_\gamma + a) & 0 & 0 & \dots \\ 0 & 0 & 0 & 3(A_\gamma + 2a) & 0 & \dots \\ 0 & 0 & 0 & 0 & \ddots & \\ \vdots & \vdots & \vdots & \vdots & & \end{pmatrix}.$$

Donde los elementos de la diagonal $(N_\gamma)_n$ juegan el papel de autovalores de la secuencia de polinomios ortogonales, siendo estos últimos sus autofunciones, es decir,

$$F_\gamma := p_2\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (p_2' + p_{1,\gamma})\frac{\mathrm{d}}{\mathrm{d}x} \qquad \Longrightarrow \qquad F_\gamma\left[P_\gamma(x)\right] = N_\gamma P_\gamma(x).$$

Por completitud, y con ánimo de ilustrar las conclusiones de esta sección, enumeramos para cada peso clásico los resultados correspondientes.

- **Hermite**. Ecuación diferencial: $\frac{du}{dx} = -2xu$; valores de los coeficientes: $A = -2,\ B = 0,\ a = 0,\ b = 0,\ c = 1$; condición inicial: $h_0 = \sqrt{\pi}$; resto de coeficientes:

$$S_{n+1,n} = 0, \qquad h_n = \sqrt{\pi}\frac{n!}{2^n}, \qquad N_n = -n.$$

- **Laguerre**. Ecuación diferencial: $x\frac{du_\alpha}{dx} = (\alpha - x)u_\alpha$; valores de los coeficientes: $A_\alpha = -1,\ B_\alpha = (1+\alpha),\ a = 0,\ b = 1,\ c = 0$; condición inicial: $(h_\alpha)_0 = \Gamma(\alpha+1)$; resto de coeficientes:

$$(S_\alpha)_{n+1,n} = -(n+1)[(n+1)+\alpha], \qquad (h_\alpha)_n = n!\Gamma(\alpha+n+1), \qquad (N\alpha)_n = -2n.$$

- **Jacobi**. Ecuación diferencial: $(1-x^2)\frac{du_{\alpha,\beta}}{dx} = -[(\alpha-\beta)+(\alpha+\beta)x]u_{\alpha,\beta}$; valores de los coeficientes: $A_{\alpha,\beta} = -[(\beta+\alpha)+2],\ B_{\alpha,\beta} = -(\alpha-\beta),\ a = -1,\ b = 0,\ c = 1$; condición inicial: $(h_{\alpha,\beta})_0 = \frac{\Gamma(\alpha+1)\Gamma(\beta+1)2^{\alpha+\beta+1}}{(\alpha+\beta+1)\Gamma(\alpha+\beta+1)}$; resto de coeficientes:

$$\begin{aligned}
(S_{\alpha,\beta})_{n+1,n} &= \frac{(n+1)(\alpha-\beta)}{(\alpha+\beta+2)+2n},\\
(h_{\alpha,\beta})_n &= n!2^{(\alpha+\beta+2n+1)}\frac{\Gamma(\alpha+\beta+n+1)\Gamma(\alpha+n+1)\Gamma(\beta+n+1)}{(\alpha+\beta+2n+1)\Gamma^2(\alpha+\beta+2n+1)},\\
(N_{\alpha,\beta})_n &= -n(\beta-\alpha+1+n).
\end{aligned}$$

1.5 Deformaciones discretas

Partiendo de la idea de que cualquier nueva secuencia de polinomios ortogonales siempre será bien recibida, en esta sección me encargaré de explicar distintos métodos que permiten construir nuevas secuencias a partir de una previa conocida. La idea consiste en proponer transformaciones sencillas de la medida cuya secuencia de polinomios se conoce y preguntarse por la posibilidad de que la nueva medida deformada tenga una secuencia deformada de polinomios asociada, y en el caso afirmativo, por su expresión en términos de elementos de la teoría original.

1.5.1 Transformaciones espectrales lineales de la medida

Esta primera parte trata del tipo de deformaciones sencillas que se engloban bajo el nombre de transformaciones espectrales lineales de la medida[4] y que consisten en multiplicar a la medida inicial por una función racional $\mathrm{d}\mu(x) \longrightarrow \mathrm{d}\tilde{\mu}(x) := \frac{R(x)}{Q(x)}\mathrm{d}\mu(x)$. Detengámonos un instante para enumerar los siguientes comentarios:

- Para que la definición tenga sentido, los ceros de $Q(x)$ no podrán pertenecer al soporte Ω de la medida inicial.
- En caso de partir de una medida definida positiva y querer mantener esta propiedad se habrá de exigir que $\frac{R(x)}{Q(x)}$ sea positivo en Ω.
- El concepto de deformar una medida puede generalizarse de una manera sencilla si la deformación se considera actuando sobre el funcional lineal, es decir $L_\mu \longrightarrow L_{\tilde{\mu}}$ donde el funcional deformado y el inicial se relacionan del siguiente modo $L_{\tilde{\mu}}[f(x)Q(x)] = L_\mu[f(x)R(x)]$. Esta es una ecuación que debe entendenderse entre elementos del dual del espacio de polinomios $\tilde{\mu}Q = \mu R$ con lo que su solución $\tilde{\mu} = \mu\frac{R}{Q} + \nu$ donde $\nu Q = 0$ tiene un carácter funcional y será más general que la correspondiente restricción al espacio de medidas (ν no tiene por qué ser única).

[4]Este es el nombre por el que se conoce a estas transformaciones en el contexto de los polinomios ortogonales [119]; en el estudio de los sistemas integrables se les denota por transformaciones de Darboux [100]; finalmente se las llama transformaciones de Lévy en el caso de la teoría de las transformaciones de superficies preservando ciertas propiedades [57].

Entenderemos mejor estos comentarios tras introducir las siguientes definiciónes.

Definición 18. *Dados dos polinomios mónicos y coprimos* $R(x) := \prod_{i=1}^{d}(x-r_i)^{m_i}$ *y* $Q(x) := \prod_{i=1}^{s}(x-q_i)^{n_i}$ *con grados* $\sum_{i=1}^{d} m_i = M$ *y* $\sum_{i=1}^{s} n_i = N$ *respectivamente, y tales que* $\{q_1, q_2, \dots q_s\} \cap \Omega = \varnothing$*; las transformaciones de Christoffel* ($Q(x) = 1$) *y de Geronimus* ($R(x) = 1$) *de un funcional inicial* μ *son,*

$$\hat{\mu}(x) := R(x)\mu(x), \qquad \hat{\Omega} = \Omega.$$

$$\check{\mu}(x) := \frac{\mu(x)}{Q(x)} + \sum_{i=1}^{s}\sum_{l=0}^{(n_i-1)}(-1)^l \frac{\xi_{i,l}}{l!}\delta^{(l)}(x-q_i), \qquad \check{\Omega} = \Omega \cup \{q_1, q_2, \dots q_s\}.$$

La composición de las anteriores da lugar a la citada transformación espectral lineal,

$$\tilde{\mu}(x) := (\widehat{\check{\mu}})(x) = \frac{R(x)\mu(x)}{Q(x)} + \sum_{i=1}^{s}\sum_{l=0}^{(n_i-1)}(-1)^l \frac{\xi_{i,l}}{l!}R(x)\delta^{(l)}(x-q_i), \qquad \tilde{\Omega} = \Omega \cup \{q_1, q_2, \dots q_s\}.$$

Donde entenderemos por $\delta^{(l)}$ *la derivada distribucional* l*-ésima de la delta de Dirac.*

Nótese cómo salvo en el caso Christoffel en el que partiendo de una medida se llegará a también a otra medida, en los casos Geronimus y espectral lineal, de restringir la solución únicamente a medidas será necesario tomar $\xi_{i,j} = 0 \; \forall j > 0$. A misma cantidad de esfuerzo, resulta ventajoso considerar la generalización al concepto de deformación del funcional lineal. Añadamos a la notación el conjunto de elementos transformados: matrices de momentos deformadas $\hat{G}, \check{G}, \tilde{G}$, polinomios ortogonales $\hat{P}_n(x), \check{P}_n(x), \tilde{P}_n(x)$, normas (en caso de seguir siendo normas) $\hat{h}_n, \check{h}_n, \tilde{h}_n$, etc. Es de destacar que de permitir $R \to Q$ en el caso espectral lineal, se obtendrá que $\tilde{\mu}(x) = \mu(x)$, es decir, la transformación identidad, mientras que de componer las transformaciones en orden contrario se tendrá $(\check{\hat{\mu}})(x) = \mu(x) + \sum_{i=1}^{s}\sum_{l=0}^{(n_i-1)}(-1)^l \frac{\xi_{i,l}}{l!}\delta^{(l)}(x-q_i)$ que es una transformación de Uvarov generalizada y de la que se hablará más adelante.
El enfoque clásico de este tipo de deformaciones procede desde la transformada de Stieltjes de la medida que nosotros hemos denotado por $C_0(y)$, la primera entrada del vector de las funciones de segunda especie,

$$C_0(y) = \int \frac{\mathrm{d}\mu(x)}{y-x}.$$

El motivo de partir desde este punto reside en las siguientes expresiones que relacionan las C_0 transformadas en términos de las originales y que además ponen de manifiesto el adjetivo “espectral lineal” del nombre que llevan estas deformaciones[5]. Consideremos por un segundo que deformamos una medida dada mediante polinomios de orden uno; en tal caso, las C_0 transformadas se escribirán de acuerdo a las expresiones,

$$\begin{aligned}
\hat{C}_0(y) &= \int \frac{(x-r)\mathrm{d}\mu}{y-x} = (y-r)\int \frac{\mathrm{d}\mu}{y-x} - \int \mathrm{d}\mu(x) = (y-r)C_0(y) - h_0, \\
\check{C}_0(y) &= \int \frac{\mathrm{d}\mu}{(x-q)(y-x)} + \xi \int \frac{\delta(x-q)\mathrm{d}\mu}{y-x} = \frac{1}{(y-q)}\left[\int \frac{\mathrm{d}\mu}{y-x} - \int \frac{\mathrm{d}\mu(x)}{x-q}\right] + \xi\frac{1}{y-q} = \frac{C_0(y) - C_0(q) + \xi}{(y-q)}, \\
\tilde{C}_0(y) &= \int \frac{(x-r)\mathrm{d}\mu}{(x-q)(y-x)} + \xi \int \frac{\delta(x-q)(x-r)\mathrm{d}\mu}{y-x} = \frac{(y-r)C_0(y) - (q-r)C_0(q) + (q-r)\xi}{(y-q)}.
\end{aligned}$$

[5] El adjetivo espectral lineal es acertado únicamente en el contexto escalar y univariado. Aunque resulta algo menos apropiado en sus versiones generalizadas lo mantendré a lo largo de la memoria.

La transformación de Christoffel

Presentada por primera vez por Elwin Chrisfoffel en sus discusiones realacionadas con la cuadratura gaussiana [41] y considerada posteriormente en el contexto de la teoría de Sturm-Liouville en [47], la transformación de Christoffel $\hat{\mu}$, puede describirse por la siguiente relación entre las respectivas matrices de momentos,

$$\hat{G} = R(\Lambda)G = GR(\Lambda^\top).$$

Asumiendo la posibilidad de factorizar LU ambas matrices de momentos se considera la siguiente matriz,

Definición 19. *Sea el conector de la transformación la siguiente matriz:*

$$\hat{\omega} := \hat{S}R(\Lambda)S^{-1} = \hat{H}\left(S\hat{S}^{-1}\right)^\top H^{-1}.$$

Proposición 14. *El conector permite relacionar los polinomios transformados y los originales,*

$$\hat{\omega}P(x) = R(x)\hat{P},$$

y tiene únicamente $(M+1)$ diagonales no nulas:

$$\hat{\omega} = \begin{pmatrix} \hat{\omega}_{0,0} & \hat{\omega}_{0,1} & \dots & \hat{\omega}_{0,(M-1)} & \hat{\omega}_{0,M} & 0 & & & \\ 0 & \hat{\omega}_{1,1} & & & \hat{\omega}_{1,M} & \hat{\omega}_{1,(M+1)} & 0 & \dots & \\ 0 & 0 & \ddots & & & & \ddots & & \\ & & & \hat{\omega}_{k,k} & & & \hat{\omega}_{k,k+M-1} & \hat{\omega}_{k,k+M} & 0 \\ & & & & \ddots & & & & \ddots \end{pmatrix}.$$

Siendo cierto que $\hat{\omega}_{k,k+M} = 1$ y $\hat{\omega}_{k,k} = \frac{\hat{h}_k}{h_k}$.

Demostración. La fórmula de conexión es la simple consecuencia de la definición de $\hat{\omega}$, mientras que su estructura de $(M+1)$ diagonales no nulas se sigue de la ecuación que resulta de la factorización LU de las matrices de momentos. □

Con el propósito de manejar una notación lo más clara posible, se propone la siguiente definición.

Definición 20. *Dado un conjunto de tuplas, en nuestro caso ceros y sus multiplicidades $r = \{(r_i, m_i)\}_{i=1}^d$ definimos para cualquier funcion $f(x)$ el siguiente operador $\mathcal{J}_r[f] : \mathcal{F}(x) \longrightarrow \mathbb{R}^M$ cuyas entradas son los coeficientes de $(m_i - 1)$-jets donde $i = 1, 2, \dots, d$:*

$$\mathcal{J}_r[f] := \left(\frac{f^{(0)}(r_1)}{0!}, \frac{f^{(1)}(r_1)}{1!}, \dots, \frac{f^{(m_1-1)}(r_1)}{(m_1-1)!}; \frac{f^{(0)}(r_2)}{0!}, \dots, \frac{f^{(m_2-1)}(r_2)}{(m_2-1)!}; \dots; \frac{f^{(0)}(r_d)}{0!}, \dots, \frac{f^{(m_d-1)}(r_d)}{(m_d-1)!}\right).$$

En caso de que la función f dependiera de más de una variable $f(x_1, x_2, \dots)$, $\mathcal{J}_r^{(j)}[f]$ denotará al correspondiente vector de jets tomado respecto de la variable j-ésima y considerando al resto de estas como parámetros.

Por fin estamos listos para dar la relación entre las secuencias de polinomios transformadas y las originales que se presenta a continuación.

Proposición 15. *Los polinomios deformados bajo una transformación de Christoffel se pueden expresar en función de los originales como sigue:*

$$\hat{P}_n(x) = \frac{1}{R(x)}\Theta_*\left(\begin{array}{c|c} \mathcal{J}_r\begin{bmatrix} P_n \\ P_{n+1} \\ \vdots \\ P_{n+M-1}\end{bmatrix} & \begin{matrix} P_n(x) \\ P_{n+1}(x) \\ \vdots \\ P_{n+M-1}(x)\end{matrix} \\ \hline \mathcal{J}_r[P_{n+M}] & P_{n+M}(x)\end{array}\right), \qquad \frac{\hat{h}_n}{h_n} = \Theta_*\left(\begin{array}{c|c} \mathcal{J}_r\begin{bmatrix} P_n \\ P_{n+1} \\ \vdots \\ P_{n+M-1}\end{bmatrix} & \begin{matrix} 1 \\ 0 \\ \vdots \\ 0\end{matrix} \\ \hline \mathcal{J}_r[P_{n+M}] & 0\end{array}\right).$$

Demostración. Tomando la componente *n-ésima* de la fórmula de conexión,

$$\begin{pmatrix} \hat{\omega}_{n,n} & \hat{\omega}_{n,n+1} & \dots & \hat{\omega}_{n,n+M-1} & 1 \end{pmatrix} \begin{pmatrix} P_n(x) \\ P_{n+1}(x) \\ \vdots \\ P_{n+M-1}(x) \\ P_{n+M}(x) \end{pmatrix} = R(x)\hat{P}_n(x).$$

Tras evaluar en los ceros de $R(x)$ se puede comprobar que,

$$\begin{pmatrix} \hat{\omega}_{n,n} & \hat{\omega}_{n,n+1} & \dots & \hat{\omega}_{n,n+M-1} & 1 \end{pmatrix} \mathcal{J}_r \begin{bmatrix} P_n \\ P_{n+1} \\ \vdots \\ P_{n+M-1} \\ P_{n+M} \end{bmatrix} = \begin{pmatrix} 0 & 0 & \dots & 0 \end{pmatrix}.$$

Por lo tanto,

$$\begin{pmatrix} \hat{\omega}_{n,n} & \hat{\omega}_{n,n+1} & \dots & \hat{\omega}_{n,n+M-1} \end{pmatrix} \mathcal{J}_r \begin{bmatrix} P_n \\ P_{n+1} \\ \vdots \\ P_{n+M-1} \end{bmatrix} = -\mathcal{J}_r[P_{n+M}],$$

$$\begin{pmatrix} \hat{\omega}_{n,n} & \hat{\omega}_{n,n+1} & \dots & \hat{\omega}_{n,n+M-1} \end{pmatrix} = -\mathcal{J}_r[P_{n+M}] \left(\mathcal{J}_r \begin{bmatrix} P_n \\ P_{n+1} \\ \vdots \\ P_{n+M-1} \end{bmatrix} \right)^{-1}.$$

Desde donde se sigue el resultado que se quería demostrar. □

Proposición 16. *Los núcleos de Christoffel–Darboux transformados y originales están relacionados,*

$$\begin{aligned} K^{[n+1]}(x,y) &= R(y)\hat{K}^{[n+1]}(x,y) \\ &- \begin{pmatrix} \hat{P}_{n+1-M} & \dots & \hat{P}_n \end{pmatrix} \begin{pmatrix} \hat{h}^{-1}_{n+1-M} & & \\ & \ddots & \\ & & \hat{h}^{-1}_n \end{pmatrix} \begin{pmatrix} \hat{\omega}_{n+1-M,n+1} & & 0 \\ \vdots & \ddots & \\ \hat{\omega}_{n,n+1} & \dots & \hat{\omega}_{n,n+M} \end{pmatrix} \begin{pmatrix} P_{n+1}(y) \\ \vdots \\ P_{n+m}(y) \end{pmatrix}. \end{aligned}$$

Demostración. La prueba se basa en las fórmulas de conexión,

$$\hat{H}^{-1}\hat{\omega}P(y) = R(y)\hat{H}^{-1}\hat{P}(y),$$
$$\left(\hat{P}(x)\right)^{\top} \hat{H}^{-1}\hat{\omega} = P(x)H^{-1}.$$

□

Dicha relación permite dar una expresión alternativa a la anterior para los polinomios transformados.

Proposición 17. *La siguiente expresión alternativa para los polinomos transformados en función del núcleo de Christoffel–Darboux original es cierta,*

$$\frac{\hat{P}_n(x)}{\hat{h}_n} = \Theta_* \left(\begin{array}{c|c} \mathcal{J}_r \begin{bmatrix} P_{n+1} \\ \vdots \\ P_{n+M} \end{bmatrix} & \begin{matrix} 0 \\ \vdots \\ 1 \end{matrix} \\ \hline \mathcal{J}^{(2)}_r[K^{[n+1]}(x,y)] & 0 \end{array} \right).$$

Demostración. Para probar esta expresión es necesario seguir un procedimiento idéntico al de la prueba de la proposoción 15, pero esta vez partiendo desde la ecuación (16). □

Antes de concluir con la sección, y volviendo a la definición 19 pongamos nombre a la matriz $S\hat{S}^{-1}$.

Definición 21. *Se define la siguiente matriz unitriangular inferior relacionada con el conector,*

$$\hat{\Omega} := S\hat{S}^{-1}, \qquad \hat{\omega} = \hat{H}\hat{\Omega}^{\top}H^{-1}.$$

Esta definición adquiere sentido una vez se observa la siguiente proposición,

Proposición 18. *La $\hat{\omega}$ y la $\hat{\Omega}$ son los factores triangular superior e inferior que proporcionan las siguientes factorizaciones LU y UL del polinomio perturbador evaluado en las matrices de Jacobi J y $\hat{J}$.*

$$\hat{\Omega}\hat{\omega} = R(J), \qquad \hat{\omega}\hat{\Omega} = R(\hat{J}).$$

Demostración. Este resultado es consecuencia directa de expresar cada uno de los elementos que aparecen en el mismo en términos de las matrices de la factorización y operar desde allí. □

La transformación de Geronimus

El primero en considerar la transformación de Geronimus $\check{\mu}$ fue Y. L Geronimus en [72]. Una vez más, nos interesa la definición de esta deformación en términos de las matrices de momentos involucradas,

$$Q(\Lambda)\check{G} = \check{G}Q(\Lambda^{\top}) = G.$$

A partir de los coeficientes del polinomio perturbador $Q(x) = (Q_0, Q_1, \dots, Q_{N-1}, 1, 0, \dots, 0, \dots)\chi(x)$, se construye la siguiente matriz semi infinita que será de utilidad en los razonamientos que siguen,

$$\mathbf{Q} := \begin{pmatrix} Q_1 & Q_2 & Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots \\ Q_2 & Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots & \\ Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots & & \\ \dots & Q_{N-1} & 1 & 0 & \dots & & & \\ Q_{N-1} & 1 & 0 & \dots & & & & \\ 1 & 0 & \dots & & & & & \\ 0 & \dots & & & & & & \end{pmatrix}.$$

Asumiendo que ambas matrices de momentos admiten su factorización LU correspondiente es posible defirnir el conector.

Definición 22. *El conector de la transformación de Geronimus es:*

$$\check{\omega} := \check{S}S^{-1} = \check{H}\left(\check{S}^{-1}\right)^{\top}Q(\Lambda^{\top})S^{\top}H^{-1}.$$

Proposición 19. *El conector permite relacionar elementos deformados con los originales,*

$$\check{\omega}P(x) = \check{P}(x), \qquad \check{\omega}C(x) = Q(x)\check{C}(x) - \check{H}\left(\check{S}^{-1}\right)^{\top}\mathbf{Q}\chi(x).$$

Y de entre todas sus diagonales, tiene únicamente $(N+1)$ subdiagonales no nulas,

$$\hat{\omega} = \begin{pmatrix} \check{\omega}_{0,0} & 0 & 0 & \dots & & & \dots & 0 \\ \check{\omega}_{1,0} & \check{\omega}_{1,1} & 0 & \dots & & & \dots & 0 \\ \vdots & \vdots & \ddots & 0 & & & & \\ \check{\omega}_{N,0} & \check{\omega}_{N,1} & & \check{\omega}_{N,N} & 0 & & & \\ 0 & \check{\omega}_{N+1,1} & & \check{\omega}_{N+1,N} & \check{\omega}_{N+1,N+1} & & & \\ 0 & 0 & \ddots & & & \ddots & & \\ \vdots & \vdots & & \check{\omega}_{k,k-N} & & & \check{\omega}_{k,k} & \\ 0 & 0 & \dots & 0 & \ddots & & & \ddots \end{pmatrix}.$$

Especificando: $\check{\omega}_{k,k-N} = \frac{\check{h}_k}{h_{k-N}}$, $\forall k > N$ *y* $\check{\omega}_{k,k} = 1$.

Demostración. La fórmula de conexión para los polinomios es consecuencia directa de la definición de $\breve{\omega}$. La conexión para las funciones de segunda especie es algo más delicada, y se basa en la de los polinomios. En primer lugar recordemos que $C_l(x) = L_\mu\left[\frac{P_l(y)}{x-y}\right]$, volvamos ahora a la fórmula de conexión entre los polinomios para entender que

$$\begin{aligned}(\breve{\omega}C)_l(x) - Q(x)\check{C}_l(x) &= L_\mu\left[\frac{(\breve{\omega}P)_l(y)}{x-y}\right] - Q(x)L_{\check{\mu}}\left[\frac{\check{P}_l(y)}{x-y}\right] \\ &= L_\mu\left[\frac{\check{P}_l(y)}{x-y}\right] - L_{\check{\mu}}\left[\frac{\check{P}_l(y)}{x-y}Q(x)\right] = L_{\check{\mu}}\left[\check{P}_l(y)\frac{Q(y)-Q(x)}{x-y}\right].\end{aligned}$$

En segundo lugar notemos que denotando por $S_n(x,y) := x^n + x^{n-1}y + x^{n-2}y^2 + \cdots + y^n$ al polinomio homogéneo complétamente simétrico en dos variables, y ayudándonos de una matriz de entrelazamiento $\left(\eta_{(n+1)}\right)_{i,j} := \delta_{i,n-j}$ es posible inferir las siguientes propiedades:

$$S_n(x,y) = \chi(y)^\top \eta_{(n+1)}\chi(x) \quad\implies\quad (y-x)S_n(x,y) = \chi(y)^\top\left[\Lambda^\top\eta_{(n+1)} - \eta_{(n+1)}\Lambda\right]\chi(x) = y^n - x^n.$$

De modo que,

$$\frac{Q(y)-Q(x)}{x-y} = -\sum_{n=0}^{N} Q_n\frac{y^n-x^n}{y-x} = -\sum_{n=0}^{N} Q_nS_n(x,y) = -\sum_{n=0}^{N} Q_n\chi(y)^\top\eta_{(n+1)}\chi(x) = -\chi(y)^\top\mathbf{Q}\chi(x).$$

Con lo cual se obtiene lo siguiente:

$$L_{\check{\mu}}\left[\check{P}_l(y)\frac{Q(y)-Q(x)}{x-y}\right] = -L_{\check{\mu}}\left[\check{P}_l(y)\chi(y)^\top\mathbf{Q}\chi(x)\right] = -\check{S}L_{\check{\mu}}\left[\chi(y)\chi(y)^\top\right]\mathbf{Q}\chi(x) = -\check{S}\check{G}\mathbf{Q}\chi(x).$$

□

Es preciso hacer dos definiciones adicionales antes de enunciar la siguiente proposición.

Definición 23. *En primer lugar se definen polinomios reducidos como sigue:*

$$Q_i(x) := \frac{Q(x)}{(x-q_i)^{n_i}}, \qquad\qquad i = 1,2,\dots,s.$$

En segundo lugar, condensando la información de los parámetros libres $\xi_{j,l}$ presentes en la definición de $\check{\mu}$ en una matriz $\Xi \in N\times N$ diagonal por bloques, cada cual siendo $\Xi_j \in n_j\times n_j$ triangular superior , se construyen las siguientes matrices:

$$\Xi := \begin{pmatrix} \Xi_1 & 0 & \dots & 0 \\ 0 & \Xi_1 & 0 & \\ \vdots & 0 & \ddots & \\ 0 & & & \Xi_s \end{pmatrix},$$

$$\Xi_j := \begin{pmatrix} \xi_{j,n_j-1} & \xi_{j,n_j-2} & \cdots & \xi_{j,1} & \xi_{j,0} \\ & \xi_{j,n_j-1} & & & \xi_{j,1} \\ & & \ddots & & \vdots \\ & & & \ddots & \xi_{j,n_j-2} \\ & & & & \xi_{j,n_j-1} \end{pmatrix} \begin{pmatrix} \frac{Q_j^{(0)}(q_j)}{0!} & \frac{Q_j^{(1)}(q_j)}{1!} & \dots & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} & \frac{Q_j^{(n_j-1)}(q_j)}{(n_j-1)!} \\ & \frac{Q_j^{(0)}(q_j)}{0!} & & & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} \\ & & \ddots & & \vdots \\ & & & \ddots & \frac{Q_j^{(1)}(q_j)}{1!} \\ & & & & \frac{Q_j^{(0)}(q_j)}{0!} \end{pmatrix}.$$

Proposición 20. *Los polinomios deformados bajo una transformación de Geronimus son expresables, mediante cuasideterminantes, en términos de los polinomios y las funciones de segunda especie originales: Para $\forall k \geq N$ se tiene,*

$$\check{P}_k = \Theta_* \left(\begin{array}{c|c} \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ \vdots \\ C_{k-1} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k-1} \end{bmatrix} \Xi & \begin{matrix} P_{k-N} \\ \vdots \\ P_{k-1} \end{matrix} \\ \hline \mathcal{J}_q[C_k] - \mathcal{J}_q[P_k]\Xi & P_k(x) \end{array} \right), \quad \check{h}_k(x) = h_{k-N}\Theta_* \left(\begin{array}{c|c} \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ \vdots \\ C_{k-1} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ P_{k+1-N} \\ \vdots \\ P_{k-1} \end{bmatrix} \Xi & \begin{matrix} 1 \\ 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \mathcal{J}_q[C_k] - \mathcal{J}_q[P_k]\Xi & 0 \end{array} \right).$$

Donde $\mathcal{J}_q$ es el correspondiente vector de jets pero en este caso del conjunto $q = \{(q_i, n_i)\}_{i=1}^s$ asociado al polinomio $Q(x)$.

Demostración. Comencemos analizando las funciones de segunda especie deformadas,

$$\check{C}_k(y) = L_{\check{\mu}} \left[\frac{\check{P}_k(x)}{y-x} \right] = L_{\frac{\mu}{Q}} \left[\frac{\check{P}_k(x)}{y-x} \right] + \sum_{j=1}^{s} \sum_{l=1}^{m_j-1} \frac{\xi_{j,l}}{l!} \left(\frac{\check{P}_k(x)}{y-x} \right)^{(l)}_{x=q_j} = L_{\frac{\mu}{Q}} \left[\frac{\check{P}_k(x)}{y-x} \right]$$

$$+ \sum_{j=1}^{s} \begin{pmatrix} \frac{\check{P}_k^{(0)}(q_j)}{0!} & \frac{\check{P}_k^{(1)}(q_j)}{1!} & \dots & \frac{\check{P}_k^{(n_j-1)}(q_j)}{(n_j-1)!} \end{pmatrix} \begin{pmatrix} \xi_{j,n_j-1} & \xi_{j,n_j-2} & \cdots & \xi_{j,1} & \xi_{j,0} \\ & \xi_{j,n_j-1} & & & \xi_{j,1} \\ & & \ddots & & \vdots \\ & & & \ddots & \xi_{j,n_j-2} \\ & & & & \xi_{j,n_j-1} \end{pmatrix} \begin{pmatrix} \left(\frac{1}{y-q_j}\right)^{n_j} \\ \left(\frac{1}{y-q_j}\right)^{n_j-1} \\ \vdots \\ \frac{1}{y-q_j} \end{pmatrix}.$$

Multiplicando esta expresión por $Q(y)$ se tiene que,

$$Q(y)\check{C}_k(y) = Q(y) L_{\frac{\mu}{Q}} \left[\frac{\check{P}_k(x)}{y-x} \right]$$

$$+ \sum_{j=1}^{s} Q_j(y) \begin{pmatrix} \frac{\check{P}_k^{(0)}(q_j)}{0!} & \frac{\check{P}_k^{(1)}(q_j)}{1!} & \dots & \frac{\check{P}_k^{(n_j-1)}(q_j)}{(n_j-1)!} \end{pmatrix} \begin{pmatrix} \xi_{j,n_j-1} & \xi_{j,n_j-2} & \cdots & \xi_{j,1} & \xi_{j,0} \\ & \xi_{j,n_j-1} & & & \xi_{j,1} \\ & & \ddots & & \vdots \\ & & & \ddots & \xi_{j,n_j-2} \\ & & & & \xi_{j,n_j-1} \end{pmatrix} \begin{pmatrix} 1 \\ (y-q_j) \\ \vdots \\ \vdots \\ (y-q_j)^{n_j-1} \end{pmatrix}.$$

Tomando ahora el límite $y \to q_j$ nos encontraremos con el siguiente desarrollo de Taylor,

$$Q(y)\check{C}_k(y) = \begin{pmatrix} \frac{\check{P}_k^{(0)}(q_j)}{0!} & \frac{\check{P}_k^{(1)}(q_j)}{1!} & \dots & \frac{\check{P}_k^{(n_j-1)}(q_j)}{(n_j-1)!} \end{pmatrix} \Xi_j \begin{pmatrix} 1 \\ (y-q_j) \\ \vdots \\ (y-q_j)^{n_j-1} \end{pmatrix} + O(y-q_j)^{n_j}.$$

Este proceso podría haberse realizado para todos y cada uno de los subíndices j. Tras hacerlo y condensando todos los resultados en una matriz escribimos,

$$\mathcal{J}_q[Q\check{C}_k] = \mathcal{J}_q[\check{P}_k]\Xi.$$

Volviendo a la fórmula de conexión para las funciones de segunda especie y dejando a $\mathcal{J}_q$ actuar a ambos lados,

$$\check{\omega}\mathcal{J}_q[C] = \mathcal{J}_q[Q\check{C}] - \check{H}\left(\check{S}^{-1}\right)^\top \mathbf{Q}\mathcal{J}_q[\chi(x)] \qquad \Longrightarrow \qquad \check{\omega}\mathcal{J}_q[C] = \mathcal{J}_q[\check{P}]\Xi - \check{H}\left(\check{S}^{-1}\right)^\top \mathbf{Q}\mathcal{J}_q[\chi(x)].$$

Reordenando términos y empleando la fórmula de conexión de los polinomios,

$$\check{\omega}\left(\mathcal{J}_q[C]-\mathcal{J}_q[P]\Xi\right)=-\check{H}\left(\check{S}^{-1}\right)^{\top}\mathbf{Q}\mathcal{J}_q[\chi(x)].$$

Con lo que encontramos que,

$$\begin{pmatrix}\check{\omega}_{k,k-N} & \dots & \check{\omega}_{k,k-1} & 1\end{pmatrix}\left(\mathcal{J}_q\begin{bmatrix}C_{k-N}\\ C_{k-N+1}\\ \vdots\\ C_k\end{bmatrix}-\mathcal{J}_q\begin{bmatrix}P_{k-N}\\ P_{k-N+1}\\ \vdots\\ P_k\end{bmatrix}\Xi\right)=0, \qquad \forall k\geq N.$$

Desde donde se deduce finalmente el resultado de la proposición. $\square$

Proposición 21. *Los núcleos deformados bajo una transformación de Geronimus están relacionados con los originales:*

$$\check{K}^{[k]}(x,y)=Q(x)K^{[k]}(x,y)-$$
$$\begin{pmatrix}\check{P}_k(x) & \dots & \check{P}_{k+N-1}\end{pmatrix}\begin{pmatrix}\check{h}_k^{-1} & & & \\ & \check{h}_{k+1}^{-1} & & \\ & & \ddots & \\ & & & \check{h}_{k+N-1}^{-1}\end{pmatrix}\begin{pmatrix}\check{\omega}_{k,k-N} & \check{\omega}_{k,k-N+1} & \dots & \check{\omega}_{k,k-1}\\ & \check{\omega}_{k+1,k+1-N} & & \vdots\\ & & \ddots & \vdots\\ & & & \check{\omega}_{k+N-1,k-1}\end{pmatrix}\begin{pmatrix}P_{k-N}\\ P_{k+1-N}\\ \vdots\\ P_{k-1}\end{pmatrix}.$$

Demostración. Como viene siendo ya habitual al encontrar este tipo de expresiones, basta con tener en cuenta las fórmulas de conexión para probarlas,

$$\check{H}^{-1}\check{\omega}P(x)=\check{H}^{-1}\check{P}(x), \qquad \check{P}^{\top}(x)\check{H}^{-1}\check{\omega}=Q(x)P^{\top}(x)H^{-1}.$$

$\square$

Un razonamiento análogo pero en base al núcleo mixto de Christoffel–Darboux permite enunciar la siguiente proposición.

Proposición 22. *Los núcleos mixtos transformados y originales $\forall k\geq N$, están relacionados como sigue:*

$$\begin{aligned}&Q(x)\mathcal{K}^{[k]}(x,y)\\ &-\begin{pmatrix}\check{P}_k(x) & \dots & \check{P}_{k+N-1}(x)\end{pmatrix}\begin{pmatrix}\check{h}_k^{-1} & & & \\ & \check{h}_{k+1}^{-1} & & \\ & & \ddots & \\ & & & \check{h}_{k+N-1}^{-1}\end{pmatrix}\begin{pmatrix}\check{\omega}_{k,k-N} & \dots & \check{\omega}_{k,k-1}\\ & \ddots & \vdots\\ & & \check{\omega}_{k+N-1,k-1}\end{pmatrix}\begin{pmatrix}C_{k-N}(y)\\ C_{k+1-N}(y)\\ \dots\\ C_{k-1}(y)\end{pmatrix}\\ &=Q(y)\check{\mathcal{K}}^{[k]}(x,y)-\left(\chi^{[N]}(x)\right)^{\top}\mathbf{Q}\chi^{[N]}(y).\end{aligned}$$

Demostración. Nuevamente, basta con recordar las fórmulas de conexión para probar el resultado.

$$\check{H}^{-1}\check{\omega}C(y)=Q(y)\check{H}^{-1}\check{C}(y)-(\check{S}^{-1})^{\top}\mathbf{Q}\chi(y) \qquad \check{P}^{\top}(x)\check{H}^{-1}\check{\omega}=Q(x)P^{\top}(x)H^{-1}.$$

$\square$

La combinación de los resultados de estas dos últimas proposiciones y de manera acorde al espíritu del resto de la sección tiene por consecuencia la siguiente proposición.

Proposición 23. *La siguiente fórmula alternativa, para los polinomios deformados bajo una transformación de Geronimus, en términos de los núcleos y los núcleos mixtos de Christoffel–Darboux se cumple:*

$$\frac{\check{P}_k(x)}{\check{h}_k} = \Theta_* \left(\begin{array}{c|c} \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ \vdots \\ C_{k-1} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k-1} \end{bmatrix} \Xi & \begin{matrix} 1 \\ \vdots \\ 0 \end{matrix} \\ \hline Q(x)\left(\mathcal{J}_q^{(2)}[\mathcal{K}^{[k]}(x,y)] - \mathcal{J}_q^{(2)}[K^{[k]}(x,y)]\Xi\right) + \left(\chi^{[N]}(x)\right)^\top \mathbf{Q}\mathcal{J}_q[\chi^{[N]}] & 0 \end{array} \right).$$

Antes de finalizar con el análisis de esta deformación, volvamos a la definición 22 y demos nombre a la matriz $SQ(\Lambda)\check{S}^{-1}$.

Definición 24. *Se define la siguiente matriz triangular superior relacionada con el conector de la transformación de Geronimus:*

$$\check{\Omega} := SQ(\Lambda)\check{S}^{-1}, \qquad \check{\omega} = \check{H}\check{\Omega}^\top H^{-1}.$$

La pareja de matrices triangulares inferior y superior $\check{\omega}$ y $\check{\Omega}$ respectivamente, son relevantes por ser los factores matriciales de las siguientes factorizaciones.

Proposición 24. *Las siguientes factorizaciones UL y LU que involucran las matrices de Jacobi son ciertas:*

$$\check{\Omega}\check{\omega} = Q(J), \qquad \check{\omega}\check{\Omega} = Q(\check{J}).$$

Demostración. Para probarlo basta con emplear las definiciones de cada elemento involucrado en términos de las matrices de la factorización. □

La transformación espectral lineal

Esta transformación que está a punto de ser analizada contiene como casos particulares a las dos anteriores. Inicialmente estudiada por Vasily Uvarov en [124] y por Alexei Zhedanov en [131] desde la perspectiva de la transformada de Stieltjes de la medida. Me interesaré por la composición en el orden $\tilde{\mu} := \widehat{\check{\mu}}$ en lugar de la que se sigue al componer en orden opuesto $\check{\hat{\mu}}$ que se deja para la sección próxima. El primer paso consiste en relacionar las matrices de momentos del caso deformado y el de partida.

$$Q\tilde{\mu} = R\mu \qquad \Longrightarrow \qquad \tilde{G}Q(\Lambda^\top) = R(\Lambda)G.$$

El asumir la existencia de la factorización LU de ambas matrices de momentos sugiere la siguiente definición.

Definición 25. *El conector de la transformación espectral lineal es:*

$$\tilde{\omega} := \tilde{H}\left(\tilde{S}^{-1}\right)^\top Q(\Lambda^\top)S^\top H^{-1} = \tilde{S}R(\Lambda)S^{-1}.$$

Proposición 25. *El conector permite relacionar los elementos deformados con los originales:*

$$\tilde{\omega}P(x) = R(x)\tilde{P}(x), \qquad \tilde{\omega}C(x) = Q(x)\tilde{C}(x) - \tilde{H}\left(\tilde{S}^{-1}\right)^\top \mathbf{Q}\chi(x).$$

Tiene únicamente M superdiagonales y N subdiagonales no nulas,

$$\tilde{\omega} = \begin{pmatrix} \tilde{\omega}_{0,0} & \tilde{\omega}_{0,1} & \dots & & \tilde{\omega}_{0,M} & 0 & & & \dots & 0 \\ \tilde{\omega}_{1,0} & \tilde{\omega}_{1,1} & & & & \tilde{\omega}_{1,M+1} & & & & \vdots \\ \vdots & \vdots & \ddots & & & & \ddots & 0 & & \\ \tilde{\omega}_{N,0} & \tilde{\omega}_{N,1} & & \tilde{\omega}_{N,N} & & & & \tilde{\omega}_{N,N+M} & 0 & \\ 0 & \tilde{\omega}_{N+1,1} & & & & & & & \tilde{\omega}_{N+1,N+M+1} & \\ & 0 & \ddots & & & \ddots & & & \ddots & 0 \\ \vdots & & & \tilde{\omega}_{k,k-N} & & & \tilde{\omega}_{k,k} & & & \tilde{\omega}_{k,k+M+1} \\ 0 & \dots & & 0 & \ddots & & & \ddots & & \end{pmatrix},$$

y entradas explícitas destacadas son: $\tilde{\omega}_{k,k+M} = 1$ *y* $\tilde{\omega}_{k-N,k} = \frac{\tilde{h}_k}{h_{k-N}}$ $\forall k \geq N$.

Demostración. Ambos resultados se prueban desde la definición 25. Para probar la primera igualdad basta con usar la segunda expresión, $\tilde{S}R(\Lambda)S^{-1}$, de la definición del conector, mientras que para la segunda igualdad es necesario usar la primera, $\tilde{H}\left(\tilde{S}^{-1}\right)^{\top} Q(\Lambda^{\top})S^{\top}H^{-1}$. La segunda parte de la proposición es consecuencia necesaria para que las dos expresiones alternativas en la definición 25 sean compatibles. □

Proposición 26. *Los polinomios deformados bajo una transformación espectral lineal pueden expresarse en términos de los polinomios y funciones de segunda especie originales mediante el uso de cuasideterminantes. Para* $\forall k \geq N$ *se tiene,*

$$\tilde{P}_k = \frac{1}{R(x)}\Theta_* \left(\begin{array}{c|c} \mathcal{J}_r \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k+M-1} \end{bmatrix}, \quad \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ \vdots \\ C_{k+M-1} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k+M-1} \end{bmatrix} \Xi & \begin{matrix} P_{k-N}(x) \\ \vdots \\ P_{k+M-1}(x) \end{matrix} \\ \hline \mathcal{J}_r[P_{k+M}], \quad \mathcal{J}_q[C_{k+M}] - \mathcal{J}_q[P_{k+M}]\Xi & P_{k+M}(x) \end{array}\right),$$

$$\tilde{h}_k(x) = h_{k-N}\Theta_* \left(\begin{array}{c|c} \mathcal{J}_r \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k+M-1} \end{bmatrix}, \quad \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ \vdots \\ C_{k+M-1} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ \vdots \\ P_{k+M-1} \end{bmatrix} \Xi & \begin{matrix} 1 \\ 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \mathcal{J}_r[P_{k+M}], \quad \mathcal{J}_q[C_{k+M}] - \mathcal{J}_q[P_{k+M}]\Xi & 0 \end{array}\right).$$

Donde los subíndices que acompañan a $\mathcal{J}_r, \mathcal{J}_q$ *sirven para especificar el conjunto de ceros y multiplicidades a los que hacen referencia* $r = \{(r_i, m_i)\}_{i=1}^d$ *o* $q = \{(q_j, n_j)\}_{j=1}^s$.

Demostración. Una vez las transformaciones de Christoffel y Geronimus han sido estudiadas, debiera no causar excesiva sorpresa que al tomar el límite $y \to q_j$ se obtenga el siguiente desarrollo de Taylor:

$$Q(y)\tilde{C}_k(y) = \begin{pmatrix} \frac{(R(x)\tilde{P}_k(x))^{(0)}_{x=q_j}}{0!} & \frac{(R(x)\tilde{P}_k(x))^{(1)}_{x=q_j}}{1!} & \dots & \frac{(R(x)\tilde{P}_k(x))^{(n_j-1)}_{x=q_j}}{(n_j-1)!} \end{pmatrix} \Xi_j \begin{pmatrix} 1 \\ (y-q_j) \\ \vdots \\ (y-q_j)^{n_j-1} \end{pmatrix} + O(y-q_j)^{n_j}.$$

Considerando esta ecuación para todos los posibles j,

$$\mathcal{J}_q[Q\tilde{C}_k] = \mathcal{J}_q[(R(x)\tilde{P}(x))]\Xi.$$

Volviendo ahora a la fórmula de conexión para las funciones de segunda especie y dejando a $\mathcal{J}_q$ actuar a ambos lados de la igualdad,

$$\tilde{\omega}\mathcal{J}_q[C] = \mathcal{J}_q[Q\tilde{C}] - \tilde{H}\left(\tilde{S}^{-1}\right)^{\top} \mathbf{Q}\mathcal{J}_q[\chi(x)],$$

$$\tilde{\omega}\mathcal{J}_q[C] = \mathcal{J}_q[R\tilde{P}]\Xi - \tilde{H}\left(\tilde{S}^{-1}\right)^{\top} \mathbf{Q}\mathcal{J}_q[\chi(x)],$$

$$\tilde{\omega}\left(\mathcal{J}_q[C] - \mathcal{J}_q[P]\Xi\right) = -\tilde{H}\left(\tilde{S}^{-1}\right)^{\top} \mathbf{Q}\mathcal{J}_q[\chi(x)].$$

En la última línea ha sido necesaria la fórmula de conexión para los polinomios, que además satisfacen,

$$\tilde{\omega}\mathcal{J}_r[P] = \mathcal{J}_r[R\tilde{P}] = 0.$$

Agrupando ambos resultados en una misma ecuación queda,

$$\begin{pmatrix} \check{\omega}_{k,k-N} & \dots & \check{\omega}_{k,k+M-1} & 1 \end{pmatrix} \left(\mathcal{J}_r \begin{bmatrix} P_{k-N} \\ P_{k-N+1} \\ \vdots \\ P_{k+M} \end{bmatrix}, \quad \mathcal{J}_q \begin{bmatrix} C_{k-N} \\ C_{k-N+1} \\ \vdots \\ C_{k+M} \end{bmatrix} - \mathcal{J}_q \begin{bmatrix} P_{k-N} \\ P_{k-N+1} \\ \vdots \\ P_{k+M} \end{bmatrix} \Xi \right) = 0, \qquad \forall k \geq N.$$

Desde donde el resultado de la proposición se deduce. □

Proposición 27. *Núcleos deformados y originales están relacionados entre sí como sigue:*

$$R(y)\tilde{K}^{[k]}(x,y)-\begin{pmatrix}\tilde{P}_{k-M} & \dots & \tilde{P}_{k-1}\end{pmatrix}\begin{pmatrix}\tilde{h}_{k-M}^{-1} & & \\ & \ddots & \\ & & \tilde{h}_{k-1}^{-1}\end{pmatrix}\begin{pmatrix}\tilde{\omega}_{k-M,k} & & 0\\ \vdots & \ddots & \\ \tilde{\omega}_{k-1,k} & \dots & \tilde{\omega}_{k-1,k+M-1}\end{pmatrix}\begin{pmatrix}P_k(y)\\ \vdots \\ P_{k+M-1}(y)\end{pmatrix}$$
$$=Q(x)K^{[k]}(x,y)-\begin{pmatrix}\tilde{P}_k(x) & \dots & \tilde{P}_{k+N-1}\end{pmatrix}\begin{pmatrix}\tilde{h}_{k}^{-1} & & \\ & \ddots & \\ & & \tilde{h}_{k+N-1}^{-1}\end{pmatrix}\begin{pmatrix}\tilde{\omega}_{k,k-N} & \dots & \tilde{\omega}_{k,k-1}\\ & \ddots & \vdots \\ & & \tilde{\omega}_{k+N-1,k-1}\end{pmatrix}\begin{pmatrix}P_{k-N}\\ \vdots \\ P_{k-1}\end{pmatrix}.$$

Demostración. La expresión la podemos deducir desde las fórmulas de conexión,

$$\tilde{H}^{-1}\tilde{\omega}P(x)=R(x)\tilde{H}^{-1}\tilde{P}(x), \qquad \tilde{P}^{\top}(x)\tilde{H}^{-1}\tilde{\omega}=Q(x)P^{\top}(x)H^{-1}.$$

□

1.5.2 Perturbaciones aditivas

Conocida la secuencia ortogonal asociada a un funcional μ uno podría interesarse, tomando otro funcional ν (compartiendo o no soporte con el anterior), por el funcional resultante de la adición de ambos $\breve{\mu}=\mu+\nu$. Tras las secciones anteriores, resulta ahora natural preguntarse por la existencia de una posible secuencia de polinomios ortogonales asociada a este nuevo funcional. Las observaciones que se siguen de emplear las técnicas de la factorización LU vuelven a ser esclarecedoras a la hora de conectar, en caso de existir, ambas secuencias. Bien es cierto que la conexión existente en el caso general es bastante esperable dando cuenta de un simple cambio de base, pero resulta también interesante la existencia de casos particulares en los que dicha conexión puede simplificarse drásticamente permitiendo dar cualquier elemento de la nueva secuencia en términos de un número constante de elementos de la original. Analizaremos aquí uno de estos casos particulares; la transformación de Uvarov del funcional.
En primer lugar, desde $\breve{\mu}=\mu+\nu$, y construyendo la matriz de Gram de cada funcional involucrado se puede escribir $\breve{G}=G+G_\nu$. Dado que partimos de una secuencia de polinomios ortogonal conocida, la factorización LU de G está garantizada. Impondremos sobre $\breve{G}$ la condición de admitir también una factorización LU, dado que nuestro propósito es encontrar una secuencia ortogonal asociada a la misma. Finalmente, no le exigiremos directamente nada a G_ν. Hechas estas consideraciones, y factorizando aquellas matrices que lo permiten,

$$\breve{G}=G+G_\nu \qquad \Longrightarrow \qquad \breve{H}\left(\breve{S}^{-1}\right)^{\top}S^{\top}=\breve{S}S^{-1}H+\breve{S}G_\nu S^{\top}.$$

La expresión anterior da pie a las siguientes definiciones:

Definición 26. *Se consideran las siguientes matrices:*

$$M:=\breve{S}S^{-1}, \qquad A:=SG_\nu S^{\top}.$$

Proposición 28. *La matriz M conecta los polinomios nuevos con los originales,*

$$MP(x)=\breve{P}(x).$$

En términos de estas matrices definidas, la expresión de partida se escribe como sigue,

$$\breve{H}\left(\breve{S}^{-1}\right)^{\top}S^{\top}=M(H+A) \qquad \Longrightarrow \qquad M^{-1}\breve{H}M^{-\top}=(H+A).$$

Esta relación puede interpretarse como una factorización LU de la matriz $(H+A)$, y por tanto, tenemos a nuestra disposición las expresiones en términos de cuasideterminantes para las entradas de las matrices de dicha factorización, es decir,

Proposición 29. *La secuencia de polinomios ortogonales y sus normas al cuadrado, ambas asociadas a la perturbación aditiva de μ mediante ν, se pueden escribir usando cuasideterminantes:*

$$\breve{P}_k(x) = \Theta_* \left[\begin{array}{cccc|c} & & & & P_0(x) \\ & (H+A)^{[k]} & & & P_1(x) \\ & & & & \vdots \\ & & & & P_{k-1}(x) \\ \hline (H+A)_{k,0} & (H+A)_{k,1} & \dots & (H+A)_{k,k-1} & P_k(x) \end{array}\right],$$

$$\breve{h}_k = \Theta_* \left[\begin{array}{cccc|c} & & & & (H+A)_{0,k} \\ & (H+A)^{[k]} & & & (H+A)_{1,k} \\ & & & & \vdots \\ & & & & (H+A)_{k-1,k} \\ \hline (H+A)_{k,0} & (H+A)_{k,1} & \dots & (H+A)_{k,k-1} & (H+A)_{k,k} \end{array}\right].$$

Véase cómo a la hora de expresar el polinomio k-*ésimo* de la nueva secuencia, precisamos de los primeros $k+1$ polinomios de la secuencia original.

La transformación de Uvarov

Usaremos ahora el resultado de la sección anterior para un caso particular en el que la perturbación aditiva ν está compuesta por una suma de deltas de Dirac y derivadas de esta (no podremos hablar de medidas en esta ocasión tampoco) en una serie de puntos. Fue V.B. Uvarov en [124] el primero en interesarse por este tipo de deformación.

Definición 27. *Se denota por transformación de Uvarov a la siguiente perturbación aditiva del funcional,*

$$\breve{\mu}(x) = \mu(x) + \nu(x), \qquad \nu = \sum_{i=0}^{s} \sum_{m=1}^{n_i-1} (-1)^m \delta^{(m)}(x-q_i)\xi_{i,m}, \qquad n_1 \leq n_2 \leq n_3 \leq \dots \leq n_s.$$

Inspirados por las secciones anteriores, incluimos a continuación las definiciones adaptadas a este caso.

Definición 28. *Dado un conjunto de tuplas, $\{(q_i, n_i)\}_{i=1}^s$ se define para una función dada $f(x)$, el vector*[6] *$\Pi[f] : \mathcal{F}(x) \longrightarrow \mathbb{R}^M$ como sigue:*

$$\Pi[f] := \Big(f^{(0)}(q_1), f^{(1)}(q_1), \dots, f^{(n_1-1)}(q_1); f^{(0)}(q_2), \dots, f^{(n_2-1)}(q_2); \dots; f^{(0)}(q_s), \dots, f^{(n_s-1)}(q_s)\Big).$$

Esta definición permite, denotando por $\left(K^{[k]}(q_i,q_j)\right)^{(t,d)} := \frac{\partial^{t+d} K^{[k]}(x,y)}{\partial x^t \partial y^d}\Big|_{(x,y)=(q_i,q_j)}$, dar la siguiente

Definición 29. *Las entradas de la siguiente matriz de tamaño $(\sum_{i=1}^s n_i) \times (\sum_{i=1}^s n_i)$ involucran la evaluación en los puntos $\{q_i\}_{i=1}^s$ de las derivadas hasta el orden $\{(n_i-1)\}_{i=1}^s$ del núcleo de Chrisfoffel–Darboux.*

$$\mathbb{K}^{[k]} := \Pi[P^{[k]}]^\top \left(H^{[k]}\right)^{-1} \left(\Pi[P^{[k]}]\right) = \begin{pmatrix} \mathbb{K}^{[k]}_{[1][1]} & \mathbb{K}^{[k]}_{[1][2]} & \dots & \mathbb{K}^{[k]}_{[1][s]} \\ \mathbb{K}^{[k]}_{[2][1]} & \mathbb{K}^{[k]}_{[2][2]} & \dots & \mathbb{K}^{[k]}_{[2][s]} \\ & & & \\ \mathbb{K}^{[k]}_{[s][1]} & \mathbb{K}^{[k]}_{[s][2]} & \dots & \mathbb{K}^{[k]}_{[s][s]} \end{pmatrix},$$

$$\mathbb{K}^{[k]}_{[i][j]} := \begin{pmatrix} \left(K^{[k]}(q_i,q_j)\right)^{(0,0)} & \left(K^{[k]}(q_i,q_j)\right)^{(0,1)} & \dots & \left(K^{[k]}(q_i,q_j)\right)^{(0,n_j-1)} \\ \left(K^{[k]}(q_i,q_j)\right)^{(1,0)} & \left(K^{[k]}(q_i,q_j)\right)^{(1,1)} & \dots & \left(K^{[k]}(q_i,q_j)\right)^{(1,n_j-1)} \\ & & & \\ \left(K^{[k]}(q_i,q_j)\right)^{(n_i-1,0)} & \left(K^{[k]}(q_i,q_j)\right)^{(n_i-1,1)} & \dots & \left(K^{[k]}(q_i,q_j)\right)^{(n_i-1,n_j-1)} \end{pmatrix}.$$

[6] Relacionado con el vector de jets $\mathcal{J}[f]$ pero sin la presencia de los factoriales en los denominadores.

Definición 30. *Dados los parámetros $\{\xi_{i,m}\}_m$ se construye la siguiente matriz $\xi^{(i)} \in n_i \times n_i$:*

$$\xi^{(i)} := \begin{pmatrix} \xi_{i,0} & \binom{1}{0}\xi_{i,1} & \binom{2}{0}\xi_{i,2} & \binom{3}{0}\xi_{i,3} & \dots & \binom{n_i-1}{0}\xi_{i,n_i-1} \\ \binom{1}{1}\xi_{i,1} & \binom{2}{1}\xi_{i,2} & \binom{3}{1}\xi_{i,3} & & & 0 \\ \binom{2}{2}\xi_{i,2} & \binom{3}{2}\xi_{i,3} & & & & \vdots \\ \binom{3}{3}\xi_{i,3} & \vdots & & & & \\ \vdots & \binom{n_i-1}{n_i-2}\xi_{i,n_i-1} & & & & \vdots \\ \binom{n_i-1}{n_i-1}\xi_{i,n_i-1} & 0 & \dots & & \dots & 0 \end{pmatrix}.$$

A continuación se sitúan estas a lo largo de la diagonal de una de mayor tamaño $(\sum_{i=1}^{s} n_i) \times (\sum_{i=1}^{s} n_i)$*,*

$$\Xi := \begin{pmatrix} \xi^{(1)} & & & \\ & \xi^{(2)} & & \\ & & \ddots & \\ & & & \xi^{(s)} \end{pmatrix}.$$

Todas estas definiciones son útiles a la hora de discutir la transformación de Uvarov. Nótese en primer lugar que,

$$\breve{G} = G + \Pi[\chi]\Xi\left(\Pi[\chi]\right)^\top, \qquad A = S\Pi[\chi]\Xi\left(\Pi[\chi]\right)^\top S^\top = \Pi[P]\Xi\left(\Pi[P]\right)^\top.$$

Proposición 30. *La siguiente fórmula se cumple,*

$$\left[(H+A)^{[k]}\right]^{-1} = (H^{[k]})^{-1} - (H^{[k]})^{-1}\Pi[P^{[k]}]\Xi\left(\mathbb{I} + \mathbb{K}^{[k]}\Xi\right)^{-1}\left(\Pi[P^{[k]}]\right)^\top (H^{[k]})^{-1}.$$

donde $(\mathbb{I} \in (\sum_{i=1}^{s} n_i) \times (\sum_{i=1}^{s} n_i))$*.*

Demostración. Es sencillo seguir las siguientes igualdades,

$$\begin{aligned} \left[(H+A)^{[k]}\right]^{-1} &= (H^{[k]})^{-1}\left[\left(\mathbb{I} + AH^{-1}\right)^{[k]}\right]^{-1} = (H^{[k]})^{-1}\left(\mathbb{I} + \Pi[P^{[k]}]\Xi(\Pi[P^{[k]}])^\top (H^{[k]})^{-1}\right)^{-1} \\ &= (H^{[k]})^{-1}\Big(\mathbb{I} - \Pi[P^{[k]}]\Xi(\Pi[P^{[k]}])^\top (H^{[k]})^{-1} \\ &\qquad + \Pi[P^{[k]}]\Xi(\Pi[P^{[k]}])^\top (H^{[k]})^{-1}\Pi[P^{[k]}]\Xi(\Pi[P^{[k]}])^\top (H^{[k]})^{-1} - \dots\Big) \\ &= (H^{[k]})^{-1} - (H^{[k]})^{-1}\Pi[P^{[k]}]\Xi\left(\mathbb{I} - \mathbb{K}^{[k]}\Xi + (\mathbb{K}^{[k]}\Xi)^2 - \dots\right)(\Pi[P^{[k]}])^\top (H^{[k]})^{-1}. \end{aligned}$$

Lo que coincide con la expresión de la proposición pero en forma de serie. □

El resultado principal de la sección es:

Proposición 31. *La secuencia de polinomios ortogonales asociados a la transformación de Uvarov de un funcional original viene dada por las siguientes fórmulas en términos de cuasideterminantes:*

$$\breve{P}_k(x) = \left(\begin{array}{c|c} \mathbb{I} + \mathbb{K}^{[k]}\Xi & \left(\Pi^{(1)}[K^{[k]}(y,x)]\right)^\top \\ \hline \Pi[P_k]\Xi & P_k(x) \end{array}\right), \qquad \breve{h}_k(x) = \left(\begin{array}{c|c} \mathbb{I} + \mathbb{K}^{[k]}\Xi & \mathbb{K}^{[k]}\Xi\left(\Pi[P_k]\right)^\top \\ \hline \Pi[P_k]\Xi & h_k + \Pi[P_k]\Xi\left(\Pi[P_k]\right)^\top \end{array}\right).$$

Nótese que lo que hace de estas unas expresiones realmente útiles es que, en caso de tener la fórmula de Christoffel–Darboux a nuestra disposición (1.5), a la hora de expresar el polinomio transformado *k-ésimo*, sólo serán necesarios dos polinomios consecutivos de la secuencia original.

Demostración. Desde el resultado para una transformación aditiva general y la reciente expresión para la inversa de $(H+A)$, se obtiene para los polinomios:

$$\begin{aligned}\breve{P}_k(x) &= P_k(x) - \Pi[P_k]\Xi\left(\Pi[P^{[k]}]\right)^\top\left[(H^{[k]})^{-1} - (H^{[k]})^{-1}\Pi[P^{[k]}]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}\left(\Pi[P^{[k]}]\right)^\top(H^{[k]})^{-1}\right]P^{[k]}(x)\\ &= P_k(x) - \Pi[P_k]\Xi\left[(\Pi^{(1)}[K(y,x)])^\top - \mathbb{K}^{[k]}\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}(\Pi^{(1)}[K(y,x)])^\top\right]\\ &= P_k(x) - \Pi[P_k]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}(\Pi^{(1)}[K(y,x)])^\top.\end{aligned}$$

Igualmente para las normas:

$$\begin{aligned}\breve{h}_k &= \left(h_k + \Pi[P_k]\Xi(\Pi[P_k])^\top\right) - \Pi[P_k]\Xi\left(\Pi[P^{[k]}]\right)^\top\Big[(H^{[k]})^{-1}\\ &\qquad - (H^{[k]})^{-1}\Pi[P^{[k]}]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}\left(\Pi[P^{[k]}]\right)^\top(H^{[k]})^{-1}\Big]\Pi[P]\Xi(\Pi[P])^\top\\ &= h_k + \Pi[P_k]\Xi(\Pi[P_k])^\top - \Pi[P_k]\Xi\left(\mathbb{K}^{[k]} - \mathbb{K}^{[k]}\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}\mathbb{K}^{[k]}\right)\Xi(\Pi[P_k])^\top\\ &= h_k + \Pi[P_k]\Xi(\Pi[P_k])^\top - \Pi[P_k]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}\mathbb{K}^{[k]}\Xi(\Pi[P_k])^\top.\end{aligned}$$

□

En el caso de ser un único punto el encargado de definir la perturbación obtenemos, a modo de ejemplo, las conocidas expresiones,

$$\begin{aligned}\breve{P}_k(x) &= \Theta_*\left(\begin{array}{c|c}1+\xi K^{[k]}(a,a) & K^{[k]}(a,x)\\ \hline \xi P_k(a) & P_k(x)\end{array}\right) = P_k(x) - \xi P_k(a)\frac{K^{[k]}(a,x)}{1+\xi K^{[k]}(a,a)},\\ \breve{h}_k &= \Theta_*\left(\begin{array}{c|c}1+\xi K^{[k]}(a,a) & \xi K^{[k]}(a,a)p_k(a)\\ \hline \xi P_k(a) & h_k+\xi p_k^2(a)\end{array}\right) = h_k + \frac{\xi p_k^2(a)}{1+\xi K^{[k]}(a,a)}.\end{aligned}$$

1.6 Deformaciones continuas. Sistemas integrables

En esta sección trataré de explicar la conexión que existe entre los polinomios ortogonales y la red de Toda[7] (caso particular de la red de Fermi-Pasta-Ulam [61]): una sucesión de partículas con masa la unidad, situadas en los puntos x_n, y unidas entre sí mediante muelles cuya energía potencial U viene dada por la exponencial de la separación entre estas (por lo tanto, no lineal),

$$\frac{\mathrm{d}^2x_n}{\mathrm{d}t^2} = U'(x_{n+1}-x_n) - U'(x_n - x_{n-1}) = e^{(x_{n+1}-x_n)} - e^{(x_n-x_{n-1})}. \tag{1.14}$$

Para establecer la conexión prometida, se habrán de introducir un conjunto de parámetros a los que llamaremos tiempos, encargados estos de parametrizar un nuevo tipo de deformación de la matriz de momentos. Esta dependencia de los parámetros la hemos de interpretar como si de una evolución temporal se tratara, no solo para la matriz de momentos, sino para todos los elementos que se construyen a partir de la misma:

[7] Fue el físico Morizaku Toda [122] quien propuso el sistema integrable que se va a considerar. Posteriormente se observó que dicho sistema es equivalente a la reducción unidimensional de una construcción geométrica propuesta por el matemático Jean Gaston Darboux [48].

polinomios ortogonales, normas, matriz de Jacobi, etc. Precisamente de la derivada temporal la matriz de Jacobi se obtiene un par de Lax equivalente a la ecuación de Toda, y por lo tanto que permite relacionar las x_n con los principales elementos de la teoría de los polinomios ortogonales (h_n y $S_{n+1,n}$) utilizando un cambio de variables propuesto por H. Flaschka.
Con el único pretexto de conseguir una conexión lo más simplificada y clara posible, en esta sección emplearemos polinomios ortonormales $p_k(x)$ en lugar de los mónicos $P_k(x)$ que veníamos usando hasta ahora. Para conseguir esto basta con reescribir la matriz diagonal de la factorización como $H = \sqrt{H}\sqrt{H}$ [8] y multiplicar estos factores por las matrices unitriangulares de la factorización, es decir,

$$G = S^{-1}HS^{-\top} = Z^{-1}Z^{-\top}, \qquad Z := \left(\sqrt{H}\right)^{-1}S, \qquad p(x) := Z\chi(x), \qquad \langle p_k, p_j\rangle_\mu = \delta_{k,j}.$$

Es interesante y sencillo comprobar que en este caso, la matriz de Jacobi para los polinomios ortonormales (que se construye desde la simetría $\Lambda G = G\Lambda^\top$ pero empleando esta vez la factorización con los factores Z) y que vamos a denotar por L, (es decir, $Lp_n(x) = xp_n(x)$) es una matriz simétrica cuyas entradas son:

$$L = L^\top := Z\Lambda Z^{-1} = \left(\sqrt{H}\right)^{-1}J\sqrt{H} = \begin{pmatrix} -S_{10} & \sqrt{h_1h_0^{-1}} & 0 & 0 & 0 & \dots \\ \sqrt{h_1h_0^{-1}} & S_{10}-S_{21} & \sqrt{h_2h_1^{-1}} & 0 & 0 & \dots \\ 0 & \sqrt{h_2h_1^{-1}} & S_{21}-S_{32} & \sqrt{h_3h_2^{-1}} & 0 & \dots \\ 0 & 0 & \sqrt{h_3h_2^{-1}} & S_{32}-S_{43} & \sqrt{h_4h_3^{-1}} & \dots \\ 0 & 0 & 0 & \sqrt{h_4h_3^{-1}} & S_{43}-S_{54} & \\ \vdots & \vdots & \vdots & \vdots & & \ddots \end{pmatrix}.$$

1.6.1 Deformación continua de la matriz de momentos

Se introducen dos conjuntos infinitos de parámetros reales $\left\{\{t_{\alpha,j}\}_{j=0}^{\infty}\right\}_{\alpha=1,2}$, con $t_{\alpha,0} = 0$ por conveniencia, de los que va a depender la deformación.

Definición 31. *La matriz de momentos dependiente del tiempo es,*

$$G(t) := W_0(t_1)G\big(W_0(t_2)\big)^\top, \qquad W_0(t_\alpha) := e^{\sum_j t_{\alpha,j}\Lambda^j}.$$

Nótese cómo esta deformación tiene por condición inicial la matriz de momentos original.

Proposición 32. *La matriz de momentos dependiente del tiempo $G(t)$ es la correspondiente a la siguiente medida con dependencia temporal:* $\mathrm{d}\mu(x,t) = e^{\sum_j (t_{1,j}+t_{2,j})x^j}\mathrm{d}\mu(x)$.

Demostración. La prueba es sencilla desde la deficinición de $G(t)$,

$$G(t) = W_0(t_1)G\big(W_0(t_2)\big)^\top = e^{\sum_j t_{1,j}\Lambda^j}\left[\int \chi(x)\mathrm{d}\mu(x)\chi(x)^\top\right]\left(e^{\sum_j t_{2,j}\Lambda^j}\right)^\top,$$

y la propiedad $\Lambda\chi = x\chi$, de forma que $e^{\sum_j t_{1,j}\Lambda^j}\chi = e^{\sum_j t_{1,j}x^j}\chi$. □

Nótese que esta proposición implica que la estructura Hankel de la matriz de momentos no va a perderse con la evolución temporal y lo mismo va a ocurrir con el carácter definido positivo de la misma. Por ello, la factorización LU con dependencia temporal va a seguir teniendo sentido. Se introducen ahora un par de matrices auxiliares.

[8] Recuérdese que estamos en un caso definido positivo por lo que si $H_{i,j} = \delta_{i,j}h_j$ se tiene $h_k > 0$; de este modo $\left(\sqrt{H}\right)_{i,j} = \delta_{i,j}\sqrt{h_j}$.

Definición 32. *Sean las siguientes las matrices de onda:*

$$W_1(t) = Z(t)W_0(t_1), \qquad\qquad W_2(t) = Z(t)W_0(t_2).$$

Dado que $\frac{\partial W_0(t_\alpha)}{\partial t_{\alpha,j}} = \Lambda^j W_0(t_\alpha)$ es sencillo comprobar que las siguientes expresiones se cumplen:

$$\frac{\partial W_1(t)}{\partial t_{1,j}} = \Big(\frac{\partial Z}{\partial t_{1,j}} Z^{-1} + L^j(t)\Big) W_1, \qquad \frac{\partial W_1(t)}{\partial t_{2,j}} = \Big(\frac{\partial Z}{\partial t_{2,j}} Z^{-1}\Big) W_1,$$
$$\frac{\partial W_2(t)}{\partial t_{1,j}} = \Big(\frac{\partial Z}{\partial t_{1,j}} Z^{-1}\Big) W_2, \qquad \frac{\partial W_2(t)}{\partial t_{2,j}} = \Big(\frac{\partial Z}{\partial t_{2,j}} Z^{-1} + L^j(t)\Big) W_2.$$

Proposición 33. *Las potencias de la matriz de Jacobi satisfacen la siguiente ecuación diferencial:*

$$L^j(t) = -\Big(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big) - \Big(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big)^\top.$$

Denotando por $M_\pm$ a la parte estríctamente superior/inferior de la matriz M la ecuación anterior es equivalente a:

$$\big(L^j(t)\big)_- = -\Big(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big)_-, \quad \big(L^j(t)\big)_+ = -\Big(\Big[\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big]^\top\Big)_+, \quad \operatorname{diag}\big[L^j(t)\big] = -2\operatorname{diag}\Big[\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big].$$

Demostración. Desde la definición de $G(t)$ se tiene,

$$G(t) = W_0(t_1) G \big(W_0(t_2)\big)^\top \qquad \Longrightarrow \qquad G = W_1^{-1}(t) W_2^{-\top}(t).$$

Derívese ahora respecto de $t_{\alpha,j}$ y úsense las relaciones para las derivadas de W_α. □

Nótese cómo dicha proposición en primer lugar encaja con la observación de que la dependencia temporal de la medida $\mathrm{d}\mu(x, t_1, t_2)$ puede realmente entenderse como $\mathrm{d}\mu\big(x, (t_1 + t_2)\big)$, por lo que uno puede derivar indistintamente respecto de $t_{1,j}$ o respecto de $t_{2,j}$, y en segundo lugar, recuerda la propiedad de simetría $L = L^\top$. Las consecuencias más relevantes de la proposición se enuncian ahora en forma de teorema.

Teorema 1. *Las matrices de onda satisfacen las siguientes relaciones diferenciales:*

$$\frac{\partial W_1(t)}{\partial t_{1,j}} = -\Big(\frac{\partial Z}{\partial t_{1,j}} Z^{-1}\Big)^\top W_1, \qquad \frac{\partial W_1(t)}{\partial t_{2,j}} = \Big(\frac{\partial Z}{\partial t_{2,j}} Z^{-1}\Big) W_1,$$
$$\frac{\partial W_2(t)}{\partial t_{1,j}} = \Big(\frac{\partial Z}{\partial t_{1,j}} Z^{-1}\Big) W_2, \qquad \frac{\partial W_2(t)}{\partial t_{2,j}} = -\Big(\frac{\partial Z}{\partial t_{2,j}} Z^{-1}\Big)^\top W_2.$$

La matriz de Jacobi evoluciona de acuerdo al siguiente par de Lax:

$$\frac{\partial L^n(t)}{\partial t_{\alpha,j}} = \Big[\Big(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big), L^n\Big] = \Big[L^n, \Big(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\Big)^\top\Big].$$

Las siguientes ecuaciones de Zakharov-Shabat se satisfacen:

$$\frac{\partial}{\partial t_{\alpha,i}}\left(\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}\right) - \frac{\partial}{\partial t_{\alpha,j}}\left(\frac{\partial Z}{\partial t_{\alpha,i}} Z^{-1}\right) + \left[\frac{\partial Z}{\partial t_{\alpha,j}} Z^{-1}, \frac{\partial Z}{\partial t_{\alpha,i}} Z^{-1}\right] = 0.$$

Demostración. Las relaciones para las matrices de onda se siguen desde la proposición 33. La ecuación que involucra el par de Lax se obtiene tras derivar $L^n(t) = Z(t)\Lambda^n Z(t)^{-1}$ y usar de nuevo la proposición 33. En último lugar, las ecuaciones de Zakharov-Shabat son la condición de compatibildad de las ecuaciones de Lax. □

1.6.2 Conexión con la red de Toda

Considérese el caso en que $t_{\alpha,j} = 0\ \forall j \neq 1$ y hágase el siguiente cambio de variables para los dos únicos parámetros no nulos que quedan: $t_{1,1} = t+s$ y $t_{2,1} = t-s$. Tras este cambio, dada $f(t_{1,1}, t_{2,1})$ se tendrá que:

$$\mathrm{d}f = \frac{\partial f}{\partial t_{1,1}}\mathrm{d}t_{1,1} + \frac{\partial f}{\partial t_{2,1}}\mathrm{d}t_{2,1} = \left(\frac{\partial f}{\partial t_{1,1}} + \frac{\partial f}{\partial t_{2,1}}\right)\mathrm{d}t + \left(\frac{\partial f}{\partial t_{1,1}} - \frac{\partial f}{\partial t_{2,1}}\right)\mathrm{d}s.$$

Tómese la matriz de Jacobi $L(t_{1,1}, t_{2,1})$. Debido a su expresión en términos de las matrices de factorización, que a su vez provienen de la medida $\mathrm{d}\mu\big(x,(t_{1,1},t_{1,2})\big) = \mathrm{d}\mu(x)e^{(t_1+t_2)x}$, se puede observar que los elementos de la teoría van a depender de t y no de s (algo que ya se anticipaba tras los resultados de la proposición 33).

$$\frac{\partial L}{\partial s} = \frac{\partial L}{\partial t_{1,1}} - \frac{\partial L}{\partial t_{2,1}} = 0, \qquad\qquad \frac{\partial L}{\partial t} = \frac{\partial L}{\partial t_{1,1}} + \frac{\partial L}{\partial t_{2,1}} = \frac{\mathrm{d}L}{\mathrm{d}t}.$$

Por lo tanto, volviendo al teorema 1 y tomando ambos $n=1$ y $j=1$ reescribamos,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \frac{\partial L(t)}{\partial t_{1,1}} + \frac{\partial L(t)}{\partial t_{2,1}} = \left[\Big(\frac{\partial Z}{\partial t_{1,1}}Z^{-1}\Big) - \Big(\frac{\partial Z}{\partial t_{2,1}}Z^{-1}\Big)^{\top}, L\right] = \left[\Big(\frac{\partial Z}{\partial t_{1,1}}Z^{-1}\Big)_{-} - \Big(\Big[\frac{\partial Z}{\partial t_{2,1}}Z^{-1}\Big]^{\top}\Big)_{+}, L\right].$$

Pero la proposición 33 finalmente implica que,

$$\frac{\mathrm{d}L}{\mathrm{d}t} = \Big[(L_+ - L_-), L\Big].$$

Que en componentes queda,

$$\frac{\mathrm{d}(S_{n,n-1} - S_{n+1,n})}{\mathrm{d}t} = 2\left(h_{n+1}h_n^{-1} - h_n h_{n-1}^{-1}\right), \quad \frac{\mathrm{d}\sqrt{h_{n+1}h_n^{-1}}}{\mathrm{d}t} = \sqrt{h_{n+1}h_n^{-1}}\left(2S_{n+1,n} - S_{n+2,n+1} - S_{n,n-1}\right).$$

Este es precisamente un par de Lax equivalente a la ecuación de Toda. Se resume este resultado en forma de proposición:

Proposición 34. *El par de Lax asociado a la red de Toda*

$$\frac{\mathrm{d}^2 x_n}{\mathrm{d}t^2} = e^{(x_{n+1}-x_n)} - e^{(x_n - x_{n-1})},$$

Coincide con la matriz de Jacobi L y la matriz $(L_+ - L_-)$ que se obtienen tras deformar la matriz de momentos del siguiente modo:

$$G(t) = W_0(t_{1,1})GW_0(t_{2,1})^{\top} \qquad \Longrightarrow \qquad \frac{\mathrm{d}L}{\mathrm{d}t} = \Big[(L_+ - L_-), L\Big].$$

La conexión entre las variables x_n y los elementos de la teoría de los polinomios ortogonales h_n, $S_{n+1,n}$ vienen dados por el cambio de variables de Flaschka [62],

$$\sqrt{h_{n+1}h_n^{-1}} = \frac{1}{2}e^{\frac{x_{n+1}-x_n}{2}}, \qquad\qquad (S_{n,n-1} - S_{n+1,n}) = \frac{1}{2}\frac{\mathrm{d}x_n}{\mathrm{d}t}.$$

1.6.3 Conexión entre deformaciones continuas y discretas

A continuación se explica cómo mediante pequeños cambios, llamados desplazamientos de Miwa, en los parámetros de las deformaciones continuas es posible recuperar cualquier transformación espectral lineal.

Definición 33. *Se definen los desplazamientos de Miwa como sigue:*

$$t_{\alpha,j} \to \tilde{t}_{\alpha,j} := t_{\alpha,j} + \frac{k}{ja^j}, \qquad j > 0.$$

Proposición 35. *Un desplazamiento de Miwa de uno de los conjuntos de los parámetros que dan cuenta de la transformación continua, equivale a la siguiente transformación espectral lineal de la medida:*

$$\mathrm{d}\mu(x, \tilde{t}_1, t_2) = \mathrm{d}\mu(x, t_1, \tilde{t}_2) = \frac{(-1)^k}{a^k}(x-a)^k \mathrm{d}\mu(x, t_1, t_2).$$

Demostración. La prueba se basa en la expansión $\ln\left(1 - \frac{x}{a}\right) = \sum_{n=1}^{\infty} \frac{1}{na^n} x^n$.

$$e^{\sum_{j=0} \tilde{t}_j x^j} = \left(e^{\sum_{j=0} t_j x^j}\right)\left(e^{\sum_{j=1}^{\infty} \frac{k}{ja^j} x^j}\right) = \left(e^{\sum_{j=0} t_j x^j}\right)\left(1 - \frac{x}{a}\right)^k.$$

□

Nótese cómo, con del teorema fundamental del álgebra a nuestra disposición, tras un número finito de desplazamientos de Miwa, uno podría obtener cualquier cociente de polinomios como factor de la medida. Dicho esto, cabe destacar que en el momento en que el teorema fundamental del álgebra no sea aplicable (por ejemplo en el caso matricial o el multivariable), habrá polinomios a los que uno nunca podrá llegar y por lo tanto, la conexión no será tan completa como lo es en este caso.

2 Polinomios biortogonales generalizados en la recta real

Con este capítulo se busca dar una serie de pinceladas introductorias y que pretenden motivar la generalización, en diferentes direcciones, del concepto de ortogonalidad presentado hasta este punto del texto. La idea es, para los casos matricial, multivariable, múltiple y Sobolev, dedicar unas líneas a cada uno de los cinco pasos enumerados en el apartado *Metodología* del resumen introductorio de la tesis. Este capítulo ha de considerarse como nexo de unión entre lo escrito hasta este punto y aquellos artículos que integran esta tesis.
El concepto principal que precisa de generalización es la factorización LU. Una vez entendida la factorización LU usual de una matriz semi infinita A dada, no es complicado aceptar que, dado un conjunto infinito de enteros positivos $\rho = \{n_j\}_j^\infty$ la imposición $\det A^{[\sum_{j=0}^{l} n_j]} \neq 0$, $\forall l$ va a permitir que A factorice del siguiente modo que se ha de entender como una factorización LU por bloques.

$$A = L_\rho D_\rho U_\rho := \begin{pmatrix} \mathbb{I}_{n_0\times n_0} & 0 & \dots \\ L_{1,0} & \mathbb{I}_{n_1\times n_1} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} D_0 & 0 & \dots \\ 0 & D_1 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \mathbb{I}_{n_0\times n_0} & U_{0,1} & \dots \\ 0 & \mathbb{I}_{n_1\times n_1} & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}. \tag{2.1}$$

Donde $(L_\rho)_{i,j}, (U_\rho)_{i,j} \in \mathbb{R}^{n_i\times n_j}$ y $D_j \in \mathbb{R}^{n_j\times n_j}$. Nos interesa por supuesto el caso en que $A = G$ es una matriz de Gram dada. El conjunto $\rho = \{n_j\}_j^\infty$ quedará determinado por el tipo de ortogonalidad bajo consideración. Como estamos a punto de ver, $n_j = p$, $\forall j$ dará una factorización por bloques de igual tamaño adaptada al caso matricial, mientras que si $n_j = \binom{D+j-1}{j}$, $\forall j$ estaremos frente a una factorización por bloques de tamaño creciente inherente al caso multivariable.

2.1 Polinomios matriciales biortogonales en la recta real

M. G. Krein es el autor a quien se vinculan los orígenes de este tipo de polinomios [91]; denotados abreviadamente por su acrónimo MBPRL. Un texto clásico al respecto es [29], mientras que otros más recientes son [46],[54] y también [16], [70]. Resultados que involucran complementos de Schur como los que proponemos aquí se pueden consultar en [102]. Los MBPRL también son de utilidad a la hora de estudiar productos internos no estándar [55]. A pesar de que son muchas las propiedades que comparten con sus homólogos escalares, no están exentos de sorpresas por ejemplo ligadas a la caracterización tipo Bochner (mediante un operador diferencial de segundo orden y con coeficientes polinómicos) de los que se podrían considerar como los análogos matriciales de los polinomios clásicos [56]. Al contrario de lo que sucede en el caso escalar, el matricial permite definir polinomios matriciales "clásicos" caracterizados mediante un operador diferencial lineal de primer orden [38]. Un enfoque desde la factorizción LU de la teoría de los MBPRL se puede encontrar en nuestras publicaciones [14] y [13].

- **Forma sesquilineal**. Dada una medida de Borel μ matricial de tamaño $p\times p$ y con soporte Ω en un conjunto (infinito) de puntos de la recta real, se define la siguiente forma sesquilineal entre dos polinomios matriciales $f(x) = \sum_{k=0} f_j x^j$ y $h(x) = \sum_{k=0} h_j x^j$ donde f_j, $h_j \in \mathbb{R}^{p\times p}$:

$$\langle f(x), h(x)\rangle_\mu := \int_\Omega f(x)\mathrm{d}\mu(x)\big(h(x)\big)^\top.$$

- **Matriz de momentos** G. Puesto que queremos tratar con polinomios matriciales, el vector de monomios ha de capturar esta cualidad y transmitirlo de algún modo a la organización de las entradas de la matriz de momentos,

$$\chi(x) := \begin{pmatrix}\mathbb{I}_{p\times p} & x\mathbb{I}_{p\times p} & x^2\mathbb{I}_{p\times p} & \dots, x^2\mathbb{I}_{p\times p} & \dots\end{pmatrix}^\top, \qquad G := \langle \chi, \chi\rangle_\mu := \int_\Omega \chi(x)\mathrm{d}\mu(x)\big(\chi(x)\big)^\top,$$

$$f(x) = \sum_{j=0}^{k} f_j\chi_j(x), \quad f_j \in \mathbb{R}^{p\times p}, \qquad G_{i,j} = \int_\Omega \chi_i(x)\mathrm{d}\mu(x)\big(\chi_j(x)\big)^\top \in \mathbb{R}^{p\times p}.$$

Nótese cómo, para una misma matriz de momentos, hay un conjunto de formas sesquilineales alternativas y relacionadas entre sí, como por ejemplo $(f,h)_\mu := \int_\Omega \big(f(x)\big)^\top \mathrm{d}\mu(x)h(x)$ y variaciones de esta. No es complicado ver que cualquiera de las posibilidades va a dar la misma matriz de momentos (cosa que no ocurrirá en el caso de biortogonalidad matricial con soporte en la circunferencia unidad, pero nos encargaremos de esa situación más adelante).

- **Factorización LU** de G. La observación de que el vector de monomios $\chi(x)$ tenga sus entradas proporcionales a $\mathbb{I}_{p\times p}$ sugiere que la factorización LU de G ha de respetar este hecho y ser por bloques de tamaño $p\times p$ (de acuerdo con (2.1), $n_j = p$).

$$G = S_1^{-1}HS_2^{-\top}, \qquad H_{i,j} = \delta_{i,j}h_j \in \mathbb{R}^{p\times p}.$$

Esta factorización dará lugar a dos familias mónicas ($(S_\alpha)_{k,k} = \mathbb{I}$) de MBPRL,

$$P_\alpha(x) := S_\alpha\chi(x) = \begin{pmatrix} P_{\alpha,0}x) \\ P_{\alpha,1}(x) \\ \vdots \\ P_{\alpha,k}(x) \\ \vdots \end{pmatrix}, \qquad P_{\alpha,k}(x) = \sum_{j=0}^{k} \big(S_\alpha\big)_{k,j}\chi_j(x), \qquad \langle P_{1,i}, P_{2,j}\rangle_\mu = \delta_{i,j}h_j.$$

Cuyas leyes de ortogonalidad son:

$$\langle P_{1,k}, x^j\rangle_\mu = 0, \qquad \langle x^j, P_{2,k}\rangle_\mu = 0, \qquad j = 0,1,2,\dots,k-1.$$

De darse que $\mathrm{d}\mu(x) = \mathrm{d}\mu(x)^\top$ supondría que G fuera una matriz simétrica y por lo tanto no habría necesidad del subíndice α. En este caso, la biortogonalidad se vería reducida a ortogonalidad.

- **Simetrías de** G. Dada la construcción de la matriz de momentos es casi esperable algún tipo de simetría tipo Hankel, pero esta vez, por bloques. La manera de evidenciar esta organización de las entradas es definir una matriz de translación (al estilo de 11) adaptada al caso matricial:

$$\Lambda := \begin{pmatrix} 0 & \mathbb{I}_{p\times p} & 0 & 0 & \dots \\ 0 & 0 & \mathbb{I}_{p\times p} & 0 & \dots \\ 0 & 0 & 0 & \mathbb{I}_{p\times p} & \dots \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix}, \qquad \Lambda\chi(x) = x\chi(x), \qquad \Lambda G = G\Lambda^\top.$$

Esto realmente supone, al igual que en el caso escalar, que $\langle xf(x), g(x)\rangle = \langle f(x), xg(x)\rangle$, por lo que dicha simetría va a dar cuenta de la relación de recurrencia a tres términos-bloques así como las fórmulas de Chrisfoffel–Darboux.

- **Deformaciones de** G. Las transformaciones de Darboux o discretas en este caso tienen una serie de matices propios del carácter matricial de que el caso escalar carece. A parte de tratarse de transformaciones no abelianas (uno podría multiplicar la matriz de medidas desde su derecha o bien desde su izquierda siendo estas, dos transformaciones diferentes), uno puede encontrarse con transformaciones polinómicas en las que el coeficiente director del polinomio resulta ser una matriz singular, careciendo en tal caso, de parte de su información espectral (que es de gran utilidad para el estudio de estas deformaciones). Las transformaciones continuas vendrán parametrizadas por dos ($\alpha = 1, 2$) conjuntos infinitos $j = 1, 2 \ldots$ de matrices diagonales $t_{\alpha,j}$ que van a permitir conectar la teoría de los MBPRL con la jerarquía no abeliana de la red de Toda.

2.2 Polinomios ortogonales en varias variables reales

Rebuscando en los orígenes de estos polinomios uno puede encontrar algunas referencias a los mismos en los trabajos de C. Hermite, aunque primeras publicaciones al respecto las encontramos en [15] y [84]. Siguientes artículos destacados son [58] y [90] en los que la notación vectorial para los polinomios (que también usaremos nosotros) se presenta por primera vez y donde se consideran polinomios análogos a los polinomios clásicos, esta vez en dos variables (también en [88]). Como versiones excelentes más actuales mencionar [117] y [53]. Esta tesis incluye un enfoque desde la factorización LU de este tipo de ortogonalidad [21] y su correspondiente conexión con los sistemas integrables. Dos publicaciones que complementan esta última son: por un lado [22], en la que generalizamos el concepto de la transformación de Christoffel presentado en [21], y por por otro [23], en la que estudiamos la teoría de las deformaciones en el contexto multivariable en su versión más general (desde el punto de vista de las transformaciones de un funcional).

- **Forma sesquilineal**. Mediante la notación $\boldsymbol{x} = (x_1, x_2, \ldots, x_D) \in \mathbb{R}^D$ y tomando una medida de Borel dependiente de D variables $\mathrm{d}\mu(\boldsymbol{x})$ con soporte $\Omega \in \mathbb{R}^D$, definimos la siguiente forma bilineal:

$$\langle f(\boldsymbol{x}), h(\boldsymbol{x})\rangle_\mu := \int_\Omega f(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})h(\boldsymbol{x}).$$

- **Matriz de momentos** G. Un paso previo a la construcción de la matriz de momentos es la elección de un vector de monomios adaptado $\chi(\boldsymbol{x})$. Empleamos la notación habitual en la que multi-índices $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_D)$ con $\alpha_j \in \mathbb{N}$ sirven para denotar cualquier monomio $\boldsymbol{x}^{\boldsymbol{\alpha}} := (x_1^{\alpha_1}, x_2^{\alpha_2}, \ldots, x_D^{\alpha_D})$. De esta forma, un polinomio arbitrario va a poderse escribir de acuerdo a la expresión $f(\boldsymbol{x}) = \sum_{\boldsymbol{\alpha}_j} f_{\boldsymbol{\alpha}_j}\boldsymbol{x}^{\boldsymbol{\alpha}_j}$. El motivo por el que queremos construir un vector de monomios adaptado es el de ordenar todos ellos de algún modo razonable. En primer lugar definimos la longitud de un multi-índice: $|\boldsymbol{\alpha}| := \sum_{a=1}^D \alpha_a$, lo que permite definir aquel conjunto de multi-índices de igual longitud, cuyo cardinal es sencillo de obtener

$$[k] := \{\boldsymbol{\alpha} \text{ tales que } |\boldsymbol{\alpha}| = k\}, \qquad |[k]| = \binom{D+k-1}{k}.$$

El concepto de longitud del multi-índice supone un primer paso a la hora de ordenar los monomios como deseábamos, si $|\boldsymbol{\alpha}_1| < |\boldsymbol{\alpha}_2|$ entonces diremos que $\boldsymbol{x}^{\boldsymbol{\alpha}_1} < \boldsymbol{x}^{\boldsymbol{\alpha}_2}$. El segundo paso pretende ordenar los multi-índices dentro de cada conjunto $[k]$. Para ello, seguimos el orden dado por la graduación lexicográfica inversa: Denotando por $\boldsymbol{\alpha}_1^{(k)}, \boldsymbol{\alpha}_2^{(k)}$ a dos multi-índices $\in [k]$ en caso de afirmar que $\boldsymbol{\alpha}_1^{(k)} < \boldsymbol{\alpha}_2^{(k)}$ se ha de cumplir que $\alpha_{1,1}^{(k)} = \alpha_{2,1}^{(k)}$, $\alpha_{1,2}^{(k)} = \alpha_{2,2}^{(k)}$, ..., $\alpha_{1,l}^{(k)} > \alpha_{2,l}^{(k)}$ para $l \leq D$. De este modo, por fin tenemos una organización de los monomios que nos permite definir el vector que agrupa los mismos y la matriz de momentos que a partir de ellos podemos construir.

$$\chi(\boldsymbol{x}) := \begin{pmatrix} \chi_{[0]} \\ \chi_{[1]} \\ \vdots \\ \chi_{[k]} \\ \vdots \end{pmatrix}, \qquad \chi_{[k]} := \begin{pmatrix} \boldsymbol{x}^{\boldsymbol{\alpha}_1^{(k)}} \\ \boldsymbol{x}^{\boldsymbol{\alpha}_2^{(k)}} \\ \vdots \\ \boldsymbol{x}^{\boldsymbol{\alpha}_{|[k]|}^{(k)}} \end{pmatrix}, \qquad G := \int_\Omega \chi(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})\Big(\chi(\boldsymbol{x})\Big)^\top$$

$$G_{\boldsymbol{\alpha}_j^{(n)},\boldsymbol{\alpha}_i^{(m)}} = \int_\Omega \boldsymbol{x}^{\boldsymbol{\alpha}_j^{(n)}+\boldsymbol{\alpha}_i^{(m)}} \mathrm{d}\mu(\boldsymbol{x}), \qquad G_{[i],[j]} := \int_\Omega \chi_{[i]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})\Big(\chi_{[j]}(\boldsymbol{x})\Big)^\top \in \mathbb{R}^{|[i]|\times|[j]|}.$$

Desde luego $G = G^\top$ es una matriz simétrica.

- **Factorización LU** de G. Tanto la longitud creciente de cada una de las entradas $\chi_{[j]}$ como el tamaño creciente de cada bloque $G_{[i],[j]}$ sugieren una factorización LU (Cholesky en este caso) que respete tal hecho. En este caso, desde (2.1) se tiene $n_j = |[j]|$.

$$G = S^{-1}HS^{-\top}, \qquad H_{[i][j]} = \delta_{i,j}h_j \in \mathbb{R}^{|[i]|\times|[j]|},$$

$$P(\boldsymbol{x}) := S\chi(\boldsymbol{x}) = \begin{pmatrix} P_{[0]} \\ P_{[1]} \\ P_{[2]} \\ \vdots \end{pmatrix}, \qquad P_{[k]}(\boldsymbol{x}) = \sum_{j=0}^{k} \big(S\big)_{[k][j]}\chi_{[j]}(\boldsymbol{x}) = \begin{pmatrix} P_{\boldsymbol{\alpha}_1^{(k)}} \\ P_{\boldsymbol{\alpha}_2^{(k)}} \\ \vdots \\ P_{\boldsymbol{\alpha}_{|[k]|}^{(k)}} \end{pmatrix}.$$

Quienes satisfacen $\langle P_{[i]}, P_{[j]}\rangle_\mu = \delta_{i,j}h_j$ o equivalentemente,

$$\langle P_{\boldsymbol{\alpha}_j^{(k)}}(\boldsymbol{x}), \boldsymbol{x}^{\boldsymbol{\alpha}}\rangle_\mu = 0, \quad \forall|\boldsymbol{\alpha}| < k; \qquad \langle P_{\boldsymbol{\alpha}_j^{(k)}}(\boldsymbol{x}), \boldsymbol{x}^{\boldsymbol{\alpha}_i^{(k)}}\rangle_\mu = \big(h_k\big)_{\alpha_j^{(k)},\alpha_i^{(k)}}.$$

- **Simetrías de** G. Del estudio de las ortogonalidades anteriores sabemos que la estructura Hankel de la matriz de momentos viene expresada como una ecuación entre la matriz de momentos y la de translación. Dicha ecuación también da cuenta del carácter autoadjunto del operador multiplicación por la variable x. Puesto que en el caso multivariable es D el número de variables a tener en cuenta, tiene sentido presentar ya no una, sino un conjunto de D matrices de translación (que además conmutan entre ellas):

$$[\Lambda_a, \Lambda_b] = 0, \qquad \Lambda_a\chi(\boldsymbol{x}) = x_a, \qquad \Lambda_a G = G\Lambda_a^\top, \qquad a, b = 1, 2, \dots, D.$$

La simetría que esto supone para la matriz de momentos tendrá que ver con la recurrencia y las fórmulas de Christoffel–Darboux.

- **Deformaciones de** G. Nuestra herramienta para estudiar las deformaciones discretas precisa de los datos espectrales de los polinomios perturbadores. En el caso unidimensional son suficientes una serie de puntos (los ceros) y sus multiplicidades, en el caso matricial hay que añadir a la lista anterior las multiplicidades parciales y las cadenas de Jordan, y en el caso multivariable van a ser necesarias una serie de nociones básicas de geometría algebraica. Las deformaciones continuas de este tipo de medidas permitirán construir ecuaciones no lineales, tanto en derivadas parciales como en diferencias, con coeficientes matriciales de tamaño variable cuyas soluciones vendrán dadas por elementos matriciales de la teoría de los polinomios ortogonales.

2.3 Polinomios múltiples biortogonales de tipo mixto

Los orígenes de este tipo de ortogonalidad los podemos situar al final del siglo XIX ligados a los esfuerzos de C. Hermite para probar la transcendencia del número de Euler [80]. Un tema al que la ortogonalidad múltiple está íntimamente ligada [115] es el de la aproximación de Hermite-Padé (o aproximación racional simultánea) también desarrollada entre el citado C. Hermite y su estudiante H. Padé. Buenos trabajos respecto de este tipo de ortogonalidad son [17],[34],[105]. Para un enfoque desde el problema de Riemann–Hilbert véanse [45], [125]. Por último, para una mirada desde la perspectiva de la factorización LU citemos [11] junto con [20], este último incluido en esta tesis.

- **Forma sesquilineal**. Partiendo de los siguientes dos conjuntos de pesos $\{w_{\alpha,j}\}_{j=1}^{p_\alpha}$ para $\alpha = 1, 2$ podemos definir una forma bilineal a partir de los mismos como sigue,

$$\langle f(x), h(x)\rangle_{\boldsymbol{w}} := \sum_{r=1}^{p_1}\sum_{l=1}^{p_2}\langle f(x), h(x)\rangle_{\mu_{r,l}}, \quad \langle f(x), h(x)\rangle_{\mu_{r,l}} := \int_\Omega f(x)\Big(\mathrm{d}\mu(x)w_{1,r}(x)w_{2,l}(x)\Big)h(x).$$

- **Matriz de Gram** G. En el caso múltiple, la tarea de definir una matriz de Gram es algo más delicada de lo que ha sido hasta el momento. La primera idea que a uno le viene a la mente es tomar por matriz de Gram a la siguiente suma $\tilde{G} = \sum_{r,l} G_{r,l}$, donde $G_{r,l} = \langle \chi, \chi\rangle_{\mu_{r,l}}$ y χ es el vector de monomios (escalar y en una única variable). Sin embargo, esta no es la elección más general a la que uno puede aspirar (es en realidad un caso límite como veremos brevemente). Para conseguir esta generalización es necesario introducir el siguiente par de vectores $\boldsymbol{n}_\alpha = \big(n_{\alpha,1}, n_{\alpha,2}, \ldots, n_{\alpha,p_\alpha}\big)$ con $\alpha = 1, 2$ y en el que $n_{\alpha,j} \in \mathbb{Z}_+$, mediante los que construimos los dos vectores de monomios adaptados al caso múltiple. En primer lugar es necesario definir,

$$(\chi_\alpha)_{[0]} := \Big(\big(1 \quad z \quad \ldots z^{(n_{\alpha,1}-1)}\big)\,\big(1 \quad z \quad \ldots z^{(n_{\alpha,2}-1)}\big)\ldots\big(1 \quad z \quad \ldots z^{(n_{\alpha,p_\alpha}-1)}\big)\Big)^\top.$$

En segundo lugar, desde el vector anterior, tenemos cualquier entrada expresada como sigue,

$$\chi_\alpha(x) := \begin{pmatrix} (\chi_\alpha)_{[0]} \\ (\chi_\alpha)_{[1]} \\ \vdots \\ (\chi_\alpha)_{[k]} \\ \vdots \end{pmatrix}, \qquad (\chi_\alpha)_{[k]} := \begin{pmatrix} z^{kn_{\alpha,1}}\mathbb{I}_{n_{\alpha,1}} & 0 & 0 & 0 \\ 0 & z^{kn_{\alpha,2}}\mathbb{I}_{n_{\alpha,2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & z^{kn_{\alpha,p_\alpha}}\mathbb{I}_{n_{\alpha,p_\alpha}} \end{pmatrix}(\chi_\alpha)_{[0]}.$$

En tercer y último lugar,

$$\xi_\alpha(x) := \begin{pmatrix} (\xi_\alpha)_{[0]} \\ (\xi_\alpha)_{[1]} \\ \vdots \\ (\xi_\alpha)_{[k]} \\ \vdots \end{pmatrix}, \quad (\xi_\alpha)_{[k]} := \begin{pmatrix} w_{\alpha,1}z^{kn_{\alpha,1}}\mathbb{I}_{n_{\alpha,1}} & 0 & 0 & 0 \\ 0 & w_{\alpha,1}z^{kn_{\alpha,2}}\mathbb{I}_{n_{\alpha,2}} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & w_{\alpha,1}z^{kn_{\alpha,p_\alpha}}\mathbb{I}_{n_{\alpha,p_\alpha}} \end{pmatrix}(\chi_\alpha)_{[0]}.$$

Una vez establecida esta notación, dada una medida, el par de vectores de pesos y el par de vectores para organizar los monomios $(\mathrm{d}\mu, w_\alpha, \boldsymbol{n}_\alpha)$ definimos la matriz de Gram como sigue:

$$G := \int_\omega \xi_1(x)\mathrm{d}\mu(x)\Big(\xi_2(x)\Big)^\top.$$

Nótese la gran cantidad de casos que este escenario múltiple engloba; el caso $p_1 = p_2 = 1$ nos lleva directamente a los OPRL respecto de $\mathrm{d}\mu = w_1w_2$; el caso $p_1 = p_2 = p$ y $n_{1,j} = n_{2,j} = 1, \forall j = 1, \ldots p$ no es otro que el de los MOPRL de tamaño $p \times p$ respecto de $\mathrm{d}\mu = \boldsymbol{w}_1(\boldsymbol{w}_2)^\top$, donde hemos llamado $\boldsymbol{w}_\alpha = (w_{\alpha,1}, w_{\alpha,2}, \ldots, w_{\alpha,p})$. Por último, el caso límite que anteriormente mencionábamos toma lugar en caso de permitir $n_{\alpha,j} \to \infty$.

- **Factorización LU** de G. Como ocurría al definir la matriz de Gram más apropiada, el tipo de factorización LU a elegir que se adapte mejor al caso bajo estudio es algo delicado. Mientras que una factorización por bloques cuyos tamaños tuvieran en cuenta los diferentes $n_{\alpha,j}$ sería lo deseable, nosotros empleamos una factorización usual, que siendo algo más restrictiva que la anterior, simplifica bastante los razonamientos y también nos permite obtener los resultados que buscamos.

- **Simetrías de G.** La discusión anterior deja entrever que la matriz de Gram en el caso múltiple no es sino una especie de *collage* de recortes de las $p_1 \cdot p_2$ diferentes matrices de momentos parciales cuyo tamaño y posición vendrá dado por los $n_{\alpha,j}$ y que se irá repitiendo periódicamente. Algo similar va a ocurrir con las dos posibles matrices de translación adaptadas al caso; serán *collages* similares al de G pero en lugar de tomar recortes de las matrices de momentos, tomarán recortes de la matriz de translación del caso estándar (escalar y única variable). Cada una de estas dos matrices de translación tendrá como autovector el vector de monomios que le corresponda. Estas matrices van a modelar la simetría de la matriz de Gram, que tendrá por consecuencia una matriz de Jacobi por bloques organizados recordando la forma de una serpiente en la que queda codificada la ley de recurrencia de los polinomios múltiples.

- **Deformaciones de G.** Puesto que la medida y los pesos involucrados esta vez son todos escalares, no resulta sorprendente que el estudio de las deformaciones del caso múltiple sea bastante más parecido al caso escalar de lo que lo fueron el caso matricial y el multivariado. Las deformaciones continuas conectan la ortogonalidad múltiple con la jerarquía de Toda multicomponente, mientras que los flujos discretos pueden correctamente considerarse elementales y están conectados a los anteriores a base de desplazamientos de Miwa al estilo explicado en la sección 1.6.3.

2.4 Polinomios biortogonales de Sobolev en la recta real

D. C. Lewis en [93] fue quien plantó la primera semilla al tratar de aproximar una función a base de polinomios e involucrando las derivadas de estos. Motivado por este trabajo, P. Althammer publicó [10], el primer artículo en considerar un producto de Sobolev propiamente dicho. En [110] y [111] encontramos tempranas pero importantes contribuciones a las que siguió un periodo de inactividad en el tema cuyo final llegó gracias a un nuevo concepto, los "pares coherentes" de medidas [81]. Una clasificación de estos la podemos encontrar en [94]. A las publicaciones anteriores merece la pena incluir [101] y [96] como recomendables revisiones de la historia, resultados y bibliografía referentes al tema. Selecciono [27] (incluido en esta tesis) de entre nuestras contribuciones para aquel lector interesado en un estudio de este tipo de biortogonalidad desde la perspectiva de la factorizción LU junto con una serie de generalizaciones que también atañen al concepto de los "pares coherentes".

- **Forma sesquilineal.** El primer concepto a definir antes de introducir la forma bilineal de Sobolev es la matriz de medidas de orden $\mathcal{N}$. Esta matriz de tamaño $(\mathcal{N}+1)\times(\mathcal{N}+1)$ que denotaremos por $\mathscr{W}$, tiene por entradas $\mathscr{W}_{i,j} = \mathrm{d}\mu_{i,j}(x)$ medidas de Borel ($\mathrm{d}\mu_{i,j} = 0\ \forall i,j > \mathcal{N}$) con soporte $\Omega_{i,j}$:

$$\mathscr{W}(x) := \begin{pmatrix} \mathrm{d}\mu_{0,0} & \mathrm{d}\mu_{0,1} & \dots & \mathrm{d}\mu_{0,\mathcal{N}} & 0 & \dots \\ \mathrm{d}\mu_{1,0} & \mathrm{d}\mu_{1,1} & \dots & \mathrm{d}\mu_{1,\mathcal{N}} & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ \mathrm{d}\mu_{\mathcal{N},0} & \mathrm{d}\mu_{\mathcal{N},1} & \dots & \mathrm{d}\mu_{\mathcal{N},\mathcal{N}} & 0 & \dots \\ 0 & 0 & & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & & \ddots \end{pmatrix}, \qquad \mathrm{d}\mu_{i,j} : \Omega_{i,j} \subseteq \mathbb{R} \longrightarrow \mathbb{R}.$$

Nótese que alguna $\mathrm{d}\mu_{i,j}(x)$ con $i,j \leq \mathcal{N}$ puede ser cero siempre que exista al menos una medida no nula a lo largo de la columna o fila $\mathcal{N}$-*ésima*. El caso límite $\mathcal{N} \to \infty$ también puede estudiarse, pero con algo más de atención. Empleando la matriz de medidas escribimos la forma bilineal de Sobolev como sigue,

$$(f,h;\mathscr{W}) := \sum_{n,r=0}^{\mathcal{N}} \langle f^{(n)}, h^{(r)} \rangle_{\mu_{n,r}}, \qquad \langle f^{(n)}, h^{(r)} \rangle_{\mu_{n,r}} := \int_{\Omega_{n,r}} f^{(n)}(x)\, \mathrm{d}\mu_{n,r}(x)\, h^{(r)}(x).$$

- **Matriz de Gram** $G_{\mathscr{W}}$. La matriz de Gram de Sobolev asociada a la matriz de medidas $\mathscr{W}$ admite diferentes expresiones, pero quizás, la más sugerente, en estas líneas introductorias sea la siguiente:

$$G_{\mathscr{W}} = \sum_{l,r=0}^{\mathcal{N}} D^l G_{l,r} (D^r)^\top, \qquad G_{l,r} = \langle \chi, \chi^\top \rangle_{\mu_{l,r}}. \tag{2.2}$$

Donde hemos usado la matriz de derivación D de la definición 17.

- **Factorización LU** de $G_{\mathscr{W}}$. En el contexto Sobolev, dado que tratamos con polinomios escalares en una única variable, la elección más razonable del tipo de factorización es la estándar. (Será Cholesky en los casos en que $\mathscr{W} = \mathscr{W}^\top$).

$$G_{\mathscr{W}} := S_1^{-1} H \left(S_2^{-1}\right)^\top \quad \Longrightarrow \quad P_\alpha(x) := S_\alpha \chi(x) := \begin{pmatrix} P_{\alpha,0}(x) \\ P_{\alpha,1}(x) \\ \vdots \\ P_{\alpha,k}(x) \\ \vdots \end{pmatrix}, \quad (P_{1,r}, P_{2,k}; \mathscr{W}) := h_r \delta_{r,k}.$$

Con las siguientes condiciones de ortogonalidad:

$$(P_{1,l}, x^r; \mathscr{W}) := \delta_{l,r} h_r, \quad \forall r \leq l \quad \Longrightarrow \quad \sum_{k=0}^{l} \sum_{j=0}^{r} \left\langle P_{1,l}^{(k)}, \frac{\mathrm{d}^j x^r}{\mathrm{d}x^j} \right\rangle_{\mu_{k,j}} = \begin{cases} 0 & \forall r < l \\ h_l & r = l \end{cases}.$$

$$(x^r, P_{2,l}, ; \mathscr{W}) := h_r \delta_{r,l}, \quad \forall r \leq l \quad \Longrightarrow \quad \sum_{k=0}^{r} \sum_{j=0}^{l} \left\langle \frac{\mathrm{d}^j x^r}{\mathrm{d}x^j}, P_{2,l}^{(k)} \right\rangle_{\mu_{j,k}} = \begin{cases} 0 & \forall r < l \\ 1 & r = l \end{cases}.$$

- **Simetrías y deformaciones de** $G_{\mathscr{W}}$. En este caso Sobolev, lo más sensato es permitir a los pasos cuatro y cinco entrelazarse como uno único. El por qué de esta decisión lo encontramos en la falta de la hasta hora ubicua (en cualquiera de sus versiones) simetría de Hankel asociada a la matriz de translación apropiada (carácter autoadjunto del operador multiplicación por x), lo que además priva (en el caso general) a la secuencia biortogonal de polinomios de Sobolev de una recurrencia y correspondiente fórmula de Christoffel–Darboux. También es cierto que existen casos particulares, como por ejemplo aquellos ligados a formas bilineales de Sobolev discretas que sí presentan una simetría del tipo $Q_1(\Lambda)G_{\mathscr{W}} = G_{\mathscr{W}} Q_2(\Lambda^\top)$ donde los $Q_\alpha(\Lambda)$ son polinomios (relacionados con el conjunto discreto que define la forma bilineal) evaluados en la matriz de translación usual. El caso en que las entradas de $\mathscr{W}$ son medidas semiclásicas (o bien estas multiplicadas polinomios) es otro ejemplo en el que la matriz de Gram de Sobolev presenta cierta simetría codificada por un operador diferencial lineal $F(\Lambda, D)$ (relacionado con las ecuaciones de Pearson que las medidas semiclásicas satisfacen) del tipo $F(\Lambda, D) G_{\mathscr{W}} = G_{\mathscr{W}} \Big(F(\Lambda, D)\Big)^\top$.
Esta falta, en el caso general, de simetrías "obvias" se puede superar parcialmente desde nuestro enfoque LU. La idea consiste en considerar cada falta de simetría como una transformación de la matriz de medidas inicial $\mathscr{W}$. Nótese que en este contexto, las transformaciones van a tener un carácter bastante más general, puesto que ahora, no tendremos únicamente deformaciones polinómicas a nuestra disposición, sino que además tendremos aquellas involucrando operadores diferenciales con coeficientes polinómicos. Cerramos el conjunto de deformaciones de la forma bilineal de Sobolev introduciendo dos matrices con las que incorporar a $\mathscr{W}$ una dependencia temporal $\mathscr{W}(t, x) = \mathscr{W}_1(t_1, x) \mathscr{W}(x) \big(\mathscr{W}_2(t_2, x)\big)^{-1}$ permitiendo así conectar los coeficientes de los polinomios biortogonales de Sobolev con ecuaciones de tipo Toda.

Parte II

Biortogonalidad en la circunferencia unidad

1 Polinomios biortogonales en la circunferencia unidad

El capítulo anterior lo comencé revisando el estudio de los polinomios escalares ortogonales respecto de un funcional lineal dado, concepto al que le siguieron una serie de generalizaciones: coeficientes matriciales, dependencia de varias variables, ortogonalidad simultánea respecto de más de una medida y ortogonalidad definida en función de las derivadas de los polinomios. En el capítulo que comienza se analizará una generalización en otra dirección diferente; nótese cómo todos los ejemplos anteriores tienen en común que el soporte (o soportes) del funcional respecto del cual se define la biortogonalidad es siempre un conjunto infinito de puntos contenidos en la recta real (o varias rectas reales en el caso multivariable). Dicho esto, ¿por qué no generalizar esta situación por ejemplo a cualquier curva σ del plano complejo $\mathbb{C}$? Pues bien, de todas las posibles curvas, el caso más estudiado (véanse por ejemplo [73] y por supuesto los dos monumentales volúmenes [113], [114]) toma $\sigma = \mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ la circunferencia unidad, caso al que dedicaremos los siguientes párrafos. El primer paso consistirá pues en definir una forma sesquilineal adapatada a la circunferencia unidad.

Definición 34. *Una medida de Borel μ cuyo soporte sea un conjunto infinito de puntos sobre la circunferencia unidad, permite definir la siguiente forma sesquilineal* [1]*:*

$$\langle f, h \rangle := \oint_{\mathbb{T}} f(z) \frac{\mathrm{d}\mu(z)}{iz} \overline{h(z)}.$$

La literatura se centra especialmente en aquellas formas sesquilineales hermíticas y definidas positivas, puesto que en estos casos uno tiene detrás la estructura de espacio de Hilbert. Un enfoque desde la factorización LU, permite sin demasiado esfuerzo adicional la generalización de estos dos adjetivos que acompañan a la forma sesquilineal. Al relajar la condición de hermiticidad, pasaremos a tener polinomios biortogonales en lugar de ortogonales. Sabemos también, tras el capítulo anterior que, si en lugar de pedir que la forma sesquilineal sea definida positiva imponemos únicamente que la matriz de Gram G asociada a la misma satisfaga que $\det G^{[k]} \neq 0\ \forall k$ (lo que daría una forma cuadrática no degenerada entre elementos de $\mathbb{C}[z]$) también seguiremos podiendo construir toda la estructura de la que se siguen nuestros resultados. En resumidas cuentas, la única condición que realmente vamos a imponer sobre la forma sesquilineal es que la matriz de Gram asociada a la misma tenga todos sus menores principales no nulos.

1.1 Matriz de momentos de Szegő y polinomios de Szegő

El término polinomios de Szegő es equivalente al de aquellos OPUC asociados a un producto interno hermítico. Seguiré llamándolos de este modo aun en el caso no hermítico y aunque en lugar de ortogonalidad necesitemos hablar de biortogonalidad.

Definición 35. *La matriz de momentos de Szegő se define del siguiente modo:*

$$G_{sz} := \langle \chi, \chi^\top \rangle_\mu = \oint_{\mathbb{T}} \chi(z) \frac{\mathrm{d}\mu(z)}{iz} (\chi(z))^\dagger .$$

[1]Usaremos la notación $\frac{\mathrm{d}\mu(z)}{iz} \equiv \mathrm{d}\mu(\theta)$ motivada por el caso absolutametnte continuo $\mathrm{d}\mu(z) = \omega(z)\mathrm{d}z$ donde se puede obervar que $\frac{\mathrm{d}\mu(z)}{iz} = \omega(z)\frac{\mathrm{d}z}{iz} = \omega(e^{i\theta})\mathrm{d}\theta = \mathrm{d}\mu(\theta)$.

Obsérvese cómo, dado que $z \in \mathbb{T}$, se cumple $\bar{z} = \frac{1}{z}$ lo que supone una simetría Toeplitz [2] para G_{sz}. La siguiente condición sobre sus menores permite enunciar la siguiente proposición.

Proposición 36. *Siempre que* $\det G_{sz}^{[k]} \neq 0\ \forall k$ *la matriz de momentos de Szegő va a admitir una factorización LU,*

$$G_{sz} := Z_1^{-1} H Z_2^{-\dagger}.$$

En términos de las matrices de la factorización Z_r, $r = 1, 2$, se construyen los siguientes polinomios:

Definición 36. *Los polinomios biortogonales de Szegő se definen a continuación:*

$$P_r(z) := Z_r \chi(z) = \begin{pmatrix} P_{r,0}(z) & P_{r,1}(z) & \dots & P_{r,l}(z) & \dots \end{pmatrix}^\top, \qquad r = 1, 2.$$

Los polinomios reversos o recíprocos de los anteriores se denotan y definen como: $P_{r,l}^*(z) := z^l \overline{P_{r,l}(\frac{1}{\bar{z}})}$. *Coeficientes de reflexión, parámetros de Verblunsky o Schur es el nombre por el que se conoce a los valores* $P_{r,l}(0) := \alpha_{r,l}$.

Como se deduce de su definición, estos polinomios son mónicos. Su propiedad de biortogonalidad se entiende mejor desde la siguiente proposición.

Proposición 37. *Los polinomios biortogonales de Szegő y sus reversos cumplen las siguientes relaciones:*

$$\begin{array}{ccc} \langle P_{1,l}, P_{2,j} \rangle_\mu = h_l \delta_{l,j} & \textit{donde} & H := \operatorname{diag}\{h_k\}. \\ \langle P_{1,l}, z^k \rangle_\mu = 0, & \langle z^k, P_{2,l} \rangle_\mu = 0, & k = 0, 1, \dots l-1. \\ \langle P_{2,l}^*, z^k \rangle_\mu = 0. & \langle z^k, P_{1,l}^* \rangle_\mu = 0, & k = 1, 2, \dots l. \end{array}$$

Demostración. Son consecuencia directa de su definición en términos de las matrices de la factorización. $\square$

Por supuesto, en el caso hermítico tan sólo se habría de lidiar con una única familia de polinomios $P_1(z) = P_2$, $(G_{sz} = G_{sz}^\dagger \Rightarrow Z_1 = Z_2)$, y en lugar de biortogonalidad, se volvería al caso ortogonal. Puesto que los BPUC se expresan desde las matrices de la factorización, son esperables las siguientes expresiones para estos:

Proposición 38.

$$P_{1,k}(z) = \Theta_* \left[\begin{array}{cccc|c} & & & & 1 \\ & & G_{sz}^{[k]} & & z \\ & & & & \vdots \\ & & & & z^{k-1} \\ \hline (G_{sz})_{k,0} & (G_{sz})_{k,1} & \dots & (G_{sz})_{k,k-1} & z^k \end{array}\right],$$

$$P_{2,k}(z) = \Theta_* \left[\begin{array}{cccc|c} & & & & 1 \\ & & \left(G_{sz}^\dagger\right)^{[k]} & & z \\ & & & & \vdots \\ & & & & z^{k-1} \\ \hline \left(G_{sz}^\dagger\right)_{k,0} & \left(G_{sz}^\dagger\right)_{k,1} & \dots & \left(G_{sz}^\dagger\right)_{k,k-1} & z^k \end{array}\right].$$

Demostración. Para los polinomios $P_{1,k}$ la prueba es análoga a la dada para probar la proposición 2. Para probar la correspondiente a la segunda familia, basta con repetir el razonamiento pero teniendo en cuenta que $Z_2 G_{sz}^\dagger = H^\dagger Z_1^{-\dagger}$. $\square$

[2] Las matrices Toeplitz y sus determinantes llevan el nombre de Otto Toeplitz [120], [121]

Proposición 39. *Los polinomios biortogonales de Szegő satisfacen las siguientes relaciones:*

$$\begin{pmatrix} P_{1,l}(z) \\ P^*_{2,l}(z) \end{pmatrix} = \begin{pmatrix} z & \alpha_{1,l} \\ \bar{\alpha}_{2,l} z & 1 \end{pmatrix} \begin{pmatrix} P_{1,l-1}(z) \\ P^*_{2,l-1}(z) \end{pmatrix}, \qquad \begin{pmatrix} P^*_{1,l}(z) \\ P_{2,l}(z) \end{pmatrix} = \begin{pmatrix} 1 & \bar{\alpha}_{1,l} z \\ \alpha_{2,l} & z \end{pmatrix} \begin{pmatrix} P^*_{1,l-1}(z) \\ P_{2,l-1}(z) \end{pmatrix}.$$

En la siguiente sección se deducirán estas expresiones desde el contexto de las secuencias biortogonales de Laurent.

1.2 Matriz de momentos CMV y polinomios biortogonales de Laurent

Hasta el momento, las secuencias biortogonales consideradas han sido todas polinómicas. Sin embargo, para el caso de funcionales soportados sobre conjuntos del plano complejo es una opción interesante considerar polinomios de Laurent. El interés que despierta tal opción reside en que dichas secuencias tienen propiedades más parecidas a las del caso real que las que presentan los polinomios. Por ejemplo, la ley de recurrencia para los polinomios de Laurent será a cinco términos en lugar de las más exóticas del tipo de la proposición 39 y que han de echar mano de los polinomios reversos. Sea cual fuere la secuencia ortogonal elegida, ambas han de complementarse y converger en cuanto a resultados se refiere. Se tratará de mostrar esta afirmación, para lo que será esencial introducir el vector CMV de monomios $\chi_{cmv}(z)$.

Definición 37. *Reciben el nombre de vectores semi infinitos CMV de monomios los siguientes:*

$$\chi_{(1)}(z) := (1, 0, z, 0, z^2, \dots)^\top, \qquad \chi_{(2)}(z) := (0, 1, 0, z, 0, z^2, \dots)^\top, \qquad \chi^*_{(a)}(z) := z^{-1}\chi_{(a)}(z^{-1}) \qquad a = 1, 2.$$

$$\chi_{cmv}(z) := \chi_{(1)}(z) + \chi^*_{(2)}(z) = \Big((\chi_{cmv})_0(z), (\chi_{cmv})_1(z), (\chi_{cmv})_2(z), \dots \Big)^\top = (1, z^{-1}, z, z^{-2}, z^2, \dots)^\top.$$

Para no cargar la notación, en lo que resta de sección, suprimiré el subíndice que encontramos en χ_{cmv}. Dicho esto, dada una medida de Borel μ soportada sobre la circunferencia unidad, se define la siguiente matriz semi infinita.

Definición 38. *La matriz de momentos CMV es la siguiente matriz semi infinita:*

$$G := \oint_{\mathbb{T}} \chi(z) \frac{\mathrm{d}\mu(z)}{iz} \left(\chi(z)\right)^\dagger.$$

Como se viene haciendo cada vez que se ha definido una matriz de momentos, se puede afirmar, sin necesidad de probarlo que:

Proposición 40. *Siempre que* $\det G^{[k]} \neq 0$, *la matriz CMV de momentos admitirá una factorizción LU,*

$$G := S_1^{-1} K S_2^{-\dagger}.$$

En términos de las matrices de la factorización se construyen los BLPUC,

$$\phi_r(z) := S_r \chi(z) := \begin{pmatrix} \varphi_{r,0}(z) \\ \varphi_{r,1}(z) \\ \vdots \\ \varphi_{r,k} \\ \vdots \end{pmatrix}, \qquad \varphi_{r,k}(z) = \sum_{j=0}^{k} (S_r)_{k,j}\, \chi_j(z) \qquad r = 1, 2.$$

Cuya versión en términos de cuasideterminantes viene dada por:

$$\varphi_{1,k}(z) = \Theta_* \left[\begin{array}{cccc|c} & & & & \chi_0(z) \\ & & G^{[k]} & & \chi_1(z) \\ & & & & \vdots \\ & & & & \chi_{k-1}(z) \\ \hline G_{k,0} & G_{k,1} & \dots & G_{k,k-1} & \chi_k(z) \end{array}\right],$$

$$\varphi_{2,k}(z) = \Theta_* \left[\begin{array}{cccc|c} & & & & \chi_0(z) \\ & & \left(G^\dagger\right)^{[k]} & & \chi_1(z) \\ & & & & \vdots \\ & & & & \chi_{k-1}(z) \\ \hline \left(G^\dagger\right)_{k,0} & \left(G^\dagger\right)_{k,1} & \dots & \left(G^\dagger\right)_{k,k-1} & \chi_k(z) \end{array}\right].$$

La biortogonalidad de los ϕ_r ha de ser especificada.

Proposición 41. *Los BLPUC satisfacen las siguientes reglas de biortogonalidad:*

$$\begin{aligned} \langle \varphi_{1,j}, \varphi_{2,j} \rangle_\mu &= k_j, & K &= \operatorname{diag}\{k_j\}, \\ \langle \varphi_{1,2k}, z^l \rangle_\mu &= 0, & -k \leq& l \leq k-1, \\ \langle \varphi_{1,2k+1}, z^l \rangle_\mu &= 0, & -k \leq& l \leq k, \\ \langle z^l, \varphi_{2,2k} \rangle_\mu &= 0, & -k \leq& l \leq k-1, \\ \langle z^l, \varphi_{2,2k+1} \rangle_\mu &= 0, & -k \leq& l \leq k-1. \end{aligned}$$

Demostración. Para probarlas, basta con tener en cuenta las definiciones de estos polinomios de Laurent en términos de las matrices de la factorización. □

Si se comparan las relaciones de biortogonalidad de que gozan los BLPUC con aquellas que satisfacen los BPUC, se llegará a la conclusión que se resume a continuación en forma de proposición.

Proposición 42. *Los BLPUC y los BPUCS están relacionandos entre ellos como sigue:*

$$\begin{aligned} P_{1,2l}(z) &= z^l \varphi_{1,2l}(z), & P_{2,2l}(z) &= z^l \varphi_{2,2l}(z), \\ P^*_{2,2l+1}(z) &= z^{l+1} \varphi_{1,2l+1}(z), & P^*_{1,2l+1}(z) &= z^{l+1} \varphi_{2,2l+1}(z). \end{aligned}$$

Estas relaciones permiten escribir para las matrices de la factorización lo siguiente:

$$S_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ \bar{\alpha}_{2,1} & 1 & 0 & 0 & \dots \\ * & \alpha_{1,2} & 1 & 0 & \dots \\ * & * & \bar{\alpha}_{2,3} & 1 & \dots \\ * & * & * & \alpha_{1,4} & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}, \quad S_2 = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ \bar{\alpha}_{1,1} & 1 & 0 & 0 & \dots \\ * & \alpha_{2,2} & 1 & 0 & \dots \\ * & * & \bar{\alpha}_{1,3} & 1 & \dots \\ * & * & * & \alpha_{2,4} & \\ \vdots & \vdots & \vdots & & \ddots \end{pmatrix}, \quad K = H = \operatorname{diag}\{h_0, h_1, \dots\}.$$

Donde,

$$h_j = \Theta_* \left[\begin{array}{cccc|c} & & & & G_{0,k} \\ & & G^{[k]} & & G_{1,k} \\ & & & & \vdots \\ & & & & G_{k-1,k} \\ \hline G_{k,0} & G_{k,1} & \dots & G_{k,k-1} & G_{k,k} \end{array}\right] = \Theta_* \left[\begin{array}{ccc|c} & & & (G_{sz})_{0,k} \\ & G^{[k]}_{sz} & & (G_{sz})_{1,k} \\ & & & \vdots \\ & & & (G_{sz})_{k-1,k} \\ \hline (G_{sz})_{k,0} & \dots & (G_{sz})_{k,k-1} & (G_{sz})_{k,k} \end{array}\right].$$

Demostración. Probémoslo para uno de los polinomios, por ejemplo $\varphi_{1,2k}(z)$. Para comprobar el primer resultado, obsérvese que

$$\langle \varphi_{1,2k}, z^l \rangle_\mu = 0, \qquad -k \leq l \leq k-1 \qquad \Longrightarrow \qquad \langle z^k \varphi_{1,2k}, z^{l+k} \rangle_\mu = 0, \qquad 0 \leq (l+k) \leq 2k-1.$$

Renombrando $l+k=j$ es fácil ver que esta es justo la relación de ortogonalidad satisfecha por $P_{1,2k}(z)$ y puesto que sólo existe un único polinomio mónico que satisfaga dicha condición, ambos han de coincidir. En cuanto a las entradas de las matrices de la factorización, usemos el resultado que acabamos de probar:

$$P_{1,2k}(z) = z^k \varphi_{1,2k}(z) = z^k \Big(z^k + (S_1)_{2k,2k-1} z^{-(k)} + (S_1)_{2k,2k-2} z^{k-1} + \cdots + (S_1)_{2k,0} \Big).$$

Desde la definición de los coeficientes de Verblunsky no es difícil ver que $\alpha_{1,2k} := P_{1,2k}(0) = (S_1)_{2k,2k-1}$. Para el resto de polinomios ha de seguirse un proceso similar. □

Definición 39. *Se define para todo $z \neq \mathbb{T}$, el vector semi infinito de las funciones de segunda especie como las transformadas de Cauchy de los BLPUC,*

$$C_1(z) := \oint_{\mathbb{T}} \phi_1(x) \frac{\mathrm{d}\mu(x)}{ix} \overline{\left(\frac{1}{\bar{z}-x}\right)} = \langle \phi_1(x), \frac{1}{\bar{z}-x} \rangle,$$
$$(C_2(z))^\dagger := \oint_{\mathbb{T}} \left(\frac{1}{\bar{z}-x}\right) \frac{\mathrm{d}\mu(x)}{ix} (\phi_2(x))^\dagger = \langle \frac{1}{\bar{z}-x}, \phi_2(x) \rangle.$$

Proposición 43. *Las funciones de segunda especie admiten la siguiente representación en términos de las matrices de la factorización:*

$$C_1(z) = S_1 G \chi^*_{(1)}(z), \qquad C_2(z) = S_2 G^\dagger \chi^*_{(1)}(z), \qquad \forall\ 1 < |z|,$$
$$C_1(z) = -S_1 G \chi_{(2)}(z), \qquad C_2(z) = -S_2 G^\dagger \chi_{(2)}(z), \qquad \forall\ |z| < 1.$$

Demostración. Las expresiones se siguen de expresar G en su versión integral (integrando en la variable x) teniendo en cuenta lo siguiente (recuérdese que dado que $x \in \mathbb{T}$ se tiene $|x| = 1$):

$$\chi\left(\frac{1}{x}\right) \chi^*_{(1)}(z) = \frac{1}{z - \frac{1}{x}}, \qquad 1 < |z|,$$
$$\chi\left(\frac{1}{x}\right) \chi_{(2)}(z) = \frac{1}{\frac{1}{x} - z}, \qquad |z| < 1.$$

□

1.3 Simetrías de la matriz de momentos: recurrencia a cinco términos

Con el objetivo de explorar las simetrías de la matriz CMV de momentos, se define el siguiente par de matrices entorno a las que girarán los resultados principales de la sección.

Definición 40. *Se definen la matriz Υ y la matriz de entrelazamiento η como sigue:*

$$\Upsilon := \left(\begin{array}{c|cc|cc|cc|cccc}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right), \qquad
\eta := \left(\begin{array}{c|cc|cc|ccc}
1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 1 & 0 & 0 & 0 & 0 & \cdots \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 1 & 0 & 0 & \cdots \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right).$$

La importancia de las matrices que acabamos de presentar queda patente en el siguiente resultado.

Proposición 44. *Se cumplen las siguientes afirmaciones:*

$$\Upsilon^\top = \Upsilon^{-1}, \qquad \eta\eta = \mathbb{I}, \qquad \eta\Upsilon = \Upsilon^{-1}\eta.$$

La acción de Υ y de η sobre el vector CMV de monomios $\chi(z)$ son de especial relevancia:

$$\Upsilon\chi(z) = z\chi(z), \qquad \Upsilon^{-1}\chi(z) = \frac{1}{z}\chi(z), \qquad \eta\chi(z) = \chi\left(\frac{1}{z}\right),$$

Las matrices Υ y η dan cuenta de las siguientes simetrías presentes en la matriz de momentos:

$$\Upsilon G = G\Upsilon, \qquad \eta G = G^\top\eta.$$

Demostración. Los primeros dos resultados se verifican sin dificultad desde la definición de las matrices Υ y η, mientras que para probar la tercera hay que volver a la expresión integral de G y usar allí los anteriores. Por supuesto, hablar de una simetría de G es hacerlo de la forma sesquilineal asociada. Con esto en mente $\Upsilon G = G\Upsilon$ es equivalente a $\langle zf(z), h(z)\rangle_\mu = \langle f(z), \frac{1}{z}h(z)\rangle_\mu$ mientras que la simetría $\eta G = G^\top\eta$ relaciona nuestra forma sesquilineal $\langle f, h\rangle_\mu = \oint f(z)\frac{\mathrm{d}\mu}{iz}\overline{h(z)}$ con otra alternativa a esta y dada por $(f,h)_\mu = \oint \overline{f(z)}\frac{\mathrm{d}\mu}{iz}h(z)$. □

El proceso de revestir, a base de las matrices de la factorización, las matrices Υ, η y su combinación $\eta\Upsilon^p$ esclarece un gran número de las propiedades que los BLPUC satisfacen. Revestir Υ permite obtener la ley de recurrencia a cinco términos de los BLPUC; el mismo proceso pero partiendo de las matrices η y $\eta\Upsilon^{-1}$ ayuda a recuperar la ley de recurrencia de Szegő así como presentar la relación que existe entre los coeficientes de Verblunsky y las h_j. En lo que sigue, siguiendo la notación que predomina en la literatura usaremos $\rho_{k+1} := \frac{h_{k+1}}{h_k}$.

Definición 41. *Las matrices de Jacobi son:*

$$J_r := S_r\Upsilon S_r^{-1}, \qquad r = 1,2.$$

Estas matrices son de especial interés como se puede deducir de la siguiente proposición.

Proposición 45. *Las dos matrices de Jacobi están relacionadas entre si,*

$$J_1 = HJ_2^{-\dagger}H^{-1},$$

tienen una estructura pentadiagonal $(2\cdot 2+1)$,

$$J_1 := \left(\begin{array}{c|cc|cc|cc|c}
 & \ddots & & \vdots & \vdots & & & \\
 & & \ddots & \vdots & \vdots & & & \\
\hline
\cdots\cdots & \rho_{2k-1}\rho_{2k-2} & \rho_{2k-1}\bar{\alpha}_{2,2k-2} & -\bar{\alpha}_{2,2k-1}\alpha_{1,2k} & \bar{\alpha}_{2,2k-1} & 0 & 0 & \cdots\cdots\\
\cdots\cdots & 0 & 0 & -\rho_{2k}\alpha_{1,2k+1} & -\bar{\alpha}_{2,2k}\alpha_{1,2k+1} & \alpha_{1,2k+2} & 1 & \cdots\cdots\\
\hline
 & & & \vdots & \vdots & \ddots & & \\
 & & & \vdots & \vdots & & \ddots &
\end{array}\right)$$

$$J_2 := \left(\begin{array}{c|cc|cc|cc|c}
 & \ddots & & \vdots & \vdots & & & \\
 & & \ddots & \vdots & \vdots & & & \\
\hline
\cdots\cdots & \bar{\rho}_{2k-1}\bar{\rho}_{2k-2} & \bar{\rho}_{2k-1}\bar{\alpha}_{1,2k-2} & -\bar{\alpha}_{1,2k-1}\alpha_{2,2k} & \bar{\alpha}_{1,2k-1} & 0 & 0 & \cdots\cdots\\
\cdots\cdots & 0 & 0 & -\bar{\rho}_{2k}\alpha_{2,2k+1} & -\bar{\alpha}_{1,2k}\alpha_{2,2k+1} & \alpha_{2,2k+2} & 1 & \cdots\cdots\\
\hline
 & & & \vdots & \vdots & \ddots & & \\
 & & & \vdots & \vdots & & \ddots &
\end{array}\right),$$

y dan cuenta de las siguientes relaciones de recurrencia,

$$J_1\phi_1(z) = z\phi_1(z), \qquad J_2\phi_2(z) = z\phi_2(z).$$

En componentes, para la primera familia se escriben como siguen:

$$\begin{aligned} z\varphi_{1,2k-1}(z) &= \rho_{2k-1}\rho_{2k-2}\varphi_{1,2k-3}(z) + \rho_{2k-1}\bar{\alpha}_{2,2k-2}\varphi_{1,2k-2}(z) - \bar{\alpha}_{2,2k-1}\alpha_{1,2k}\varphi_{1,2k-1}(z) + \bar{\alpha}_{2,2k-1}\varphi_{1,2k}(z), \\ z\varphi_{1,2k}(z) &= -\rho_{2k}\alpha_{1,2k+1}\varphi_{1,2k-1}(z) - \bar{\alpha}_{2,2k}\alpha_{1,2k+1}\varphi_{1,2k}(z) + \alpha_{1,2k+2}\varphi_{1,2k+1}(z) + \varphi_{1,2k+2}(z). \end{aligned}$$

Para la segunda familia se tiene:

$$\begin{aligned} z\varphi_{2,2k-1}(z) &= \bar{\rho}_{2k-1}\bar{\rho}_{2k-2}\varphi_{2,2k-3}(z) + \bar{\rho}_{2k-1}\bar{\alpha}_{1,2k-2}\varphi_{2,2k-2}(z) - \bar{\alpha}_{1,2k-1}\alpha_{2,2k}\varphi_{2,2k-1}(z) + \bar{\alpha}_{1,2k-1}\varphi_{2,2k}(z), \\ z\varphi_{2,2k}(z) &= -\bar{\rho}\alpha_{2,2k+1}\varphi_{2,2k-1}(z) - \bar{\alpha}_{1,2k}\alpha_{2,2k+1}\varphi_{2,2k}(z) + \alpha_{2,2k+2}\varphi_{2,2k+1}(z) + \varphi_{2,2k+2}(z). \end{aligned}$$

Demostración. La primera parte de la proposición responde a la factorización de G en su simetría $\Upsilon G = G\Upsilon$. La recurrencia es consecuencia de la propia definición de las matrices de Jacobi junto con la acción de Υ sobre χ. Por último, la estructura pentadiagonal se sigue de la primera parte de la proposición en la que se relacionan las dos matrices de Jacobi. □

Definición 42. *Se define la siguiente matriz en términos de los factores LU de la matriz de momentos:*

$$C_p := \bar{S}_2\eta\Upsilon^p S_1^{-1}, \qquad p \in \mathbb{Z}.$$

Esta construcción matricial es interesante porque mezcla las dos matrices de la factorización LU, lo que va a permitir relacionar las ambas familias biortogonales de polinomios como se explica en la siguiente proposición.

Proposición 46. *La matriz C_p satisface*

$$C_p = HC_p^\top H^{-1},$$

y permite relacionar diferentes BLPUC como sigue,

$$C_p\phi_1(z) = z^p\bar{\phi}_2\left(\frac{1}{z}\right).$$

Demostración. La primera relación se sigue al factorizar la matriz de momentos en la simetría $\eta\Upsilon^p G = G^\top\eta\Upsilon^p$, que es consecuencia de combinar las presentes en la proposición 44. El segundo resultado se sigue de la definición 42, las definiciones de las dos familias de polinomios biortogonales en términos de las matrices de la factorización LU y por último de las acciones de las matrices η y Υ sobre el vector CMV de monomios. □

Como se puede comprobar, de todos los posibles valores que p podría tomar en C_p, los de mayor relevacia son $p = 0, -1$. Es sencillo entender el por qué de esta afirmación. El número de diagonales no nulas de C_p se va a dar en los casos en que las entradas de $\eta\Upsilon^p$ se encuentren lo más cerca posible de la diagonal principal, que es el caso para los valores $p = 0, -1$ en cuyo caso $\eta\Upsilon^p$ va a tener entradas no nulas únicamente a lo largo de las primeras super y sub diagonales.

Proposición 47. *Las entradas de C_p para los casos $p = 0, -1$ se pueden dar en función de los coeficientes de Verblunsky y las h_j:*

$$C_0 = \left(\begin{array}{c|cc|c|cc|c}
1 & & & & & & \\ \hline
 & -\alpha_{1,2} & 1 & & & & \\
 & \rho_2 & \bar{\alpha}_{2,2} & & & & \\ \hline
 & & & \ddots & & & \\
 & & & & \ddots & & \\ \hline
 & & & & -\alpha_{1,2k} & 1 & \\
 & & & & \rho_{2k} & \bar{\alpha}_{2,2k} & \\ \hline
 & & & & & & \ddots \\
 & & & & & & & \ddots
\end{array}\right),$$

$$C_{-1} = \left(\begin{array}{c|cc|cc|cc|cc}
-\bar{\alpha}_{2,1} & 1 & 0 & & & & & & \\ \hline
\rho_1 & \alpha_{1,1} & 0 & 0 & 0 & & & & \\
0 & 0 & -\bar{\alpha}_{2,1} & 1 & 0 & & & & \\ \hline
 & & & \ddots & & & & & \\
 & & & & \ddots & & & & \\ \hline
 & & & 0 & \rho_{2k-1} & \alpha_{1,2k-1} & 0 & 0 & 0 \\
 & & & 0 & 0 & 0 & -\bar{\alpha}_{2,2k+1} & 1 & 0 \\ \hline
 & & & & & & & \ddots & \\
 & & & & & & & & \ddots
\end{array}\right).$$

Lo que implica que los ρ_k han de satisfacer,

$$\rho_{k+1} = \frac{h_{k+1}}{h_k} = 1 + \bar{\alpha}_{2,k+1}\alpha_{1,k+1}.$$

Junto con las relaciones,

$$C_0\phi_1(z) = \bar{\phi}_2\left(\frac{1}{z}\right), \qquad C_{-1}\phi_1(z) = \frac{1}{z}\bar{\phi}_2\left(\frac{1}{z}\right).$$

Por componentes,

$$\begin{aligned}
\bar{\varphi}_{2,2k-1}\left(\frac{1}{z}\right) &= -\alpha_{1,2k}\varphi_{1,2k-1}(z) + \varphi_{1,2k}(z), & \frac{1}{z}\bar{\varphi}_{2,2k-1}\left(\frac{1}{z}\right) &= \rho_{2k-1}\varphi_{1,2k-2}(z) + \alpha_{1,2k-1}\varphi_{1,2k-1}(z), \\
\bar{\varphi}_{2,2k}\left(\frac{1}{z}\right) &= \rho_{2k}\varphi_{1,2k-1}(z) + \bar{\alpha}_{2,2k}\varphi_{1,2k}(z), & \frac{1}{z}\bar{\varphi}_{2,2k}\left(\frac{1}{z}\right) &= -\bar{\alpha}_{2,2k+1}\varphi_{1,2k}(z) + \varphi_{1,2k+1}(z).
\end{aligned}$$

que son equivalentes a las relaciones de recurrencia de Szegő de la proposición 39.

Demostración. La primera y segunda partes de la proposición se obtienen sin dificultad al operar para conocer las entradas de C_p para $p = 0, -1$ desde su propia definición. La tercera parte se sigue al sustituir en estas las expresiones de $\varphi_{\alpha,j}$ en función de los $P_{\beta,i}$ siguiendo la norma propuesta en la poposición 42; demos un ejemplo: tomando la primera relación y multiplicando por z^k a ambos lados de la igualdad se tiene que,

$$z^k\bar{\varphi}_{2,2k-1}\left(\frac{1}{z}\right) = -\alpha_{1,2k}z^k\varphi_{1,2k-1}(z) + z^k\varphi_{1,2k}(z),$$

pero la proposición 42 dice que,

$$z^k\varphi_{1,2k}(z) = P_{1,2k}(z), \qquad z^k\varphi_{1,2k-1}(z) = P^*_{1,2k-1}(z), \qquad z^k\bar{\varphi}_{2,2k-1}\left(\frac{1}{z}\right) = zP_{1,2k-1}(z).$$

Por lo tanto, la igualdad queda,

$$zP_{1,2k-1}(z) = -\alpha_{1,2k}P^*_{1,2k-1}(z) + P_{1,2k}(z).$$

Que coincide con una de las relaciones de recurrencia de los polinomios de Szegő. Las que restan se obtienen del mismo modo. □

1.4 El núcleo de Christoffel–Darboux

Al igual que sucede en caso real, el núcleo de Christoffel Darboux en este contexto, también va a actuar como proyector dando la mejor aproximación a una función dada, en la base de los polinomios de Laurent.

Definición 43. *Definiciones equivalentes del núcleo CMV de Christoffel–Darboux son:*

$$K^{[l]}(x,y) := \sum_{j=0}^{l-1}\frac{\overline{\varphi_{2,j}(x)}\varphi_{1,j}(y)}{h_j} = \left(\chi^{[l]}(x)\right)^\dagger\left(G^{[l]}\right)^{-1}\chi^{[l]}(y) = \left(\phi_2^{[l]}(x)\right)^\dagger\left(H^{[l]}\right)^{-1}\phi_1^{[l]}(y).$$

Por mejor aproximación en la base de los polinomios de Laurent se entiende lo siguiente:

$$\Pi_1^{[l]}[f](y) := \langle f(x), \overline{K^{[l]}(x,y)}\rangle_\mu = \sum_{j=0}^{l-1}\beta_j\varphi_{1,j}(y), \qquad \beta_j = \frac{\langle f,\varphi_{2,j}\rangle_\mu}{h_j}.$$

$$\overline{\Pi_2^{[l]}[f](x)} := \langle K^{[l]}(x,y), f(y)\rangle_\mu = \overline{\sum_{j=0}^{l-1}\nu_j\varphi_{2,j}(y)}, \qquad \bar{\nu}_j = \frac{\langle\varphi_{1,j}, f\rangle_\mu}{h_j}.$$

Además, tratándose de un proyector satisface,

$$\langle K^{[l]}(x,z), \overline{K^{[l]}(z,y)}\rangle_\mu = K^{[l]}(x,y).$$

Proposición 48. *Las siguientes expresiones alternativas para el núcleo de Christoffel–Darboux en términos de un pequeño y constante número de los BLPUC se cumplen:*

$$\begin{aligned}
(1-y\bar{x})K^{[2k]}(x,y) &= \frac{\bar{x}}{h_{2k-1}}\left[\overline{\varphi_{2,2k+1}(x)} - \alpha_{1,2k-1}\overline{\varphi_{2,2k}(x)}\right]\varphi_{1,2k-1}(y)\\
&\quad - \frac{\bar{x}}{h_{2k-1}}\left[\rho_{2k-1}\overline{\varphi_{2,2k-2}(x)} + \bar{\alpha}_{2,2k-1}\overline{\varphi_{2,2k-1}(x)}\right]\varphi_{1,2k}(y)\\
&= \frac{\varphi_{1,2k}\left(\frac{1}{\bar{x}}\right)\varphi_{1,2k-1}(y) - \varphi_{1,2k-1}\left(\frac{1}{\bar{x}}\right)\varphi_{1,2k}(y)}{h_{2k-1}},\\
(1-y\bar{x})K^{[2k+1]}(x,y) &= \frac{\bar{x}}{h_{2k}}\left[\rho_{2k}\varphi_{1,2k-1}(y) + \bar{\alpha}_{2,2k}\varphi_{1,2k}(y)\right]\overline{\varphi_{2,2k+1}(x)}\\
&\quad + \frac{\bar{x}}{h_{2k}}\left[\varphi_{1,2k+2}(y) + \alpha_{1,2k+2}\varphi_{1,2k+1}(y)\right]\overline{\varphi_{2,2k}(x)}\\
&= \bar{x}\frac{\bar{\varphi}_{2,2k}\left(\frac{1}{y}\right)\overline{\varphi_{2,2k+1}(x)} - \bar{\varphi}_{2,2k+1}\left(\frac{1}{y}\right)\overline{\varphi_{2,2k}(x)}}{h_{2k}}.
\end{aligned}$$

Demostración. Este conjunto de expresiones se siguen de las relaciones de recurrencia. Las primeras se deducen desde la recurrencia dada por las matrices de Jacobi J_r, mientras que las segundas se prueban partiendo de las recurrencias que se obtienen al emplear C_0, C_{-1}. □

1.5 Deformaciones

A la hora de considerar transformaciones discretas, que recordamos vendrán dadas en términos de polinomios de Laurent, se ha de prestar atención a la siguiente observación: la base de polinomios de Laurent no es un conjunto graduado. Consideremos un polinomio de Laurent y tomemos el módulo de todos y cada uno de los exponentes de las potencias de z que aparecen en el mismo. Podemos llamar longitud del polinomio de Laurent al valor más alto de este conjunto. Entonces, si tenemos dos polinomios de Laurent L_1, L_2 de longitudes ℓ_1 y ℓ_2 respectivamente su producto no tiene por qué tener por longitud $\ell_1 + \ell_2$. Nos interesaremos por el subconjunto de polinomios de Laurent que sí satisfacen dicha propiedad. El método que presentamos en [25] y [26], dedicados ambos al estudio de las deformaciones en cuestión, es únicamente válido para deformaciones producidas a base de este tipo de polinomios. Para un estudio detallado de la conexión entre las deformaciones continuas y las jerarquías integrables asociadas a la red de Toda desde el enfoque de las técnicas de la factorización LU, se recomienda la lectura del artículo [12].

2 Biortogonalidad generalizada en la circunferencia unidad

Al igual que se hizo al final de la parte anterior en el contexto de la recta real, en el capítulo que sigue trataré de motivar al lector a considerar nuestras publicaciones en las que, bajo la luz de la factorización LU, extendemos las técnicas previamente explicadas para la biortogonalidad escalar en la circunferencia unidad a los casos matricial y multivariable. Habiendo ya pasado por este proceso en el contexto de la recta real, ahora seré más breve deteniéndome únicamente en aquellas particularidades de especial relevancia.

2.1 Polinomios matriciales biortogonales en la circunferencia unidad

- **Forma sesquilineal** Partiendo de una medida matricial $\mu(z)$ de tamaño $m \times m$ se presentan las dos siguientes formas sesquilineales matriciales:

$$\langle\langle f, g\rangle\rangle_L := \oint_{\mathbb{T}} g(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}\, z}f(z)^\dagger, \qquad \langle\langle f, g\rangle\rangle_R := \oint_{\mathbb{T}} f(z)^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}\, z}g(z).$$

  Nótese cómo en este caso (a diferencia del caso matricial sobre la recta real) sí que es preciso considerar, para una única medida matricial, dos formas sesquilineales (una adaptada al módulo izquierdo y la otra al derecho) que no van a ser equivalentes. El núcleo de esta cuestión reside en que cada una de las formas sesquilineales va a tener su respectiva matriz de momentos asociada; estas matrices de momentos aun estando relacionadas (como veremos luego), no son equivalentes.

- **Matrices de momentos** G^L, G^R. Dos formas sesquilineales supone tener dos matrices de momentos. Para construirlas se precisa de la versión matricial del vector CMV de monomios (ver la definición 37) $\chi_{cmv}(z\mathbb{I}_{m\times m})$.

$$G_L := \oint_{\mathbb{T}} \chi(z)\,\frac{\mathrm{d}\mu(z)}{\mathrm{i}\, z}\,(\chi(z))^\dagger, \qquad G_R := \oint_{\mathbb{T}} (\chi(z)^\top)^\dagger\,\frac{\mathrm{d}\mu(z)}{\mathrm{i}\, z}\,\chi(z)^\top.$$

- **Factorización LU**. Dos matrices de momentos suponen dos pares de familias de polinomios de Laurent biortogonales, cada par consecuencia de la factorización LU por bloques de tamaño $m \times m$ que le corresopnda.

$$G_L := S_1^{-1}D_L\widehat{S_2} = S_1^{-1}S_2, \qquad G_R := Z_2D_R\widehat{Z_1}^{-1} = Z_2Z_1^{-1}.$$

  Donde S_1, Z_2 son matrices unitriangulares inferiores por bloques mientras que $\widehat{S_2}, \widehat{Z_1}$ son unitriangulares superiores (por unitriangulares en este contexto de bloques de tamaño $m \times m$ se entiende que a lo largo de la diagonal de estas se va repitiendo la matriz identidad de dicho tamaño). Como se habrá observado, la notación que usamos aquí difiere levemente de la que venimos usando. El uso de esta queda justificado por ser más adaptada a la situación, poniendo en evidencia de una manera más clara la diferencia entre el módulo por la izquierda y el correspondiente por la derecha que conviven en la teoría. Se denotará a las cuatro familias de polinomios biortogonales como sigue:

$$\phi_j^L =: \begin{pmatrix} (\varphi_j^L)^{(0)}(z) \\ (\varphi_j^L)^{(1)}(z) \\ \vdots \end{pmatrix}, \qquad \phi_j^R =: ((\varphi_j^R)^{(0)}(z), (\varphi_j^R)^{(1)}(z), \dots), \qquad j = 1, 2.$$

Dados en términos de las matrices de factorización como:

$$\begin{aligned} \phi_1^L &:= S_1 \chi(z), & \phi_2^L &:= \left(S_2^{-1}\right)^\dagger \chi(z), \\ \phi_1^R &:= \chi^\top(z) Z_1, & \phi_2^R &:= \chi^\top(z) \left(Z_2^{-1}\right)^\dagger. \end{aligned}$$

Y que satisfacen las relaciones siguientes de biortogonalidad:

$$\langle\!\langle (\varphi_2^H)^{(j)}, (\varphi_1^H)^{(k)} \rangle\!\rangle_K = \mathbb{I}\delta_{j,k}, \qquad H = L, R, \qquad j, k = 0, 1, \dots$$

Al igual que se hizo en el caso escalar, al comparar estas relaciones de biortogonalidad con las que satisfacen los polinomios biortogonales matriciales de Szegő, se llega a identificaciones entre estos, lo que permite dar algunas de las entradas de las matrices de la factorización, S_j, Z_j con $j = 1, 2$, en términos de los parámetros de Verblunsky.

- **Simetrías de** G^L, G^R. Desde luego, es esperable una simetría que involucre al análogo matricial del operador Υ (ver definición 40) siendo esta responsable de las relaciones a cinco bloques (matriciales) de los polinomios biortogonales así como las fórmulas para los núcleos de Christoffel–Darboux. Algo menos esperable es que la versión matricial del operador de entrelazamiento η (ver la definición 40) consigue mezclar en la misma ecuación a las dos matrices de momentos asociadas a la misma medida matricial, lo cual permite a su vez relacionar los dos pares de familias biortogonales entre sí.
- **Deformaciones de** G^L, G^R. Como principal particularidad de las deformaciones en este caso, a parte de las obvias a tener en cuenta para las deformaciones discetas: se trata de un caso matricial (el teorema fundamental no se cumple y se precisa además la información espectral de los polinomios involucrados) y se involucra a polinomios de Laurent (recuérdese que no forman una base graduada) se ha de ser cuidadoso con un matiz adicional, aplicable tanto a las deformaciones continuas como a las discretas. Dicho matiz consiste en que las deformaciones han de ser compatibles con las dos matrices de momentos, es decir, las dos matrices de momentos deformadas tienen que seguir estando ligadas mediante el operador de entrelazamiento η como lo estaban las originales.

2.2 Polinomios biortogonales en varias variables en el Toro unitario

- **Forma sesquilineal**. Se denota mediante $\mathbb{T}^D$ al toro unitario (producto cartesiano de D copias de la circunferencia unidad) y se emplea $\mathrm{d}\mu \in \mathcal{B}(\Omega)$ para referirse a una medida de Borel con soporte sobre $\mathbb{T}^D$. También, para $\boldsymbol{z} \in \mathbb{T}^D$, será útil la parametrización siguiente,

$$\boldsymbol{z}(\boldsymbol{\theta}) = \left(\mathrm{e}^{\mathrm{i}\,\theta_1}, \dots, \mathrm{e}^{\mathrm{i}\,\theta_D}\right)^\top,$$

con $\boldsymbol{\theta} := (\theta_1, \dots, \theta_D)^\top \in [0, 2\pi)^D$. La forma sesquilineal entre dos funciones $f(\boldsymbol{z})$ y $g(\boldsymbol{z})$ se define a continuación:

$$\langle f, g \rangle_\mu := \oint_{\mathbb{T}^D} f(\boldsymbol{z}(\boldsymbol{\theta})) \overline{g(\boldsymbol{z}(\boldsymbol{\theta}))} \mathrm{d}\mu(\boldsymbol{\theta}).$$

- **Matriz de momentos** G. Como se hizo en el caso multivariable real, es necesario un multi-índice $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_D)^\top \in \mathbb{Z}^D$ para poder denotar los monomios $\boldsymbol{z}^{\boldsymbol{\alpha}} = z_1^{\alpha_1} \cdots z_D^{\alpha_D}$ y así poder expresar cualquier polinomio de Laurent como $L = \sum_{\boldsymbol{\alpha}} L_{\boldsymbol{\alpha}} \boldsymbol{z}^{\boldsymbol{\alpha}}$. Llamaremos longitud del multi-índice al valor $|\boldsymbol{\alpha}| := \sum_{a=1}^{D} |\alpha_a|$. [1] Para dotar de cierto orden, relacionado con su longitud, a los monomios (inspirados por el caso CMV en una variable en el que se agrupan las potencias de z^n y z^{-n}) del vector CMV de monomios, se define, para un entero positivo $k \in \mathbb{Z}_+$ el siguiente conjunto de multi-índices:

$$[k] := \{\boldsymbol{\alpha} \in \mathbb{Z}^D : |\boldsymbol{\alpha}| = k\}, \qquad |[k]| = \sum_{j=1}^{\text{mín}(k,D)} 2^j \binom{D}{j}\binom{k-1}{j-1}.$$

Este permite en primer lugar ordenar dos monomios de Laurent arbitrarios de acuerdo a la longitud de sus multi-índices, es decir, se entenderá que $z^{\boldsymbol{\alpha}} < z^{\boldsymbol{\alpha}'}$ siempre que $|\boldsymbol{\alpha}| < |\boldsymbol{\alpha}'|$. En segundo lugar, denotando por $\boldsymbol{\alpha}_j^{(k)}$ a los multi-índices cuya longitud es $|\boldsymbol{\alpha}_j^{(k)}| = k$ y donde $j = 1, 2, \dots, |[k]|$, aquellos se ordenarán de acuerdo al orden lexicográfico:

$$[k] = \left\{\boldsymbol{\alpha}_1^{(k)}, \boldsymbol{\alpha}_2^{(k)}, \dots, \boldsymbol{\alpha}_{|[k]|}^{(k)}\right\} \text{ donde } \boldsymbol{\alpha}_j^{(k)} < \boldsymbol{\alpha}_{j+1}^{(k)}.$$

Estipulada esta organización de los monomios, se procede a construir el vector CMV multivariable de monomios.

$$\chi := \begin{pmatrix} \chi_{[0]}^{(k)} \\ \chi_{[1]}^{(k)} \\ \vdots \\ \chi_{[k]}^{(k)} \\ \vdots \end{pmatrix}, \qquad \chi_{[k]} := \begin{pmatrix} \boldsymbol{z}^{\boldsymbol{\alpha}_1} \\ \boldsymbol{z}^{\boldsymbol{\alpha}_2} \\ \vdots \\ \boldsymbol{z}^{\boldsymbol{\alpha}_{|[k]|}} \end{pmatrix}.$$

Gracias a estos, la matriz de momentos se construye sin demasiado problema:

$$G := \oint_{\mathbb{T}^D} \chi(\boldsymbol{z}(\boldsymbol{\theta}))\mathrm{d}\mu(\boldsymbol{\theta})\chi(\boldsymbol{z}(-\boldsymbol{\theta}))^\top = \begin{pmatrix} G_{[0],[0]} & G_{[0],[1]} & \cdots \\ G_{[1],[0]} & G_{[1],[1]} & \cdots \\ \vdots & \vdots & \end{pmatrix},$$

$$G_{[k],[l]} := \oint_{\mathbb{T}^D} \chi_{[k]}(\boldsymbol{z}(\boldsymbol{\theta}))\mathrm{d}\mu(\boldsymbol{\theta})\chi_{[l]}(\boldsymbol{z}(-\boldsymbol{\theta}))^\top = \begin{pmatrix} G_{\boldsymbol{\alpha}_1^{(k)},\boldsymbol{\alpha}_1^{(l)}} & \dots & G_{\boldsymbol{\alpha}_1^{(k)},\boldsymbol{\alpha}_{|[l]|}^{(l)}} \\ \vdots & & \vdots \\ G_{\boldsymbol{\alpha}_{|[k]|}^{(k)},\boldsymbol{\alpha}_1^{(l)}} & \dots & G_{\boldsymbol{\alpha}_{|[k]|}^{(k)},\boldsymbol{\alpha}_{|[l]|}^{(l)}} \end{pmatrix} \in \mathbb{C}^{|[k]|\times|[l]|},$$

$$G_{\boldsymbol{\alpha}_i^{(k)},\boldsymbol{\alpha}_j^{(l)}} = \oint_{\mathbb{T}^D} \mathrm{e}^{\mathrm{i}(\boldsymbol{\alpha}_i^{(k)}-\boldsymbol{\alpha}_j^{(l)})\cdot\boldsymbol{\theta}}\,\mathrm{d}\mu(\boldsymbol{\theta}) \in \mathbb{C}.$$

- **Factorización LU**. El tipo de factorización LU más adaptado a este caso, al igual que sucedió al considerar el caso real multivariable, es uno por bloques de tamaño creciente dado por el cardinal de los conjuntos $[k]$, es decir, de acuerdo con (2.1) se tiene $n_k = |[k]|$.

$$G = S^{-1}H\big(\hat{S}^{-1}\big)^\dagger,$$

[1] Una vez más, nótese que para el anillo de los polinomios $\mathbb{C}[\boldsymbol{z}]$ la longitud coincide con el grado total del polinomio (se trata de un anillo graduado), mientras que en caso del anillo de los polinomios de Laurent la longitud del producto de dos polinomios de Laurent no tiene por qué ser la suma de las longitudes de los factores. Sólo se podrá asegurar que $\ell(L_1L_2) \leq \ell(L_1) + \ell(L_2)$.

Desde donde se definen los polinomios multivariables de Laurent en D variables complejas:

$$\Phi := S\chi = \begin{pmatrix} \phi_{[0]} \\ \phi_{[1]} \\ \vdots \end{pmatrix}, \quad \phi_{[k]}(\boldsymbol{z}) = \sum_{l=0}^{k} S_{[k],[l]}\chi_{[l]}(\boldsymbol{z}) = \begin{pmatrix} \phi_{\boldsymbol{\alpha}_1^{(k)}} \\ \vdots \\ \phi_{\boldsymbol{\alpha}_{|[k]|}^{(k)}} \end{pmatrix}, \quad \phi_{\boldsymbol{\alpha}_i^{(k)}} = \sum_{l=0}^{k}\sum_{j=1}^{|[l]|} S_{\boldsymbol{\alpha}_i^{(k)},\boldsymbol{\alpha}_j^{(l)}} \boldsymbol{z}^{\boldsymbol{\alpha}_j^{(l)}},$$

$$\hat{\Phi} := \hat{S}\chi = \begin{pmatrix} \hat{\phi}_{[0]} \\ \hat{\phi}_{[1]} \\ \vdots \end{pmatrix}, \quad \hat{\phi}_{[k]}(\boldsymbol{z}) = \sum_{l=0}^{k} \hat{S}_{[k],[l]}\chi_{[l]}(\boldsymbol{z}) = \begin{pmatrix} \hat{\phi}_{\boldsymbol{\alpha}_1^{(k)}} \\ \vdots \\ \hat{\phi}_{\boldsymbol{\alpha}_{|[k]|}^{(k)}} \end{pmatrix}, \quad \hat{\phi}_{\boldsymbol{\alpha}_i^{(k)}} = \sum_{l=0}^{k}\sum_{j=1}^{|[l]|} S_{\boldsymbol{\alpha}_i^{(k)},\boldsymbol{\alpha}_j^{(l)}} \boldsymbol{z}^{\boldsymbol{\alpha}_j^{(l)}}.$$

Estas dos familias son biortogonales entre si,

$$\oint_{\mathbb{T}^D} \phi_{[k]}(\boldsymbol{z}(\boldsymbol{\theta}))\mathrm{d}\mu(\boldsymbol{\theta})\big(\hat{\phi}_{[l]}(\boldsymbol{z}(\boldsymbol{\theta}))\big)^{\dagger} = \delta_{k,l}H_{[k]}.$$

- **Simetrías** de G. Por un lado estará presente la persimetría de la matriz de momentos, modelada por la versión multivariable del operador de entrelazamiento η y por otro, serán precisas D matrices Υ_a a partir de las que se seguirán las leyes de recurrencia y las fórmulas de Christoffel–Darboux así como relaciones entre las dos familias de MVBLPUC.

- **Deformaciones** de G. Como se destacó en el caso de una variable, el hecho de que el anillo de los polinomios de Laurent no sea graduado tiene sus consecuencias a la hora de considerar las transformaciones discretas. Sólo hemos sido capaces de encargarnos de aquellas deformaciones discretas producidas por el subconjunto de polinomios de Laurent que llamamos *nice* o equilibrado (que si es graduado y que caracterizamos mediante nociones básicas de geometría tropical). Las deformaciones continuas van a permitir mostrar que los coeficientes de los MVBLPUC (de tamaño variable) son solución de ecuaciones de tipo Toda. A base de desplazamientos de Miwa se podrán recuperar únicamente deformaciones discretas producidas por polinomios de Laurent de longitud igual a la unidad (recordemos que los polinomios irreducibles en varias variables pueden tener cualquier longitud).

2.3 Biortogonalidad de tipo múltiple y Sobolev en la circunferencia

Los dos tipos de polinomios biortogonales incluidos en el título han quedado sin su análisis bajo la perspectiva LU. Con todas las situaciones ya consideradas así como con la experiencia ganada en el proceso que ha supuesto llegar hasta este punto, puedo afirmar que el estudio de este par de temas sería accesible precisando únicamente de una cantidad finita de trabajo y motivación.

Parte III

Artículos incluidos

1 Christoffel Transformations for Matrix Orthogonal Polynomials in the Real Line and the non-Abelian 2D Toda Lattice Hierarchy

C. Álvarez-Fernández *et al.* (2017) "Christoffel Transformations for MOPRL and Toda Lattice Hierarchy," International Mathematics Research Notices, Vol. 2017, No. 5, pp. 1285–1341
Advance Access publication May 23, 2016
doi: 10.1093/imrn/rnw027

Christoffel Transformations for Matrix Orthogonal Polynomials in the Real Line and the non-Abelian 2D Toda Lattice Hierarchy

Carlos Álvarez-Fernández[1], Gerardo Ariznabarreta[2], Juan Carlos García-Ardila[3], Manuel Mañas[2,*] and Francisco Marcellán[3,4]

[1]Departamento de Métodos Cuantitativos, Universidad Pontificia Comillas, Calle de Alberto Aguilera 23, 28015-Madrid, [2]Departamento de Física Teórica II (Métodos Matemáticos de la Física), Universidad Complutense de Madrid, Ciudad Universitaria, Plaza de Ciencias 1, 28040-Madrid, [3]Departamento de Matemáticas, Universidad Carlos III de Madrid, Avenida Universidad 30, 28911 Leganés, and [4]Instituto de Ciencias Matemáticas (ICMAT), Avenida Universidad 30, 28911 Leganés, Spain

Correspondence to be sent to: e-mail: manuel.manas@ucm.es

Given a matrix polynomial $W(x)$, matrix bi-orthogonal polynomials with respect to the sesquilinear form

$$\langle P(x), Q(x)\rangle_W = \int P(x)W(x)\,\mathrm{d}\mu(x)(Q(x))^\top, \quad P, Q \in \mathbb{R}^{p\times p}[x],$$

where $\mu(x)$ is a matrix of Borel measures supported in some infinite subset of the real line, are considered. Connection formulas between the sequences of matrix bi-orthogonal polynomials with respect to $\langle\cdot,\cdot\rangle_W$ and matrix polynomials orthogonal with respect to $\mu(x)$ are presented. In particular, for the case of nonsingular leading coefficients of the perturbation matrix polynomial $W(x)$ we present a generalization of the Christoffel formula constructed in terms of the Jordan chains of $W(x)$. For

Received November 17, 2015; Revised January 28, 2016; Accepted February 11, 2016
Communicated by Prof. Percy Deift

perturbations with a singular leading coefficient, several examples by Durán and coworkers are revisited. Finally, we extend these results to the non-Abelian 2D Toda lattice hierarchy.

1 Introduction

This paper is devoted to the extension of the Christoffel formula to the Matrix Orthogonal Polynomials on the Real Line (MOPRL) and the non-Abelian 2D Toda lattice hierarchy.

1.1 Historical background and state of the art

In 1858 the German mathematician Elwin Christoffel [32] was interested, in the framework of Gaussian quadrature rules, in finding explicit formulas relating the corresponding sequences of orthogonal polynomials with respect to two measures $\mathrm{d}\mu$ (in the Christoffel's discussion was just the Lebesgue measure $\mathrm{d}\mu = \mathrm{d}x$) and $d\hat{\mu}(x) = p(x)d\mu(x)$, with $p(x) = (x-q_1)\cdots(x-q_N)$ a signed polynomial in the support of $\mathrm{d}\mu$, as well as the distribution of their zeros as nodes in such quadrature rules, see [101]. The so-called Christoffel formula is a very elegant formula from a mathematical point of view, and is a classical result which can be found in a number of orthogonal polynomials textbooks, see for example [29, 51, 95]. Despite these facts, we must mention that for computational and numerical purposes it is not so practical, see [51]. These transformations have been extended from measures to the more general setting of linear functionals. In the theory of orthogonal polynomials with respect to a moment linear functional $u \in \left(\mathbb{R}[x]\right)'$, an element of the algebraic dual (which coincides with the topological dual) of the linear space $\mathbb{R}[x]$ of polynomials with real coefficients. Given a positive definite linear moment functional, that is $\left|\left(\langle u, x^{n+m}\rangle\right)_{n,m=0}^{k}\right| > 0$, $\forall k \in \mathbb{Z}_+ := \{0, 1.2, \dots\}$, there exists a nontrivial probability measure μ such that (see [9, 29, 95]) $\langle u, x^n\rangle = \int x^m \,\mathrm{d}\mu(x)$. Given a moment linear functional u, its canonical or elementary Christoffel transformation is a new moment functional given by $\hat{u} = (x-a)u$ with $a \in \mathbb{R}$, see [24, 29, 97]. The right inverse of a Christoffel transformation is called the Geronimus transformation. In other words, if you have a moment linear functional u, its elementary or canonical Geronimus transformation is a new moment linear functional $\check{u}$ such that $(x-a)\check{u} = u$. Notice that in this case $\check{u}$ depends on a free parameter, see [54, 78]. The right inverse of a general Christoffel transformation is said to be a multiple Geronimus transformation, see [40]. All these transformations are referred as Darboux transformations, a name that was first given in the context of integrable systems in [77]. In 1878 the French mathematician Gaston Darboux, when studying the Sturm–Liouville theory in [35], explicitly treated these transformations, which appeared for the first time in [82]. In the framework of orthogonal polynomials on the real line, such a factorization of Jacobi matrices has been studied in [24, 97]. They also play an important role in the analysis of bispectral problems, see [58] and [57].

An important aspect of canonical Christoffel transformations is its relations with LU factorization, in terms of lower and upper tringular matrix factors (and its flipped version, an UL factorization) of the Jacobi matrix. A sequence of monic polynomials $\{P_n(x)\}_{n=0}^{\infty}$ associated with a nontrivial probability measure μ satisfies a three-term recurrence relation (TTRR, in short) $xP_n(x) = P_{n+1}(x) + b_nP_n(x) + a_n^2P_{n-1}(x), n \geq 0$, with the convention $P_{-1}(x) = 0$. If we denote by $P(x) = [P_0(x), P_1(x), \dots]^T$, then the matrix representation of the multiplication operator by x is directly deduced from the TTRR and reads $xP(x) = JP(x)$, where J is a tridiagonal semi-infinite matrix such that the entries in the upper diagonal are the unity. Assuming that a is a real number off the support of μ, then you have a factorization $J - aI = LU$, where L and U are, respectively, lower unitriangular and upper triangular matrices. The important observation is that the matrix $\hat{J}$ defined by $\hat{J} - aI = UL$ is again a Jacobi matrix and the corresponding sequence of monic polynomials $\{\hat{P}_n(x)\}_{n=0}^{\infty}$ associated with the multiplication operator defined by $\hat{J}$ is orthogonal with respect to the canonical Christoffel transformation of the measure μ defined as above.

For a moment linear functional u, the Stieltjes function $S(x) := \sum_{n=0}^{\infty} \frac{\langle u, x^n \rangle}{x^{n+1}}$ plays an important role in the theory of orthogonal polynomials, due to its close relation with the measure associated with u as well as its (rational) Padé Approximation, see [23, 63]. If you consider the canonical Christoffel transformation $\hat{u}$ of the linear functional u, then its Stieltjes function is $\hat{S}(x) = (x - a)S(x) - u_0$. This is a particular case of the spectral linear transformations studied in [102].

Given a bilinear form $L : \mathbb{R}[x] \times \mathbb{R}[x] \to \mathbb{R}$ one could consider the following non-symmetric and symmetric bilinear perturbations

$$\tilde{L}_1(p, q) = L(wp, q), \qquad \tilde{L}_2(p, q) = L(p, wq), \qquad \hat{L}(p, q) = L(wp, wq),$$

where $w(x)$ is a polynomial. The study of these perturbations can be found in [25]. Taking into account the matrix representation of the multiplication operator by z is a Hessenberg matrix, the authors establish a relation between the Hessenberg matrices associated with the initial and the perturbed functional by using LU and QR factorization, in terms of orthogonal and an upper triangular matrices. They also give some algebraic relations between the sequences of orthogonal polynomials associated with those bilinear forms. The above perturbations can be seen as an extension of the Christoffel transformation for bilinear forms. When the bilinear form is defined by a nontrivial probability measure supported on the unit circle, Christoffel transformations have been studied in [26] in the framework of Cantero-Moral-Valázquez (CMV) matrices, that is, the matrix representation of the multiplication operator by z in terms of an orthonormal Laurent polynomial basis. Therein, the authors state the explicit relation between the sequences of orthonormal Laurent polynomials associated with a measure and its Christoffel transformation, as well as its link with QR factorizations of such CMV matrices.

The theory of scalar orthogonal polynomials with respect to probability measures supported either on the real line or the unit circle is a standard and classic topic in approximation theory and it also has remarkable applications in many domains as discrete mathematics, spectral theory of linear differential operators, numerical integration, integrable systems, among others. Some extensions of such a theory have been developed more recently. One of the most exciting generalizations appears when you consider non-negative Hermitian-valued matrix of measures of size $p \times p$ on a σ-algebra of subsets of a space Ω such that each entry is countably additive and you are interested in the analysis of the Hilbert space of matrix-valued functions of size $p \times p$ under the inner product associated with such a matrix of measures. This question appears in the framework of weakly stationary processes, see [89]. Notice that such an inner product pays the penalty of the non-commutativity of matrices as well as the existence of singular matrices with respect to the scalar case. By using the standard Gram–Schmidt method for the canonical linear basis of the linear space of polynomials with matrix coefficients a theory of matrix orthogonal polynomials can be studied. The paper by M. G. Krein [68] is credited as the first contribution in this topic. Despite they have been sporadically studied during the last half century, there is an exhaustive bibliography focused on inner products defined on the linear space of polynomials with matrix coefficients as well as on the existence and analytic properties of the corresponding sequences of matrix orthogonal polynomials in the real line (see [42, 43, 80, 88, 94]) and their applications in Gaussian quadrature for matrix-valued functions [93], scattering theory [14, 53], and system theory [50]. The work [33] constitutes an updated overview on these topics.

But, more recently, an intensive attention was paid to the spectral analysis of second-order linear differential operators with matrix polynomials as coefficients. This work was motivated by the Bochner's characterization of classical orthogonal polynomials (Hermite, Laguerre, and Jacobi) as eigenfunctions of second-order linear differential equations with polynomial coefficients. The matrix case gives a more rich set of solutions. From the pioneering work [44] some substantial progress has been done in the study of families of matrix orthogonal polynomials associated with second-order linear differential operators as eigenfunctions and their structural properties (see [43, 59, 60] as well as the survey [45]). Moreover, in [27] the authors showed that there exist sequences of orthogonal polynomials satisfying a first-order linear matrix differential equation that constitutes a remarkable difference with the scalar case where such a situation does not appear. The spectral problem for second-order linear difference operators with polynomial coefficients has been considered in [13] as a first step in the general approach. Therein, four families of matrix orthogonal polynomials (as matrix relatives of Charlier, Meixner, Krawtchouk scalar polynomials and another one that seems not have any scalar relative) are obtained as illustrative examples of the method described therein.

It is also a remarkable fact that matrix orthogonal polynomials appear in the analysis of nonstandard inner products in the scalar case. Indeed, from the study of higher order recurrence relations that some sequences of orthogonal polynomials satisfy (see [44] where the corresponding inner product is analyzed as an extension of the Favard's theorem and [48], where the connection with matrix orthogonal polynomials is stated), to the relation between standard scalar polynomials associated with measures supported on harmonic algebraic curves and matrix orthogonal polynomials deduced by a splitting process of the first ones (see [74]) you get an extra motivation for the study of matrix orthogonal polynomials. Matrix orthogonal polynomials appear in the framework of orthogonal polynomials in several variables when the lexicographical order is introduced. Notice that in such a case, the moment matrix has a Hankel block matrix where each block is a Hankel matrix, that is, it has a doubly Hankel structure, see [39].

Concerning spectral transformations, in [40] the authors show that the so-called multiple Geronimus transformations of a measure supported in the real line yield a simple Geronimus transformation for a matrix of measures. This approach is based on the analysis of general inner products $\langle\cdot,\cdot\rangle$ such that the multiplication by a polynomial operator h is symmetric and satisfies an extra condition $\langle h(x)p(x), q(x)\rangle = \int p(x)q(x)\,\mathrm{d}\mu(x)$, where μ is a nontrivial probability measure supported on the real line. The connection between the Jacobi matrix associated with the sequence of scalar polynomials with respect to μ and the Hessenberg matrix associated with the multiplication operator by h is given in terms of the so-called *UL* factorizations. Notice that the connection between the Darboux process and the noncommutative bispectral problem has been discussed in [56]. The analysis of perturbations on the entries of the matrix of moments from the point of view of the relations between the corresponding sequences of matrix orthogonal polynomials was done in [30].

The seminal work of the Japanese mathematician Mikio Sato [91, 92] and later on of the Kyoto school [36–38] settled the basis for a Grasmannian and Lie group theoretical description of integrable hierarchies. Not much later Motohico Mulase [83] gave a mathematical description of factorization problems, dressing procedure, and linear systems as the keys for integrability. It was not necessary to wait too long, in the development of integrable systems theory, to find multicomponent versions of the integrable Toda equations, [98–100] which later on played a prominent role in the connection with orthogonal polynomials and differential geometry. The multicomponent versions of the Kadomtsev-Petviashvili (KP) hierarchy were analyzed in [21, 22] and [62, 70, 71] and in [72, 73] we can find a further study of the multi-component Toda lattice hierarchy, block Hankel/Toeplitz reductions, discrete flows, additional symmetries, and dispersionless limits. For the relation with multiple orthogonal polynomials see [8, 11].

The work of Mark Adler and Pierre van Moerbeke was fundamental to the connection between integrable systems and orthogonal polynomials. They showed that the Gauss–Borel factorization problem is the keystone for this connection. In particular, their studies in the papers on the 2D Toda hierarchy and what they called the discrete KP hierarchy [3–7] clearly established—from a group-theoretical setup—why standard orthogonality of polynomials and integrability of nonlinear equations of Toda type where so close.

The relation of multicomponent Toda systems or non-Abelian versions of Toda equations with matrix orthogonal polynomials was studied, for example, in [11, 80] (on the real line) and in [17, 81] (on the unit circle).

The approach to the Christoffel transformations in this paper, which is based on the Gauss–Borel factorization problem, has been used before in different contexts. It has been applied for the construction of discrete integrable systems connected with orthogonal polynomials of diverse types,

(i) The case of multiple orthogonal polynomials and multicomponent Toda was analyzed in [12].
(ii) In [15], we dealt with the case of matrix orthogonal Laurent polynomials on the circle and CMV orderings.
(iii) For orthogonal polynomials in several real variables see [16, 17] and [18] for orthogonal polynomials on the unit torus and the multivariate extension of the CMV ordering.

It is well known that there is a deep connection between discrete integrable systems and Darboux transformations of continuous integrable systems (see e.g., [41]). Finally, let us comment that, in the realm of several variables, in [17–19] one can find extensions of the Christoffel formula to the multivariate scenario with real variables and on the unit torus, respectively.

1.2 Objectives, results, and layout of the paper

In this contribution, we focus our attention on the study of Christoffel transformations (Darboux transformations in the language of integrable systems [77], or Lévy transformations in the language of differential geometry [49]) for matrix sesquilinear forms. More precisely, given a matrix of measures $\mu(x)$ and a matrix polynomial $W(x)$ we are going to deal with the following matrix sesquilinear forms

$$\langle P(x), Q(x)\rangle_W = \int P(x)W(x)\,\mathrm{d}\mu(x)(Q(x))^\top. \tag{1}$$

We will first focus our attention on the existence of matrix bi-orthogonal polynomials with respect to the sesquilinear form $\langle\cdot,\cdot\rangle_W$ under some assumptions about the matrix polynomial W. Once this is done, the next step will be to find an explicit representation of such bi-orthogonal polynomials in terms of the matrix orthogonal polynomials with respect to the matrix of measures $\mathrm{d}\mu(x)$. We start with what we call connection formulas in Proposition 18.

One of the main achievements of this paper is Theorem 2 where we extend the Christoffel formula to MOPRL with a perturbation given by an arbitrary degree monic matrix polynomial. For that aim we use the rich spectral theory available today for these type of polynomials, in particular tools like root polynomials and Jordan chains will be extremely useful, see [76, 88]. Following [27, 59, 60] some applications to the analysis of matrix orthogonal polynomials which are eigenfunctions of second-order linear differential operators and related to polynomial perturbations of diagonal matrix of measures $\mathrm{d}\mu(x)$ leading to sesquilinear forms as $\langle\cdot,\cdot\rangle_W$ will be considered.

Next, to have a better understanding of the singular leading coefficient case we concentrate on the study of some cases which generalize important examples given by Alberto Grünbaum and Antonio Durán in [45, 46], in relation again with second-order linear differential operators.

Finally, we see that these Christoffel transformations extend to more general scenarios in Integrable Systems Theory. In these cases we find the non-Abelian Toda hierarchy which is relevant in string theory. In general, we have lost the block Hankel condition, and we do not have anymore a matrix of measures but only a sesquilinear form. We show that Theorem 2 also holds in this general situation. At this point we must stress that for the non-Abelian Toda equation we can find Darboux transformations (or Christoffel transformations) in [90], see also [84], which contemplate only what it are called elementary transformations and their iteration. Evidently, their constructions do not cover by far what Theorem 2 does. There are many matrix polynomials that do not factor in terms of linear matrix polynomials and, therefore, they cannot be studied by means of the results in [84, 90]. We have been fortunate to have at our disposal the spectral theory of [76, 88] that at the moment of the publication of [90] was not so well known and under construction.

The layout of the paper is as follows. We continue this introduction with two subsections that give the necessary background material regarding the spectral theory of matrix polynomials and also of matrix orthogonal polynomials. Then, in Section 2, we give the connection formulas for bi-orthogonal polynomials and for the Christoffel—Darboux (CD) kernel, being this last result relevant to find the dual polynomials in the family of bi-orthogonal polynomials. We continue in Section 3 discussing the nonsingular leading coefficient case, that is the monic matrix polynomial perturbation. We find the Christoffel formula for matrix bi-orthogonal polynomials and, as an example, we consider the degree 1 monic matrix polynomial perturbations. We dedicate the rest of this section to discuss some examples. In Section 4 we start the exploration of the singular leading coefficient matrix polynomial perturbations and, despite we do not give a general theory, we have been able to successfully discuss some relevant examples. Finally, Section 5 is devoted to the study of the extension of the previous results to the non-Abelian 2D Toda lattice hierarchy.

1.3 On spectral theory of matrix polynomials

Here we give some background material regarding matrix polynomials. For further reading we refer the reader to [55]

Definition 1. Let $A_0, A_1 \dots, A_N \in \mathbb{R}^{p\times p}$ be square matrices with real entries. Then

$$W(x) = A_N x^N + A_{N-1}x^{N-1} + \cdots + A_1 x + A_0 \tag{2}$$

is said to be a matrix polynomial of degree N, $\deg W = N$. The matrix polynomial is said to be monic when $A_N = I_p$, where $I_p \in \mathbb{R}^{p\times p}$ denotes the identity matrix. The linear space of matrix polynomials with coefficients in $\mathbb{R}^{p\times p}$ will be denoted by $\mathbb{R}^{p\times p}[x]$. □

Definition 2. We say that a matrix polynomial W as in (2) is monic normalizable if $\det A_N \neq 0$ and say that $\tilde{W}(x) := A_N^{-1} W(x)$ is its monic normalization. □

Definition 3. The spectrum, or the set of eigenvalues, $\sigma(W)$ of a matrix polynomial W is the zero set of $\det W(x)$, that is,

$$\sigma(W) := Z(W) = \{a \in \mathbb{C} : \det W(a) = 0\}. \qquad \square$$

Proposition 1. A monic normalizable matrix polynomial $W(x)$, $\deg W = N$, has Np (counting multiplicities) eigenvalues or zeros; that is we can write

$$\det W(x) = \prod_{i=1}^{q} (x - x_i)^{\alpha_i}$$

with $Np = \alpha_1 + \cdots + \alpha_q$. $\square$

Remark 1. In contrast with the scalar case, there exist matrix polynomials which do not have a unique factorization in terms of degree 1 factors or even it could happen that the factorization does not exist. For example, the matrix polynomial

$$W(x) = I_2 x^2 - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} x$$

can be written as

$$W(x) = \left(I_2 x - \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \right) \left(I_2 x - \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right) \qquad \text{or} \qquad W(x) = \left(I_2 x - \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \right) I_2 x,$$

but the polynomial

$$W(x) = I_2 x^2 - \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}$$

cannot be factorized in terms of degree 1 matrix polynomials. $\square$

Definition 4.

(i) Two matrix polynomials $W_1, W_2 \in \mathbb{R}^{m\times m}[x]$ are said to be equivalent $W_1 \sim W_2$ if there exist two matrix polynomials $E, F \in \mathbb{R}^{m\times m}[x]$, with constant determinants (not depending on x), such that $W_1(x) = E(x)W_2(x)F(x)$.

(ii) A degree 1 matrix polynomial $I_{Np}x - A \in \mathbb{R}^{Np\times Np}$ is called a linearization of a monic matrix polynomial $W \in \mathbb{R}^{p\times p}[x]$ if

$$I_{Np}x - A \sim \begin{bmatrix} W(x) & 0 \\ 0 & I_{(N-1)p} \end{bmatrix}. \qquad \square$$

Definition 5. Given a matrix polynomial $W(x) = I_p x^N + A_{N-1}x^{N-1} + \cdots + A_0$ its companion matrix $C_1 \in \mathbb{R}^{Np\times Np}$ is

$$C_1 := \begin{bmatrix} 0 & I_p & 0 & \dots & 0 \\ 0 & 0 & I_p & \ddots & 0 \\ \vdots & \vdots & \ddots & \ddots & \\ 0 & 0 & 0 & & I_p \\ -A_0 & -A_1 & -A_2 & \dots & -A_{N-1} \end{bmatrix}. \qquad \square$$

The companion matrix plays an important role in the study of the spectral properties of a matrix polynomial $W(x)$, (see, e.g., [55, 75, 76]).

Proposition 2. Given a monic matrix polynomial $W(x) = I_p x^N + A_{N-1}x^{N-1} + \cdots + A_0$ its companion matrix C_1 provides a linearization

$$I_{Np}x - C_1 \sim \begin{bmatrix} W(x) & 0 \\ 0 & I_{(N-1)p} \end{bmatrix},$$

where

$$E(x) = \begin{bmatrix} B_{N-1}(x) & B_{N-2}(x) & B_{N-3}(x) & \dots & B_1(x) & B_0(x) \\ -I_p & 0 & 0 & \dots & 0 & 0 \\ 0 & -I_p & 0 & \dots & 0 & 0 \\ 0 & 0 & -I_p & & 0 & 0 \\ \vdots & & & \ddots & \vdots & \\ 0 & 0 & 0 & & -I_p & 0 \end{bmatrix},$$

$$F(x) = \begin{bmatrix} I_p & 0 & 0 & \dots & 0 & 0 \\ -I_p x & I_p & 0 & \dots & 0 & 0 \\ 0 & -I_p x & I_p & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & I_p & 0 \\ 0 & 0 & 0 & & -I_p x & I_p \end{bmatrix},$$

with $B_0(x) := I_p$, $B_{r+1}(x) = xB_r(x) + A_{N-r-1}$, for $r \in \{0, 1, \dots, N-2\}$. $\square$

From here one deduces the important

Proposition 3. The eigenvalues with multiplicities of a monic matrix polynomial coincide with those of its companion matrix. □

Proposition 4. Any nonsingular matrix polynomial $W(x) \in \mathbb{C}^{m\times m}[x]$, $\det W(x) \neq 0$, can be represented

$$W(x) = E(x_0)\,\mathrm{diag}((x-x_0)^{\kappa_1},\dots,(x-x_0)^{\kappa_m})F(x_0)$$

at $x = x_0 \in \mathbb{C}$, where $E(x_0)$ and $F(x_0)$ are nonsingular matrices and $\kappa_1 \leq \cdots \leq \kappa_m$ are nonnegative integers. Moreover, $\{\kappa_1,\dots,\kappa_m\}$ are uniquely determined by W and they are known as partial multiplicities of $W(x)$ at x_0. □

Definition 6.

(i) Given a monic matrix polynomial $W(x) \in \mathbb{R}^{p\times p}[x]$ with eigenvalues and multiplicities $\{x_k,\alpha_k\}_{k=1}^{q}$, $Np = \alpha_1+\cdots+\alpha_q$, a non-zero vector $v_{k,0} \in \mathbb{C}^p$ is said to be an eigenvector with eigenvalue x_k whenever $W(x_k)v_{k,0} = 0$, $v_{k,0} \in \operatorname{Ker} W(x_k) \neq \{0\}$.

(ii) A sequence of vectors $\{v_{i,0}, v_{i,1}\dots, v_{1,m_i-1}\}$ is said to be a Jordan chain of length m_i corresponding to $x_i \in \sigma(W)$ if $v_{0,i}$ is an eigenvector of $W(x_i)$ and

$$\sum_{r=0}^{j} \frac{1}{r!}\frac{\mathrm{d}^r W}{\mathrm{d}x^r}\Big|_{x=x_i} v_{i,j-r} = 0, \qquad j = \{0,\dots,m_i-1\}.$$

(iii) A root polynomial at an eigenvalue $x_0 \in \sigma(W)$ of $W(x)$ is a non-zero vector polynomial $v(x) \in \mathbb{C}^p[x]$ such that $W(x_0)v(x_0) = 0$. The multiplicity of this zero will be denoted by κ.

(iv) The maximal length of a Jordan chain corresponding to the eigenvalue x_k is called the multiplicity of the eigenvector $v_{0,k}$ and is denoted by $m(v_{0,k})$. □

The above definition generalizes the concept of Jordan chain for degree 1 matrix polynomials [1].

Proposition 5. The Taylor expansion of a root polynomial at a given eigenvalue $x_0 \in \sigma(W)$

$$v(x) = \sum_{j=0}^{q} v_j (x-x_0)^j$$

provides a Jordan chain $\{v_0, v_1,\dots,v_{\kappa-1}\}$. □

Proposition 6. Given an eigenvalue $x_0 \in \sigma(W)$, with multiplicity $s = \dim \operatorname{Ker} W(x_0)$, we can construct s root polynomials

$$v_i(x) = \sum_{j=0}^{\kappa_i - 1} v_{i,j}(x - x_0)^j, \qquad i \in \{1, \dots, s\},$$

where $v_i(x)$ is a root polynomial with the largest order κ_i among all root polynomials whose eigenvector does not belong to $\mathbb{C}\{v_{1,0}, \dots, v_{i-1,0}\}$. □

Definition 7. A canonical set of Jordan chains of the monic matrix polynomial $W(x)$ corresponding to the eigenvalue $x_0 \in \sigma(W)$ is, in terms of the root polynomials described in Proposition 6, the set of vectors

$$\{v_{1,0} \dots, v_{1,\kappa_1 - 1}, \dots, v_{s,0} \dots, v_{s,\kappa_r - 1}\}.$$ □

Proposition 7. For a monic matrix polynomial $W(x)$ the lengths $\{\kappa_1, \dots, \kappa_r\}$ of the Jordan chains in a canonical set of Jordan chains of $W(x)$ corresponding to the eigenvalue x_0, see Definition 7, are the non-zero partial multiplicities of $W(x)$ at $x = x_0$ described in Proposition 4. □

Definition 8. For each eigenvalue $x_i \in \sigma(W)$, with multiplicity α_i and $s_i = \dim \operatorname{Ker} W(x_i)$, we choose a canonical set of Jordan chains

$$\left\{v_{j,0}^{(i)}, \dots, v_{j,\kappa_j^{(i)} - 1}^{(i)}\right\}, \qquad j = 1, \dots, s_i,$$

and, consequently, with partial multiplicities satisfying $\sum_{j=1}^{s_i} \kappa_j^{(i)} = \alpha_i$. Thus, we can consider the following adapted root polynomials

$$v_j^{(i)}(x) = \sum_{r=0}^{\kappa_j^{(i)} - 1} v_{j,r}^{(i)}(x - x_i)^r. \tag{3}$$

□

Proposition 8. Given a monic matrix polynomial $W(x)$ the adapted root polynomials given in Definition 8 satisfy

$$\left.\frac{\mathrm{d}^r}{\mathrm{d}x^r}\right|_{x=x_i} \left(W(x) v_j^{(i)}(x)\right) = 0, \qquad r = 0, \dots, \kappa_j^{(i)} - 1, \qquad j = 1 \dots, s_i.$$ □

1.4 On orthogonal matrix polynomials

Recall that a sesquilinear form $\langle \cdot, \cdot \rangle$ on the linear space $\mathbb{R}^{p\times p}[x]$ is a map

$$\langle \cdot, \cdot \rangle : \mathbb{R}^{p\times p}[x] \times \mathbb{R}^{p\times p}[x] \longrightarrow \mathbb{R}^{p\times p},$$

such that for any triple $P, Q, R \in \mathbb{R}^{p\times p}[x]$ of matrix polynomials we have

(i) $\langle AP(x) + BQ(x), R(x) \rangle = A \langle P(x), R(x) \rangle + B \langle Q(x), R(x) \rangle$, $\forall A, B \in \mathbb{R}^{p\times p}$.

(ii) $\langle P(x), AQ(x) + BR(x) \rangle = \langle P(x), Q(x) \rangle A^\top + \langle P(x), R(x) \rangle B^\top$, $\forall A, B \in \mathbb{R}^{p\times p}$.

Here $A^\top$ denotes the transpose of A, an antiautomorphism of order 2 in the ring of matrices.

Definition 9. A sesquilinear form $\langle \cdot, \cdot \rangle$ is said to be nondegenerate if the leading principal sub-matrices of the corresponding Hankel matrix of moments $M := \left(\left\langle I_p x^i I, I_p x^j \right\rangle\right)_{i,j=0}^{\infty}$ are nonsingular, and nontrivial if $\langle \cdot, \cdot \rangle$ is a symmetric matrix sesquilinear form and $\langle P(x), P(x) \rangle$ is a positive definite matrix for all $P(x) \in \mathbb{R}^{p\times p}[x]$ with nonsingular leading coefficient. □

Given a sesquilinear form $\langle \cdot, \cdot \rangle$, two sequences of polynomials $\left\{P_n^{[1]}(x)\right\}_{n=0}^{\infty}$ and $\left\{P_n^{[2]}(x)\right\}_{n=0}^{\infty}$ are said to be bi-orthogonal with respect to $\langle \cdot, \cdot \rangle$ if

(i) $\deg(P_n^{[1]}) = \deg(P_n^{[2]}) = n$ for all $n \in \mathbb{Z}_+$,

(ii) $\left\langle P_n^{[1]}(x), P_m^{[2]}(x) \right\rangle = \delta_{n,m} H_n$ for all $n, m \in \mathbb{Z}_+$,

where $H_n \neq 0$ and $\delta_{n,m}$ is the Kronecker delta. Here, it is important to notice the order of the polynomials in the sesquilinear form; that is if $n \neq m$ then $\left\langle P_n^{[2]}(x), P_m^{[1]}(x) \right\rangle$ could be different from 0.

Remark 2. Recall that if A is a positive semidefinite (resp. definite) matrix, then there exists a unique positive semidefinite (resp. definite) matrix B such that $B^2 = A$. B is said to be the square root of A (see [61], Theorem 7.2.6) and we denote it by $B =: A^{1/2}$. As in the scalar case, when $\langle \cdot, \cdot \rangle$ is a sesquilinear form, we will write the matrix $\langle P, P \rangle^{1/2} := \|P\|$. □

Let

$$\mu = \begin{bmatrix} \mu_{1,1} & \dots & \mu_{1,p} \\ \vdots & & \vdots \\ \mu_{p,1} & \dots & \mu_{p,p} \end{bmatrix}$$

be a $p \times p$ matrix of Borel measures in $\mathbb{R}$. Given any pair of matrix polynomials $P(x), Q(x) \in \mathbb{R}^{p\times p}[x]$ we introduce the following sesquilinear form

$$\langle P(x), Q(x)\rangle = \int_{\mathbb{R}} P(x)\,\mathrm{d}\mu(x)(Q(x))^{\top}.$$

In terms of the moments of the matrix of measures μ, we define the matrix moments as

$$m_n := \int_{\mathbb{R}} x^n\,\mathrm{d}\mu(x) \in \mathbb{R}^{p\times p}$$

and arrange them in the semi-infinite block matrix and its kth truncation

$$M := \begin{bmatrix} m_0 & m_1 & m_2 & \cdots \\ m_1 & m_2 & m_3 & \cdots \\ m_2 & m_3 & m_4 & \cdots \\ \vdots & \vdots & \vdots & \cdots \end{bmatrix}, \qquad M_{[k]} := \begin{bmatrix} m_0 & \cdots & m_{k-1} \\ \vdots & & \vdots \\ m_{k-1} & \cdots & m_{2k-2} \end{bmatrix}.$$

Following [15] we can prove

Proposition 9. If $\det M_{[k]} \neq 0$ for $k \in \{1, 2, \ldots\}$, then there exists a unique Gaussian factorization of the moment matrix M given by

$$M = S_1^{-1} H (S_2)^{-\top},$$

where S_1, S_2 are lower unitriangular block matrices and H is a diagonal block matrix

$$S_i = \begin{bmatrix} I_p & 0 & 0 & \ldots \\ (S_i)_{1,0} & I_p & 0 & \cdots \\ (S_i)_{2,0} & (S_i)_{2,1} & I_p & \ddots \\ & & & \ddots \end{bmatrix}, \qquad H = \begin{bmatrix} H_0 & 0 & 0 & \cdots \\ 0 & H_1 & 0 & \ddots \\ 0 & 0 & H_2 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \qquad i = 1, 2,$$

with $(S_i)_{n,m}, H_n \in \mathbb{R}^{p\times p}$, $\forall n, m \in \{0, 1, \ldots\}$. If $\mu = \mu^{\top}$ then we are dealing with a Cholesky block factorization with $S_1 = S_2$ and $H = H^{\top}$. □

For $l \geq k$ we will also use the following bordered truncated moment matrix

$$M_{[k,l]}^{[1]} := \left[\begin{array}{ccc} m_0 & \cdots & m_{k-1} \\ \vdots & & \vdots \\ m_{k-2} & \cdots & m_{2k-3} \\ \hline m_l & \ldots & m_{l+k-1} \end{array}\right],$$

where we have replaced the last row of blocks, $\begin{bmatrix} m_{k-1} & \dots & m_{2k-2}\end{bmatrix}$, of the truncated moment matrix M_k by the row of blocks $\begin{bmatrix} m_l & \dots & m_{l+k-1}\end{bmatrix}$. We also need a similar matrix but replacing the last block column of M_k by a column of blocks as indicated

$$M^{[2]}_{[k,l]} := \left[\begin{array}{ccc|c} m_0 & \cdots & m_{k-2} & m_l \\ \vdots & & \vdots & \vdots \\ m_{k-1} & \cdots & m_{2k-3} & m_{k+l-1} \end{array}\right].$$

Using last quasi-determinants, see [52, 85], we find

Proposition 10. If the last quasi-determinants of the truncated moment matrices are nonsingular, that is

$$\det \Theta_*(M_{[k]}) \neq 0, \qquad k = 1, 2, \dots,$$

then the Gauss–Borel factorization can be performed and the following expressions

$$H_k = \Theta_*(M_{[k+1]}), \quad (S_1^{-1})_{k,l} = \Theta_*(M^{[1]}_{[k,l+1]})\Theta_*(M_{[l+1]})^{-1}, \quad (S_2^{-1})_{k,l} = \left(\Theta_*(M_{[l+1]})^{-1}\Theta_*(M^{[2]}_{[k,l+1]})\right)^\top,$$

hold. □

Definition 10. We define $\chi(x) := [I_p, I_p x, I_p x^2, \dots]^\top$ and the vectors of matrix polynomials $P^{[1]} = [P_0^{[1]}, P_1^{[1]}, \cdots]^\top$ and $P^{[2]} = [P_0^{[2]}, P_1^{[2]}, \dots]^\top$, where

$$P^{[1]} := S_1 \chi(x), \qquad P^{[2]} := S_2 \chi(x).$$ □

Proposition 11. The matrix polynomials $P_n^{[i]}(x)$ are monic and $\deg P_n^{[i]} = n$, $i = 1, 2$. □

Observe that the moment matrix can be expressed as

$$M = \int_{\mathbb{R}} \chi(x) \, \mathrm{d}\mu(x) (\chi(x))^\top.$$

Proposition 12. The families of monic matrix polynomials $\left\{P_n^{[1]}(x)\right\}_{n=0}^\infty$ and $\left\{P_n^{[2]}(x)\right\}_{n=0}^\infty$ are bi-orthogonal

$$\left\langle P_n^{[1]}(x), P_m^{[2]}(x)\right\rangle = \delta_{n,m} H_n, \qquad n, m \in \mathbb{Z}_+.$$

If $\mu = \mu^\top$ then $P_n^{[1]} = P_n^{[2]} =: P_n$ which in turn conform an orthogonal set of monic matrix polynomials

$$\langle P_n(x), P_m(x)\rangle = \delta_{n,m} H_n, \qquad n, m \in \mathbb{Z}_+,$$

and we can write $\|P_n\| = H_n^{1/2}$. These bi-orthogonal relations can be recasted as

$$\int_{\mathbb{R}} P_n^{[1]}(x)\,\mathrm{d}\mu(x) x^m = \int_{\mathbb{R}} x^m\,\mathrm{d}\mu(x)(P_n^{[2]}(x))^\top = H_n \delta_{n,m}, \qquad m \le n. \qquad \square$$

Proof. From the definition of the polynomials and the factorization problem, we get

$$\int_{\mathbb{R}} P^{[1]}(x)\,\mathrm{d}\mu(x)(P^{[2]}(x))^\top = \int_{\mathbb{R}} S_1 \chi(x)\,\mathrm{d}\mu(x)\chi(x)^T S_2{}^\top = S_1 M (S_2)^\top = H. \qquad \blacksquare$$

Remark 3. The matrix of measures $\mathrm{d}\mu(x)$ may undergo a similarity transformation, $\mathrm{d}\mu(x) \mapsto \mathrm{d}\mu_c(x)$, and be conjugate to $\mathrm{d}\mu(x) = B^{-1}\,\mathrm{d}\mu_c(x) B$, where $B \in \mathbb{R}^{p\times p}$ is a non-singular matrix. The relation between the orthogonal polynomials given by these two measures is easily seen to be

$$M = B^{-1} M_c B, \quad S_1 = B^{-1} S_{c,1} B, \quad H = B^{-1} H_c B, \quad (S_2)^\top = B^{-1}(S_{c,2})^\top B,$$
$$P_n^{[1]} = B^{-1} P_{c,n}^{[1]} B, \quad (P_n^{[2]})^\top = B^{-1}(P_{c,n}^{[2]})^\top B. \qquad \square$$

The shift matrix is the following semi-infinite block matrix

$$\Lambda := \begin{bmatrix} 0 & I_p & 0 & 0 & \dots \\ 0 & 0 & I_p & 0 & \dots \\ 0 & 0 & 0 & I_p & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots \end{bmatrix},$$

which satisfies the important spectral property

$$\Lambda \chi(x) = x \chi(x).$$

Proposition 13. The block Hankel symmetry of the moment matrix can be written as

$$\Lambda M = M \Lambda^\top. \qquad \square$$

Notice that this symmetry completely characterizes Hankel block matrices.

Proposition 14. We have the following last quasi-determinantal expressions

$$P_n^{[1]}(x) = \Theta_* \begin{bmatrix} m_0 & m_1 & \cdots & m_{n-1} & I_p \\ m_1 & m_2 & \cdots & m_n & I_p x \\ \vdots & \vdots & & \vdots & \vdots \\ m_{n-1} & m_n & \cdots & m_{2n-2} & I_p x^{n-1} \\ m_n & m_{n+1} & \cdots & m_{2n-1} & I_p x^n \end{bmatrix},$$

$$(P_n^{[2]}(x))^\top = \Theta_* \begin{bmatrix} m_0 & m_1 & \cdots & m_{n-1} & m_n \\ m_1 & m_2 & \cdots & m_n & m_{n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ m_{n-1} & m_n & \cdots & m_{2n-2} & m_{2n-1} \\ I_p & I_p x & \cdots & I_p x^{n-1} & I_p x^n \end{bmatrix}. \qquad \square$$

Given two sequences of matrix bi-orthonormal polynomials $\{P_k^{[1]}(x)\}_{k=0}^\infty$ and $\{P_k^{[2]}(x)\}_{k=0}^\infty$, with respect to $\langle\cdot,\cdot\rangle$, we define the nth Christoffel–Darboux kernel matrix polynomial

$$K_n(x,y) := \sum_{k=0}^{n} (P_k^{[2]}(y))^\top H_k^{-1} P_k^{[1]}(x). \tag{4}$$

Named after [32, 34] see also [33].

Proposition 15 (ABC theorem). An Aitken–Berg–Collar type formula

$$K_n(x,y) = \left[I_p, \dots, I_p x^n\right] (M_n)^{-1} \begin{bmatrix} I_p \\ \vdots \\ I_p y^n \end{bmatrix}$$

holds. $\square$

The scalar version was rediscovered and popularized by Berg [20], who found it in a paper of Collar [31], who attributes it to his teacher, Aitken. As we are inverting a Hankel block matrix we are dealing with a Hankel Bezoutian type expression. This is connected with the following:

Proposition 16 (Christoffel–Darboux formula). The Christoffel–Darboux kernel satisfies

$$(x-y)K_n(x,y) = (P_n^{[2]}(y))^\top (H_n)^{-1} P_{n+1}^{[1]}(x) - (P_{n+1}^{[2]}(y))^\top (H_n)^{-1} P_n^{[1]}(x). \qquad \square$$

2 Connection Formulas for Darboux Transformations of Christoffel Type

Given a monic matrix polynomial $W(x)$ of degree N we consider a new matrix of measures of the form

$$\mathrm{d}\mu(x) \mapsto \mathrm{d}\hat{\mu}(x) := W(x)\,\mathrm{d}\mu(x)$$

with the corresponding perturbed sesquilinear form

$$\langle P(x), Q(x)\rangle_W = \int P(x)W(x)\,\mathrm{d}\mu(x)(Q(x))^\top.$$

In the same way as above, the moment block matrix

$$\hat{M} := \int \chi(x)W(x)\,\mathrm{d}\mu(x)(\chi(x))^\top$$

is introduced. Let us assume that the perturbed moment matrix admits a Gaussian factorization

$$\hat{M} = \hat{S}_1^{-1}\hat{H}(\hat{S}_2)^{-\top},$$

where $\hat{S}_1, \hat{S}_2$ are lower unitriangular block matrices and $\hat{H}$ is a diagonal block matrix

$$\hat{S}_i = \begin{bmatrix} I_p & 0 & 0 & \dots \\ (\hat{S}_i)_{1,0} & I_p & 0 & \cdots \\ (\hat{S}_i)_{2,0} & (\hat{S}_i)_{2,1} & I_p & \ddots \\ & & & \ddots \end{bmatrix}, \qquad \hat{H} = \begin{bmatrix} \hat{H}_0 & 0 & 0 & \cdots \\ 0 & \hat{H}_1 & 0 & \ddots \\ 0 & 0 & \hat{H}_2 & \ddots \\ \vdots & \vdots & \ddots & \ddots \end{bmatrix}, \qquad i = 1, 2.$$

Then, we have the corresponding perturbed bi-orthogonal matrix polynomials

$$\hat{P}^{[i]}(x) = \hat{S}_i\chi(x), \qquad i = 1, 2,$$

with respect to the perturbed sesquilinear form $\langle\cdot,\cdot\rangle_W$.

Remark 4. The discussion for monic matrix polynomial perturbations and perturbations with a matrix polynomial with nonsingular leading coefficients are equivalent. Indeed, if instead of a monic matrix polynomial we have a matrix polynomial $\tilde{W}(x) = A_N x^N + \cdots + A_0$ with a nonsingular leading coefficient, $\det A_N \neq 0$, then we could factor out A_N, $\tilde{W}(x) = A_N W(x)$, where W is monic. The moment matrices are related by $\tilde{M} = A_N\hat{M}$ and, moreover, $\tilde{S}_1 = A_N\hat{S}_1(A_N)^{-1}$, $\tilde{H} = A_N\hat{H}$, $\tilde{S}_2 = \hat{S}_2$, and $\tilde{P}_k^{[1]}(x) = A_N P_k^{[1]}(x)(A_N)^{-1}$ as well as $\tilde{P}_k^{[2]}(x) = \hat{P}_k^{[2]}(x)$. □

2.1 Connection formulas for bi-orthogonal polynomials

Proposition 17. The moment matrix M and the W-perturbed moment matrix $\hat{M}$ satisfy

$$\hat{M} = W(\Lambda)M. \qquad \square$$

Definition 11. Let us introduce the following semi-infinite matrices

$$\omega^{[1]} := \hat{S}_1 W(\Lambda) S_1^{-1}, \qquad \omega^{[2]} := (S_2 \hat{S}_2^{-1})^\top,$$

which we call resolvent or connection matrices. $\square$

Proposition 18 (Connection formulas)**.** Perturbed and nonperturbed bi-orthogonal polynomials are subject to the following linear connection formulas

$$\omega^{[1]} P^{[1]}(x) = \hat{P}^{[1]}(x) W(x), \tag{5}$$

$$P^{[2]}(x) = (\omega^{[2]})^\top \hat{P}^{[2]}(x). \tag{6}$$

$\square$

Proposition 19. The following relations hold

$$\hat{H}\omega^{[2]} = \omega^{[1]} H. \qquad \square$$

Proof. From Proposition 17 and the LU factorization we get

$$\hat{S}_1^{-1} \hat{H} \hat{S}_2^{-\top} = W(\Lambda) S_1^{-1} H S_2^{-\top},$$

so that

$$\hat{H}(S_2 \hat{S}_2^{-1})^\top = \hat{S}_1 W(\Lambda) S_1^{-1} H$$

and the result follows. ∎

From this result we easily get that

Proposition 20. The resolvent matrix ω is a band upper triangular block matrix with all the block superdiagonals above the Nth one equal to zero.

$$\omega^{[1]} = \begin{bmatrix} \omega^{[1]}_{0,0} & \omega^{[1]}_{0,1} & \omega^{[1]}_{0,2} & \dots & \omega^{[1]}_{0,N-1} & I_p & 0 & 0 & \dots \\ 0 & \omega^{[1]}_{1,1} & \omega^{[1]}_{1,2} & \dots & \omega^{[1]}_{1,N-1} & \omega^{[1]}_{1,N} & I_p & 0 & \dots \\ 0 & 0 & \omega^{[1]}_{2,2} & \dots & \omega^{[1]}_{2,N-1} & \omega^{[1]}_{2,N} & \omega^{[1]}_{2,N+1} & I_p & \ddots \\ & \ddots & \ddots & \ddots & & & & \ddots & \ddots \end{bmatrix}$$

with

$$\omega^{[1]}_{k,k} = \hat{H}_k (H_k)^{-1}. \tag{7}$$

□

2.2 Connection formulas for the Christoffel–Darboux kernel

In order to relate the perturbed and nonperturbed kernel matrix polynomials, let us introduce the following truncation of the connection matrix ω.

Definition 12. We introduce the lower unitriangular matrix $\omega_{(n,N)} \in \mathbb{R}^{Np\times Np}$ with

$$\omega_{(n,N)} := \begin{cases} \begin{bmatrix} 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & & & \vdots & & \\ 0 & \dots & 0 & 0 & \dots & 0 \\ \omega^{[1]}_{0,n+1} & \dots & \omega_{0,N-1} & I_p & \ddots & 0 \\ \vdots & & & & \ddots & \vdots \\ \omega^{[1]}_{n,n+1} & \dots & & & \omega^{[1]}_{n,n+N-1} & I_p \end{bmatrix}, & n < N, \\ \begin{bmatrix} I_p & 0 & \dots & 0 & 0 \\ \omega^{[1]}_{n-N+2,n+1} & I_p & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & I_p & 0 \\ \omega^{[1]}_{n,n+1} & \dots & & \omega^{[1]}_{n,n+N-1} & I_p \end{bmatrix}, & n \geq N, \end{cases}$$

and the diagonal block matrix

$$\hat{H}_{n,N} = \mathrm{diag}(\hat{H}_{n-N+1}, \dots, \hat{H}_n).$$

□

Then, we can write the important

Theorem 1. The perturbed and original Christoffel–Darboux kernels are related by the following connection formulas

$$K_n(x,y)+\Big[(\hat P^{[2]}_{n-N+1}(x))^\top,\dots,(\hat P^{[2]}_{n}(x))^\top\Big](\hat H_{(n,N)})^{-1}\omega_{(n,N)}\begin{bmatrix}P^{[1]}_{n+1}(x)\\ \vdots\\ P^{[1]}_{n+N}(x)\end{bmatrix}=\hat K_n(x,y)W(x),$$

by convention $\hat P^{[2]}_j=0$ whenever $j<0$. □

Proof. Consider the truncation

$$(\omega^{[2]})_{[n+1]}:=\begin{bmatrix}I_p & \dots & \omega^{[2]}_{0,N-1} & \omega^{[2]}_{0,N} & 0 & 0 & \dots & 0\\ 0 & I_p & \dots & \omega^{[2]}_{1,N+1} & \omega^{[2]}_{1,N+2} & 0 & \dots & 0\\ \vdots & & \ddots & & & \ddots & \ddots & \\ \\ 0 & 0 & & 0 & I_p & \dots & & \omega^{[2]}_{n-N,n}\\ \vdots & \vdots & & & \ddots & \ddots & & \vdots\\ 0 & 0 & & & & & I_p & \omega^{[2]}_{n-1,n}\\ 0 & 0 & & & & & 0 & I_p\end{bmatrix}.$$

Recalling (6) in the form $(P^{[2]}(y))^\top=(\hat P^{[2]}(y))^\top\omega^{[2]}$ we see that $\big((\hat P^{[2]}(y))_{[n+1]}\big)^\top(\omega^{[2]})_{[n+1]}=\big((P^{[2]}(y))_{[n+1]}\big)^\top$ holds for the nth truncations of $P^{[2]}(y)$ and $\hat P^{[2]}(y)$. Therefore,

$$\begin{aligned}\big((\hat P^{[2]}(y))_{[n+1]}\big)^\top(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]}=&\big((P^{[2]}(y))_{[n+1]}\big)^\top(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]}\\ =&K_n(x,y).\end{aligned}$$

Now, we consider $(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]}$ and recall Proposition 19 in the form

$$(\omega^{[2]})_{[n+1]}\big(H_{[n+1]}\big)^{-1}=(\hat H_{[n+1]})^{-1}(\omega^{[1]})_{[n+1]}$$

which leads to

$$(\omega^{[2]})_{n+1}(H_{n+1})^{-1}(P^{[1]}(x))_{n+1}=(\hat H_{[n+1]})^{-1}(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]}.$$

Observe also that

$$(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]} = (\omega^{[1]}P^{[1]}(x))_{[n+1]} - \begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix}$$

with

$$V_N(x) = \omega_{(n,N)} \begin{bmatrix} P^{[1]}_{n+1}(x) \\ \vdots \\ P^{[1]}_{n+N}(x) \end{bmatrix}.$$

Hence, recalling (5) we get

$$(\omega^{[1]})_{[n+1]}(P^{[1]}(x))_{[n+1]} = (\hat{P}^{[1]}(x))_{[n+1]}W(x) - \begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix},$$

and consequently

$$\begin{aligned}
&\left((\hat{P}^{[2]}(y))_{[n+1]}\right)^{\top}(\omega^{[2]})_{[n+1]}(H_{[n+1]})^{-1}(P^{[1]}(x))_{[n+1]} \\
&= \left((\hat{P}^{[2]}(y))_{[n+1]}\right)^{\top}(\hat{H}_{[n+1]})^{-1}(\hat{P}^{[1]}(x))_{[n+1]}W(x) - \left((\hat{P}^{[2]}(y))_{[n+1]}\right)^{\top}(\hat{H}_{[n+1]})^{-1}\begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix} \\
&= \hat{K}_{n+1}(x,y)W(y) - \left((\hat{P}^{[2]}(y))_{[n+1]}\right)^{\top}(\hat{H}_{[n+1]})^{-1}\begin{bmatrix} 0_{(n-N)p\times p} \\ V_N(x) \end{bmatrix}.
\end{aligned}$$

■

3 Monic Matrix Polynomial Perturbations

In this section we study the case of perturbations by monic matrix polynomials $W(x)$, which is equivalent to matrix polynomials with nonsingular leading coefficients. Using the theory given in Section 1.3 we are able to extend the celebrated Christoffel formula to this context.

3.1 The Christoffel formula for matrix bi-orthogonal polynomials

We are now ready to show how the perturbed set of matrix bi-orthogonal polynomials $\{\hat{P}^{[1]}_n(x), \hat{P}^{[2]}_n(x)\}_{n=0}^{\infty}$ is related to the original set $\{P^{[1]}_n(x), P^{[2]}_n(x)\}_{n=0}^{\infty}$.

Proposition 21. Let $v_j^{(i)}(x)$ be the adapted root polynomials of the monic matrix polynomial $W(x)$ given in (3). Then, for each eigenvalue $x_i \in \sigma(W)$, $i \in \{1,\dots,q\}$,

$$\omega^{[1]}_{k,k}\frac{\mathrm{d}^r(P^{[1]}_k v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i}+\cdots+\omega^{[1]}_{k,k+N-1}\frac{\mathrm{d}^r(P^{[1]}_{k+N-1} v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i}=-\frac{\mathrm{d}^r(P^{[1]}_{k+N} v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i}, \tag{8}$$

for $r=0,\dots,\kappa_j^{(i)}-1$, and $j=1\dots,s_i$. □

Proof. From (5) we get

$$\omega^{[1]}_{k,k}P^{[1]}_k(x)+\cdots+\omega^{[1]}_{k,k+N-1}P^{[1]}_{k+N-1}(x)+P_{k+N}(x)=\hat{P}^{[1]}_k(x)W(x).$$

Now, according to Proposition 8 we have

$$\begin{aligned}\frac{\mathrm{d}^r}{\mathrm{d}x^r}\bigg|_{x=x_i}(\hat{P}^{[1]}_k W v_j^{(i)}) &= \sum_{s=0}^{r}\binom{r}{s}\frac{\mathrm{d}^{r-s}\hat{P}^{[1]}_k}{\mathrm{d}x^{r-s}}\bigg|_{x=x_i}\frac{\mathrm{d}^s(W v_j^{(i)})}{\mathrm{d}x^s}\bigg|_{x=x_i}\\ &=0,\end{aligned} \tag{9}$$

for $r=0,\dots,\kappa_j^{(i)}-1$ and $j=1\dots,s_i$. ■

Recall that $\sum_{j=1}^{s_i}\kappa_j^{(i)}=\alpha_i$ and that the sum of all multiplicities α_i is $Np=\sum_{i=1}^{q}\alpha_i$, $q=\#\sigma(W)$.

Definition 13.

(i) For each eigenvalue $x_i\in\sigma(W)$, in terms of the adapted root polynomials $v_j^{(i)}(x)$ of the monic matrix polynomial $W(x)$ given in (3), we introduce the vectors

$$\pi^{(r),(i)}_{j,k}:=\frac{\mathrm{d}^r(P^{[1]}_k v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i}\in\mathbb{C}^p$$

and arrange them in the partial row matrices $\pi^{(i)}_k\in\mathbb{C}^{p\times\alpha_i p}$ given by

$$\pi^{(i)}_k=\Big[\pi^{(0),(i)}_{1,k},\dots,\pi^{(\kappa_1^{(i)}-1),(i)}_{1,k},\dots,\pi^{(0),(i)}_{s_i,k},\dots,\pi^{(\kappa_{s_i}^{(i)}-1),(i)}_{s_i,k}\Big].$$

We collect all them as

$$\pi_k:=\big[\pi^{(1)}_k,\dots,\pi^{(q)}_k\big]\in\mathbb{C}^{p\times Np}.$$

Finally, we have

$$\Pi_{k,N} := \begin{bmatrix} \pi_k \\ \vdots \\ \pi_{k+N-1} \end{bmatrix} \in \mathbb{C}^{Np\times Np}.$$

(ii) In a similar manner, we define

$$\gamma_{j,n}^{(r),(i)}(y) := \left.\frac{\mathrm{d}^r\big(K_n(x,y)v_j^{(i)}(x)\big)}{\mathrm{d}x^r}\right|_{x=x_i} \in \mathbb{C}^p[y],$$
$$\gamma_n^{(i)}(y) := \Big[\gamma_{1,n}^{(0),(i)}(y),\dots,\gamma_{1,n}^{(\kappa_1^{(i)}-1),(i)}(y),\dots,\gamma_{s_i,n}^{(0),(i)}(y),\dots,\gamma_{s_i,n}^{(\kappa_{s_i}^{(i)}-1),(i)}(y)\Big] \in \mathbb{C}^{p\times\alpha_i}[y],$$

and, as above, collect all of them in

$$\gamma_n(y) = \big[\gamma_n^{(1)},\dots,\gamma_n^{(q)}\big] \in \mathbb{C}^{p\times Np}[y]. \qquad \square$$

Theorem 2 (The Christoffel formula for matrix bi-orthogonal polynomials)**.** The perturbed set of matrix bi-orthogonal polynomials $\{\hat{P}_k^{[1]}(x), \hat{P}_k^{[2]}(x)\}_{k=0}^{\infty}$, whenever $\det \Pi_{k,N} \neq 0$, can be written as the following last quasi-determinant

$$\hat{P}_k^{[1]}(x)W(x) = \Theta_* \left[\begin{array}{c|c} \Pi_{k,N} & \begin{matrix} P_k^{[1]}(x) \\ \vdots \\ P_{k+N-1}^{[1]}(x)\end{matrix} \\ \hline \pi_{k+N} & P_{k+N}^{[1]}(x) \end{array}\right], \tag{10}$$

$$(\hat{P}_k^{[2]}(x))^{\top}(\hat{H}_k)^{-1} = \Theta_* \left[\begin{array}{c|c} \Pi_{k+1,N} & \begin{matrix} 0 \\ \vdots \\ 0 \\ I_p \end{matrix} \\ \hline \gamma_k(x) & 0 \end{array}\right]. \tag{11}$$

Moreover, the new matrix squared norms are

$$\hat{H}_k = \Theta_* \left[\begin{array}{c|c} \Pi_{k,N} & \begin{matrix} H_k \\ 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \pi_{k+N} & 0 \end{array}\right]. \tag{12}$$

□

Proof. We assume that $P_j^{[2]} = 0$ whenever $j < 0$. To prove (10) notice that from (8) one deduces for the rows of the connection matrix that

$$[\omega_{k,k}^{[1]}, \dots, \omega_{k,k+n-1}^{[1]}] = -\pi_{k+N}(\Pi_{k,N})^{-1}. \tag{13}$$

Now, using (5) we get

$$[\omega_{k,k}^{[1]}, \dots, \omega_{k,k+N-1}^{[1]}] \begin{bmatrix} P_k^{[1]}(x) \\ \vdots \\ P_{k+N-1}^{[1]}(x) \end{bmatrix} + P_{k+N}^{[1]}(x) = \hat{P}_k^{[1]}(x) W(x).$$

and (10) follows immediately.

To deduce (11) for $k \geq N$ notice that Theorem 1 together with (9) yields

$$\gamma_{j,k}^{(r),(i)}(x) + \left[(\hat{P}_{k-N+1}^{[2]}(x))^\top, \dots, (\hat{P}_k^{[2]}(x))^\top\right](\hat{H}_{(k,N)})^{-1}\omega_{(k,N)} \begin{bmatrix} \pi_{j,k+1}^{(r),(i)} \\ \vdots \\ \pi_{j,k+N}^{(r),(i)} \end{bmatrix} = 0,$$

for $r = 0, \dots, \kappa_j^{(i)} - 1$ and $j = 1 \dots, s_i$. We arrange these equations in a matrix form to get

$$\gamma_k(x) + \left[(\hat{P}_{k-N+1}^{[2]}(x))^\top, \dots, (\hat{P}_k^{[2]}(x))^\top\right](\hat{H}_{(k,N)})^{-1}\omega_{(k,N)}\Pi_{k+1,N} = 0.$$

Therefore, assuming that $\det \Pi_{k+1,N} \neq 0$, we get

$$\left[(\hat{P}_{k-N+1}^{[2]}(x))^\top, \dots, (\hat{P}_k^{[2]}(x))^\top\right](\hat{H}_{(k,N)})^{-1}\omega_{(k,N)} = -\gamma_k(x)(\Pi_{k+1,N})^{-1},$$

which, in particular, gives

$$(\hat{P}_k^{[2]}(x))^\top(\hat{H}_k)^{-1} = -\gamma_k(x)(\Pi_{k+1,N})^{-1}\begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_p \end{bmatrix}.$$

Finally, (12) is a consequence of (7) and (13). ■

3.2 Degree 1 monic matrix polynomial perturbations

Let us illustrate the situation with the most simple case of a perturbation of degree 1 monic polynomial matrix

$$W(x) = I_p x - A.$$

The spectrum $\sigma(I_p x - A) = \sigma(A) = \{x_1, \dots, x_q\}$ is determined by the zeros of the characteristic polynomial of A

$$\det(I_p x - A) = (x - x_1)^{\alpha_1} \cdots (x - x_q)^{\alpha_q},$$

and for each eigenvalue let $s_i = \dim \operatorname{Ker}(I_p x_i - A)$ be the corresponding geometric multiplicity, and $\kappa_j^{(i)}, j = 1, \dots, s_i$, its partial multiplicities, so that $\alpha_i = \sum_{j=1}^{s_i} \kappa_j^{(i)}$ is the algebraic multiplicity (the order of the eigenvalue as a zero of the characteristic polynomial of A). After a similarity transformation of A we will get its canonical Jordan form. With no lack of generality we assume that A is already given in Jordan canonical form

$$\begin{bmatrix} \boxed{\begin{matrix} \mathcal{J}_{\kappa_1^{(1)}}(x_1) & & 0 \\ & \ddots & \\ 0 & & \mathcal{J}_{\kappa_{s_1}^{(1)}}(x_1) \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} \mathcal{J}_{\kappa_1^{(q)}}(x_q) & & 0 \\ & \ddots & \\ 0 & & \mathcal{J}_{\kappa_{s_q}^{(q)}}(x_q) \end{matrix}} \end{bmatrix},$$

where the Jordan blocks corresponding to each eigenvalue are given by

$$\mathcal{J}_{\kappa_j^{(i)}}(x_i) := \begin{bmatrix} x_i & 1 & 0 & & \\ 0 & x_i & 1 & & \\ & & \ddots & \ddots & \\ & & & x_i & 1 \\ & & & 0 & x_i \end{bmatrix} \in \mathbb{R}^{\kappa_j^{(i)} \times \kappa_j^{(i)}}, \qquad j = 1, \dots, s_i.$$

For each eigenvalue x_i we pick a basis $\{v_{0,j}^{(i)}\}_{j=1}^{s_i}$ of $\operatorname{Ker}(I_p x_i - A)$, then look for vectors $\{v_{r,j}^{(i)}\}_{r=1}^{\kappa_j^{(i)}-1}$ such that

$$(I_p x_i - A) v_{r,j}^{(i)} = -\, v_{r-1,j}^{(i)}, \quad r = 1, \dots, \kappa_j^{(i)} - 1,$$

so that $\{v_{r,j}^{(i)}\}_{r=0}^{\kappa_j^{(i)}}$ is a Jordan chain. As we are dealing with A in its canonical form the vectors $v_{r,j}^{(i)}$ can be identified with those of the canonical basis $\{e_i\}_{i=1}^{p}$ of $\mathbb{R}^p$ with $e_i = (0, \cdots 1_i, \cdots 0)_p^\top$. Indeed, we have

$$v_{r,j}^{(i)} = e_{\alpha_1 + \cdots + \alpha_{i-1} + \kappa_1^{(i)} + \cdots + \kappa_{j-1}^{(i)} + r + 1}.$$

Then, consider the polynomial vectors

$$v_j^{(i)}(x) = \sum_{r=0}^{\kappa_j^{(i)}-1} v_{r,j}^{(i)} (x - x_i)^r, \qquad j = 1, \dots, s_i, \qquad i = 1 \dots, q,$$

which satisfy

$$\left.\frac{\mathrm{d}^r}{\mathrm{d}x^r}\right|_{x=x_i} \Big((I_p x - A) v_j^{(i)} \Big) = 0, \quad r = 0, \dots, \kappa_j^{(i)} - 1, \quad j = 1, \dots, s_i, \quad i = 1, \dots, q.$$

Consequently,

$$\left.\frac{\mathrm{d}^r}{\mathrm{d}x^r}\right|_{x=x_i} \Big(\hat{P}^{[1]}(x) (I_p x - A) v_j^{(i)} \Big) = 0, \quad r = 0, \dots, \kappa_j^{(i)} - 1, \quad j = 1, \dots, s_i, \quad i = 1, \dots, q. \tag{14}$$

Now, let us notice the following simple fact

Proposition 22. For a given matrix $A \in \mathbb{R}^{p\times p}$ any matrix polynomial $P(x) = \sum_{k=0}^{n} P_k x^k \in \mathbb{R}^{p\times p}[x]$, $\deg P = n$, can be written as

$$P = \sum_{k=0}^{n} P_k^{(A)} (I_p x - A)^k.$$

In particular, we have $P_0^{(A)} = P(A) := \sum_{k=0}^{n} P_k A^k$. □

Proposition 23. We can write

$$\frac{1}{r!}\frac{\mathrm{d}^r (P v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i} = P(A) v_{r,j}^{(i)}, \qquad r = 0,\dots,\kappa_j^{(i)} - 1, \qquad j = 1,\dots,s_i, \qquad i = 1,\dots,q. \qquad \square$$

Proof. Observe that

$$\begin{aligned}\frac{\mathrm{d}^s P}{\mathrm{d}x^s}\bigg|_{x=x_i} v_{r-s,j}^{(i)} &= \sum_{l=r}^{k} \frac{l!}{(l-s)!} P_l^{(A)} (I_p x_i - A)^{l-s} v_{r-s,j}^{(i)} \\ &= \sum_{l=s}^{r} (-1)^{l-s} \frac{l!}{(l-s)!} P_l^{(A)} v_{r-l,j}^{(i)}, \qquad s = 1,\dots,r, \quad j = 1,\dots,s_i, \quad i = 1,\dots,q,\end{aligned}$$

and, consequently,

$$\begin{aligned}\frac{\mathrm{d}^r (P v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i} &= \sum_{s=0}^{r} \frac{r!}{s!} \sum_{l=s}^{r} (-1)^{l-s} \frac{l!}{(l-s)!} P_l^{(A)} v_{r-l,j}^{(i)} \\ &= r! \sum_{s=0}^{r} \sum_{l=s}^{r} (-1)^{l-s} \binom{l}{s} P_l^{(A)} v_{r-l,j}^{(i)} \\ &= r! \sum_{m=0}^{r} (-1)^{r-m} \left[\sum_{s=0}^{r-m} (-1)^s \binom{r-m}{s} \right] P_{r-m}^{(A)} v_{m,j}^{(i)} \\ &= r! P(A) v_{r,j}^{(i)},\end{aligned}$$

for $r = 0,\dots,\kappa_j^{(i)} - 1$, $j = 1,\dots,s_i$, and $i = 1,\dots,q$. ■

In the following, we'll use

$$K_n(y, A) := \sum_{m=0}^{n} P_m^{[2]}(y) H_m^{-1} P_m^{[1]}(A).$$

Proposition 24 (Degree 1 Christoffel formula). If $W(x) = I_p x - A$ and $\det P_n^{[1]}(A) \neq 0$ for $n \in \mathbb{Z}_+$, then the Christoffel formulas can be written as

$$\hat{P}_n^{[1]}(x)(I_p x - A) = \Theta_* \begin{bmatrix} P_n^{[1]}(A) & P_n^{[1]}(x) \\ P_{n+1}^{[1]}(A) & P_{n+1}^{[1]}(x) \end{bmatrix} \qquad (\hat{P}_n^{[2]}(y))^\top = \Theta_* \begin{bmatrix} P_{n+1}^{[1]}(A) & I \\ K_n(y,A) & 0 \end{bmatrix}$$
$$= P_{n+1}^{[1]}(x) - P_{n+1}^{[1]}(A)\big[P_n^{[1]}(A)\big]^{-1} P_n^{[1]}(x), \qquad = -K_n(y,A)\big[P_n^{[1]}(A)\big]^{-1}.$$

For the perturbed matrix squared norms we have

$$\begin{aligned} \hat{H}_n &= \Theta_* \begin{bmatrix} P^{[1]}(A) & H_n \\ P^{[n+1]}(A) & 0 \end{bmatrix} \\ &= -P_{n+1}^{[1]}(A)\big[P_n^{[1]}(A)\big]^{-1} H_n. \end{aligned}$$ □

Proof. According to (5) and Theorem 1

$$\omega_{n,n} P_n^{[1]}(x) + P_{n+1}^{[1]}(x) = \hat{P}_n^{[1]}(x)(I_p x - A), \quad K_n(x,y) + (\hat{P}_n^{[2]}(y))^\top P_{n+1}^{[1]}(x) = \hat{K}_n(x,y)(I_p x - A),$$

and using (14) we conclude

$$\begin{aligned} \omega_{n,n} \frac{\mathrm{d}^r (P_k^{[1]} v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i} + \frac{\mathrm{d}^r (P_{n+1}^{[1]} v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i} &= 0, \\ \frac{\mathrm{d}^r}{\mathrm{d}x^r}\bigg|_{x=x_i} \big((K_n(x,y) v_j^{(i)}(x)\big) + (\hat{P}_n^{[2]}(y))^\top \frac{\mathrm{d}^r (P_{n+1}^{[1]} v_j^{(i)})}{\mathrm{d}x^r}\bigg|_{x=x_i} &= 0, \end{aligned}$$

for $r = 0, \dots, \kappa_j^{(i)} - 1$, $j = 1, \dots, s_i$, and $i = 1, \dots, q$. From Proposition 23 and the fact that the ordered arrangement of the Jordan chain vectors $v_{r,j}^{(i)}$ gives the identity matrix I_p, we conclude

$$\begin{aligned} \omega_{n,n} P_k^{[1]}(A) + P_{n+1}^{[1]}(A) &= 0, \\ K_n(y,A) + (\hat{P}_n^{[2]}(y))^\top P_{n+1}^{[1]}(A) &= 0. \end{aligned}$$ ■

We now illustrate the Christoffel formula in the matrix orthogonal polynomial context with a simple case. We will study what is the effect of the Christoffel transformation on a positive Borel scalar measure $\mathrm{d}\mu(x)$, thus the perturbed matrix of measures is $(I_2 x - A)\,\mathrm{d}\mu(x)$. The perturbed monic orthogonal polynomials will be expressed, see Proposition 24,

$$\hat{P}_n^{[1]}(x) = \Big(I_2 p_{n+1}(x) - p_{n+1}(A)\big(p_n(A)\big)^{-1} p_n(x)\Big)(I_2 x - A)^{-1}, \tag{15}$$

$$\hat{P}_n^{[2]}(x) = -K_n(x, A)\big[P_n^{[1]}(A)\big]^{-1}, \tag{16}$$

where $p_n(x), K_n(x, y)$ are the scalar orthogonal polynomials and kernel polynomials associated with the original scalar positive Borel measure $\mathrm{d}\mu(x)$. Observe that despite starting with a set of orthogonal polynomials the perturbation generates a set of bi-orthogonal matrix polynomials. As the original measure is scalar, if we ensure that $A = A^\top$ is symmetric, we will get $\hat{P}_n(x) := P_n^{[1]}(x) = P_n^{[2]}(x)$, a new set of orthogonal matrix polynomials.

However, this will be a very trivial situation as we have

Proposition 25. The matrix orthogonal polynomials $\{\hat{P}_n(x)\}_{n=0}^\infty$ of the matrix of measures $(I_p x - A)\,\mathrm{d}\mu(x)$, where $A = A^\top$ is symmetric and $\mathrm{d}\mu$ is a positive Borel scalar measure, are similar to diagonal matrix orthogonal polynomials. □

Proof. Being the matrix A symmetric it will be always diagonalizable

$$A = QDQ^\top,$$

where Q is an orthogonal matrix $Q^\top = Q^{-1}$ and $D = \operatorname{diag}(x_1, \dots, x_p)$, is a diagonal matrix that collects the eigenvalues, not necessarily different, of A.

At the end, the new orthogonal polynomials will be

$$\hat{P}_n(x) = Q \begin{bmatrix} \frac{p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)} p_n(x)}{x - x_1} & 0 & \dots & 0 \\ 0 & \frac{p_{n+1}(x) - \frac{p_{n+1}(x_2)}{p_n(x_2)} p_n(x)}{x - x_2} & & 0 \\ \vdots & \ddots & \ddots & \vdots \\ & & & \\ 0 & 0 & & \frac{p_{n+1}(x) - \frac{p_{n+1}(x_p)}{p_n(x_p)} p_n(x)}{x - x_p} \end{bmatrix} Q^\top, \tag{17}$$

and the result is proven. ■

Thus, we have a diagonal bunch of elementary Darboux transformations of the original scalar orthogonal polynomials associated with the scalar measure $\mathrm{d}\mu$. This situation reappears even when the matrix is not symmetric but diagonalizable, as we will have again that the perturbed matrix orthogonal polynomials will be similar to a similar bunch of elementary Darboux transformations of the original scalar orthogonal polynomials.

$$\hat{P}_n^{[1]}(x) = Q \begin{bmatrix} \frac{p_{n+1}(x)-\frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} & 0 & \dots & 0 \\ 0 & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_2)}{p_n(x_2)}p_n(x)}{x-x_2} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & & \frac{p_{n+1}(x)-\frac{p_{n+1}(x_p)}{p_n(x_p)}p_n(x)}{x-x_p} \end{bmatrix} Q^{-1},$$

$$\hat{P}_n^{[2]}(x) = -Q \begin{bmatrix} \frac{K_n(x,x_1)}{p_n(x_1)} & 0 & \dots & 0 \\ 0 & \frac{K_n(x,x_2)}{p_n(x_2)} & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & & \frac{K_n(x,x_p)}{p_n(x_p)} \end{bmatrix} Q^{-1},$$

where Q does not need to be an orthogonal matrix.

If the matrix is not diagonalizable and has nontrivial Jordan blocks the situation is different. Let us explore this case when $p = 2$. Hence, we consider

$$W(x) = I_2 x - A,$$

with

$$A = M \begin{bmatrix} x_1 & 1 \\ 0 & x_1 \end{bmatrix} M^{-1}.$$

Now we have only one eigenvalue $\sigma(A) = \{x_1\}$, with a length 2 Jordan chain. Thus, there is a linear basis $\{v_1, v_2\} \subset \mathbb{R}^2$, $(A - x_1)v_1 = v_2$, $(A - x_1)v_2 = 0$, with $v_i = [v_{i,1} \quad v_{i,2}]^\top$, $i \in \{1, 2\}$, such that

$$M = \begin{bmatrix} v_{1,1} & v_{2,1} \\ v_{1,2} & v_{2,2} \end{bmatrix}.$$

Therefore,

$$p_{n+1}(A)\big(p_n(A)\big)^{-1} = M \begin{bmatrix} \frac{p_{n+1}(x_1)}{p_n(x_1)} & \frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2} \\ 0 & \frac{p_{n+1}(x_1)}{p_n(x_1)} \end{bmatrix} M^{-1}, \quad (I_2 x - A)^{-1} = M \begin{bmatrix} \frac{1}{x-x_1} & \frac{1}{(x-x_1)^2} \\ 0 & \frac{1}{x-x_1} \end{bmatrix} M^{-1},$$

where

$$W(p_n, p_{n+1})(x) = p_n(x)p'_{n+1} - p_{n+1}(x)p'_n(x)$$

is the Wronskian of two consecutive orthogonal polynomials.

Hence,

$$\hat{P}_n^{[1]}(x) = M \begin{bmatrix} \frac{p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} & \frac{p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x) - \frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2}(x-x_1)p_n(x)}{(x-x_1)^2} \\ 0 & \frac{p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x)}{x-x_1} \end{bmatrix} M^{-1}, \qquad (18)$$

$$\hat{P}_n^{[2]}(x) = -M \begin{bmatrix} \frac{K_n(x,x_1)}{p_n(x_1)} & -\frac{K_n(x,x_1)}{p_n(x_1)}p'_n(x_1) + \frac{1}{p_n(x_1)}\frac{\partial K_n(x,y)}{\partial y}\Big|_{y=x_1} \\ 0 & \frac{K_n(x,x_1)}{p_n(x_1)} \end{bmatrix} M^{-1}.$$

Observe that the polynomials

$$p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x), \qquad p_{n+1}(x) - \frac{p_{n+1}(x_1)}{p_n(x_1)}p_n(x) - \frac{W(p_n,p_{n+1})(x_1)}{(p_n(x_1))^2}(x-x_1)p_n(x)$$

have a zero at $x = x_1$ of order 1 and 2, respectively.

3.3 Examples

Example 1. In [59] the authors define the notion of a classical pair $\{w(x), D\}$, where $w(x)$ is a symmetric matrix-valued weight function and D is a second-order linear ordinary differential operator. In that paper a weight function is said to be classical if there exists a second-order linear ordinary differential operator D with matrix-valued polynomial coefficients $A_j(t)$, $\deg A_j \leq j$, of the form $D = A_2(x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + A_1(x)\frac{\mathrm{d}}{\mathrm{d}x} + A_0(x)$, such that $\langle DP, Q\rangle = \langle P, DQ\rangle$ for all matrix-valued polynomial functions $P(x)$ and $Q(x)$. Then, the pair $\{w, D\}$ is called a classical pair. In example 5.1 in [59] they present a family of Jacobi type classical pairs that contains, up to equivalence, all classical pairs of size 2 where $w(x) = x^\alpha(1-x)^\beta F(x)$, with $\alpha, \beta > -1$ and $0 < x < 1$, and such that $F(x)$ is of degree 1 and which are irreducible (in the sense that they are not equivalent to a direct sum of classical pairs of size 1). As we will show they are a direct sum of orthogonal polynomials of size 1 produced by two degree 1 Christoffel transformations of the scalar Jacobi polynomials with zeros at $x = 0, 1$. Thus, we are faced with two scalar monic Jacobi polynomials with each of the two parameters α and β shifted by one, respectively. In [96] an analysis of the reducibility of matrix weights is given. In particular, in Example 2.4 they consider the case $\alpha = \beta$. We must stress that, as was pointed in [59], reducibility of the matrix of weights $w(x)$ does not imply the reducibility of the classical pair $\{w(x), D\}$. Indeed, despite that the matrix of weights in this example is reducible the corresponding second-order linear differential operator is not.

The classical pair $\{w(x) = x^{\alpha}(1-x)^{\beta}F(x), D\}$ is given by

$$F(x) = F_1 x + F_0, \qquad F_1 = \begin{bmatrix} 0 & -a \\ -a & \frac{\beta-\alpha}{\alpha+1}a \end{bmatrix}, \qquad F_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \qquad \text{with } a = \frac{\alpha+\beta+2}{\alpha+1},$$

and a second-order matrix linear ordinary differential operator

$$D = x(1-x)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (X - xU)\frac{\mathrm{d}}{\mathrm{d}x} + V$$

where U, V, X are constant matrices depending on a parameter u. The sequence of orthogonal polynomials $\{\tilde{\hat{P}}_n^{(\alpha,\beta)}(x)\}_{n=0}^{\infty}$ associated with the classical pair is not given in [59]. Here, an explicit representation of $\tilde{\hat{P}}_n^{(\alpha,\beta)}(x)$ using Darboux transformations is deduced. In order to do it we consider that we have an initial alternative Jacobi measure $\mathrm{d}\mu(x) = x^{\alpha}(1-x)^{\beta}I_2$, with $\alpha, \beta > -1$ and $0 < x < 1$, which is perturbed by a degree 1 matrix polynomial F. This matrix polynomial is not monic but its leading coefficient is nonsingular and we can write

$$F(x) = F_1 W(x), \qquad W(x) = I_2 x - A, \qquad A := -F_1^{-1}F_0 = \frac{1}{a}\begin{bmatrix} \frac{\beta+1}{\alpha+1} & \frac{\beta+1}{\alpha+1} \\ 1 & 1 \end{bmatrix},$$

in terms of a degree 1 monic matrix polynomial $W(x)$. We have that A has two different eigenvalues $\sigma(A) = \{0, 1\}$ with corresponding eigenvectors $[1, -1]^{\top}$ and $\left[\frac{\beta+1}{\alpha+1}, 1\right]^{\top}$, the matrix $M := \begin{bmatrix} 1 & \frac{\beta+1}{\alpha+1} \\ -1 & 1 \end{bmatrix}$ allows to write $A = M \operatorname{diag}(0,1) M^{-1}$.

Remember, as was noticed in Remark 4, that from the monic orthogonal polynomials $\hat{P}_n^{(\alpha,\beta),[1]}(x)$ with respect to W, we get

$$\tilde{\hat{P}}_n^{(\alpha,\beta)}(x) = F_1 \hat{P}_n^{(\alpha,\beta),[1]}(x) F_1^{-1},$$

which are the monic orthogonal polynomials with respect to $w(x)$. As the matrix of measures $F(x)\,\mathrm{d}\mu(x)$ is symmetric, the bi-orthogonality collapses to orthogonality and the super-indexes $[1, 2]$ can be omitted. We will do the same with $\hat{P}_n^{(\alpha,\beta),[1]} = \hat{P}_n^{(\alpha,\beta)}$.

Following [28, 29] we conclude that the set of monic matrix orthogonal polynomials $\{P_n^{(\alpha,\beta)}(x)\}_{n=0}^{\infty}$ with respect to $\mathrm{d}\mu(x)$ is $P_n^{(\alpha,\beta)}(x) = p_n^{(\alpha,\beta)}(x) I_2$. (We must be careful at this point with the notation. This is not the scalar standard Jacobi polynomial usually denoted by the same symbol. In fact, if $\mathcal{P}^{(\alpha,\beta)}(z)$ denotes the standard Jacobi polynomials, then $p_n^{(\alpha,\beta)}(x) = \frac{2^n}{S_n(\alpha,\beta)}\mathcal{P}_n^{(\beta,\alpha)}(2x-1)$, notice the interchange between the parameters $\alpha \rightleftharpoons \beta$ and the linear transformation of the independent variable x.) With the alternative Jacobi polynomials $p_n^{(\alpha,\beta)}(x)$ given by

$$p_n^{(\alpha,\beta)}(x) = \frac{1}{S_n(\alpha,\beta)}\sum_{k=0}^{n}\binom{n+\alpha}{n-k}\binom{n+\beta}{k}x^{n-k}(x-1)^k, \quad \text{with} \quad S_n(\alpha,\beta) = \binom{2n+\beta+\alpha}{n}.$$

We easily see that

$$p_n^{(\alpha,\beta)}(0) = \frac{1}{S_n(\alpha,\beta)}\binom{n+\alpha}{n}, \qquad p_n^{(\alpha,\beta)}(1) = \frac{1}{S_n(\alpha,\beta)}\binom{n+\beta}{n},$$

so that

$$\frac{p_{n+1}^{(\alpha,\beta)}(0)}{p_n^{(\alpha,\beta)}(0)} = (n+1+\alpha)\rho_n^{(\alpha,\beta)}, \qquad \frac{p_{n+1}^{(\alpha,\beta)}(1)}{p_n^{(\alpha,\beta)}(1)} = (n+1+\beta)\rho_n^{(\alpha,\beta)},$$

where

$$\rho_n^{(\alpha,\beta)} := \frac{(n+\beta+\alpha+1)}{(2n+\beta+\alpha+2)(2n+\beta+\alpha+1)}.$$

From (17) we conclude

$$\hat{P}_n^{(\alpha,\beta)}(x) = M\begin{bmatrix} \frac{p_{n+1}^{(\alpha,\beta)}(x)-(n+1+\alpha)\rho_n^{(\alpha,\beta)}p_n^{(\alpha,\beta)}(x)}{x} & 0 \\ 0 & \frac{p_{n+1}^{(\alpha,\beta)}(x)-(n+1+\beta)\rho_n^{(\alpha,\beta)}p_n^{(\alpha,\beta)}(x)}{x-1}\end{bmatrix}M^{-1}.$$

However, we must notice that these two Darboux transformations correspond to the following transformations of the Jacobi measure

$$\begin{aligned} x^{\alpha}(x-1)^{\beta} &\mapsto x(x^{\alpha}(x-1)^{\beta}) = x^{\alpha+1}(x-1)^{\beta}, \\ x^{\alpha}(x-1)^{\beta} &\mapsto (x-1)(x^{\alpha}(x-1)^{\beta}) = x^{\alpha}(x-1)^{\beta+1}, \end{aligned}$$

that is, the transformations correspond to the shifts $\alpha \mapsto \alpha+1$ and $\beta \mapsto \beta+1$, respectively. Consequently,

$$\hat{P}_n^{(\alpha,\beta)}(x) = M\begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x)\end{bmatrix}M^{-1}.$$

With the matrix

$$\tilde{M} := \begin{bmatrix} -1 & 1 \\ \frac{1+\beta}{1+\alpha} & 1 \end{bmatrix}$$

we can write $F_1 M = -a\tilde{M}$. We finally get the monic matrix orthogonal polynomials

$$\tilde{\hat{P}}_n^{(\alpha,\beta)}(x) = \tilde{M} \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x) \end{bmatrix} \tilde{M}^{-1},$$

for the matrix of measures $\tilde{W}(x)\,\mathrm{d}\mu(x)$ in Example 5.1 of [59] which are explicitly expressed in terms of scalar Jacobi polynomials as follows:

$$\begin{aligned} \tilde{\hat{P}}_n^{(\alpha,\beta)}(x) &= \frac{1}{2+\alpha+\beta} \\ &\times \begin{bmatrix} (\alpha+1)p_n^{(\alpha+1,\beta)}(x) + (\beta+1)p_n^{(\alpha,\beta+1)}(x) & -(\alpha+1)(p_n^{(\alpha+1,\beta)}(x) - p_n^{(\alpha,\beta+1)}(x)) \\ -(\beta+1)(p_n^{(\alpha+1,\beta)}(x) - p_n^{(\alpha,\beta+1)}(x)) & (\beta+1)p_n^{(\alpha+1,\beta)}(x) + (\alpha+1)p_n^{(\alpha,\beta+1)}(x) \end{bmatrix}. \end{aligned}$$

To conclude with this example let us mention that in [27] it was found that these matrix orthogonal polynomials also obey a first-order ordinary differential equation. From our point of view, this is just a consequence of a remarkable fact regarding the Darboux transformations $p^{(\alpha+1,\beta)}(x), p^{(\alpha,\beta+1)}(x)$ of the original alternative Jacobi polynomials. Under the hypergeometric function description of the Jacobi polynomials one gets recurrences for the Jacobi polynomials. In particular, from the Gauss' contiguous relations one gets the first-order differential relations

$$\begin{aligned} \left(x\frac{\mathrm{d}}{\mathrm{d}x} + \alpha + 1\right) p_n^{(\alpha+1,\beta)}(x) &= (\alpha+1+n)p_n^{(\alpha,\beta+1)}(x), \\ \left((x-1)\frac{\mathrm{d}}{\mathrm{d}x} + \beta + 1\right) p_n^{(\alpha,\beta+1)}(x) &= (\beta+1+n)p_n^{(\alpha+1,\beta)}(x). \end{aligned}$$

This first-order linear ordinary differential system can be recasted as a matrix linear differential equation as follows

$$\begin{aligned} &\left(\begin{bmatrix} 0 & x-1 \\ x & 0 \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \beta+1 \\ \alpha+1 & a_2 \end{bmatrix}\right) \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x) \end{bmatrix} \\ &= \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x) \end{bmatrix} \begin{bmatrix} a_1 & n+\beta+1 \\ n+\alpha+1 & a_2 \end{bmatrix}, \end{aligned} \tag{19}$$

where $a_1, a_2 \in \mathbb{R}$. This equation is invariant under multiplication on the right- and on the left-hand sides by arbitrary diagonal matrices

$$\begin{aligned}
&\begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \left(\begin{bmatrix} 0 & x-1 \\ x & 0 \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \beta+1 \\ \alpha+1 & a_2 \end{bmatrix} \right) \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix} \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x) \end{bmatrix} \\
&\quad = \begin{bmatrix} p_n^{(\alpha+1,\beta)}(x) & 0 \\ 0 & p_n^{(\alpha,\beta+1)}(x) \end{bmatrix} \begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix} \begin{bmatrix} a_1 & n+\beta+1 \\ n+\alpha+1 & a_2 \end{bmatrix} \begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}.
\end{aligned}$$

After the similarity transformation, $A \mapsto \tilde{M}^{-\top} A \tilde{M}^{\top}$ we find out that the orthogonal polynomial $\tilde{\hat{P}}_n$ satisfies

$$\left(A_1(x) \frac{\mathrm{d}}{\mathrm{d}x} + A_0 \right) \left(\tilde{\hat{P}}_n \right)^{\top} = \left(\tilde{\hat{P}}_n \right)^{\top} \Lambda_n,$$

where

$$A_1(x) = \frac{\alpha+1}{\alpha+\beta+2} \left(\begin{bmatrix} -\delta & \delta_- \\ \delta_+ & \delta \end{bmatrix} x + d \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \right), \quad A_0 = \frac{\alpha+1}{\alpha+\beta+2} \begin{bmatrix} C_+ - \Delta & \frac{\beta+1}{\alpha+1}(C+\Delta_-) \\ C+\Delta_+ & C_- + \Delta \end{bmatrix},$$

$$\Lambda_n = \frac{\alpha+1}{\alpha+\beta+2} \begin{bmatrix} -\delta & \delta_- \\ \delta_+ & \delta \end{bmatrix} n + A_0,$$

with $d = l_1 r_2$ and

$$\begin{aligned}
\delta &= l_1 r_2 + \frac{\beta+1}{\alpha+1} l_2 r_1, & \delta_- &= -l_1 r_2 + \left(\frac{\beta+1}{\alpha+1} \right)^2 l_2 r_1, & \delta_+ &= l_1 r_2 - l_2 r_1, \\
\Delta &= (\beta+1)(l_1 r_2 + l_2 r_1), & \Delta_- &= -l_1 r_2 (\alpha+1) + l_2 r_1 (\beta+1), & \Delta_+ &= l_1 r_2 (\beta+1) - l_2 r_1 (\alpha+1), \\
C &= -l_1 r_1 a_1 + l_2 r_2 a_2, & C_- &= \frac{\beta+1}{\alpha+1} l_1 r_1 a_1 + l_2 r_2 a_2, & C_+ &= l_1 r_1 a_1 + \frac{\beta+1}{\alpha+1} l_2 r_2 a_2.
\end{aligned}$$

When we take $l_1 = l_2 = -1$, $r_1 = r_2 = 1$ and $a_1 = \beta + 1$ and $a_2 = \alpha + 1$ we get the first-order ordinary differential system in Section 4 of [27].

Remark 5. The discussion in this example, regarding the Jacobi polynomials $p^{(\alpha+1,\beta)}(x)$ and $p^{(\alpha,\beta+1)}(x)$ and the use of the Gauss' contiguous relations, connects with the results in [66], Remark 2.8., see also [64, 65]. □

□

Example 2. Here we analyze the Chebyshev example taken from [27] that gives an example of a family of matrix orthogonal polynomials which satisfy a first-order linear ordinary differential equation. In Section 3 of [27] we find a set of MOPRL related with the measure $\tilde{W}(x)\,\mathrm{d}\mu(x)$ where

$$\tilde{W}(x) := \begin{bmatrix} 1 & x \\ x & 1 \end{bmatrix}, \qquad \mathrm{d}\mu(x) = \frac{1}{\sqrt{1-x^2}}.$$

We have a nonsingular leading coefficient $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ so that

$$\tilde{W}(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} W(x), \qquad W(x) := I_2 x - \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}.$$

Following Remark 4 we shall analyze the Darboux transformations $\mathrm{d}\mu(x) \mapsto W(x)\,\mathrm{d}\mu(x)$.

Thus, using (17) we can write the perturbed monic matrix orthogonal polynomials as follows:

$$\hat{P}_n(x) = Q \begin{bmatrix} \frac{t_{n+1}(x) - \frac{t_{n+1}(-1)}{t_n(-1)} t_n(x)}{x+1} & 0 \\ 0 & \frac{t_{n+1}(x) - \frac{t_{n+1}(1)}{t_n(1)} t_n(x)}{x-1} \end{bmatrix} Q^\top, \qquad Q := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

where $\{t_n(x)\}_{n=0}^\infty$ are the monic Chebyshev polynomials of first kind, that is, $t_n(x) = 2^{-n+1} T_n(x)$ with T_n the first kind Chebyshev polynomial of degree n. Therefore, recalling that $T_n(\pm 1) = (\pm 1)^n$ we get

$$\frac{t_{n+1}(\mp 1)}{t_n(\mp 1)} = \mp \frac{1}{2}, \qquad t_{n+1}(x) - \frac{t_{n+1}(\mp 1)}{t_n(\mp 1)} t_n(x) = \frac{1}{2^n}(T_{n+1}(x) \pm T_n(x)).$$

Now, recalling the mutual recurrence relation satisfied by Chebyshev polynomials of the first and second kind, denoted these last ones by U_n,

$$T_{n+1}(x) = xT_n(x) - (1-x^2)U_{n-1}(x), \qquad T_n(x) = U_n(x) - xU_{n-1}(x),$$

which implies $T_{n+1}(x) = xU_n(x) - U_{n-1}(x)$, we deduce

$$T_{n+1}(x) \pm T_n(x) = (x \pm 1)(U_n(x) \mp U_{n-1}(x)).$$

Consequently,

$$\hat{P}_n(x) = \frac{1}{2^n} Q \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} Q^\top, \qquad Q := \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}. \qquad \square$$

The matrix orthogonal polynomials associated with the original measure $\tilde{W}(x)\,\mathrm{d}\mu(x)$ can be recovered from this by a similarity transformation with $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ so that

$$\begin{aligned} \tilde{\hat{P}}_n(x) &= \frac{1}{2^{n+1}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} \\ &= \frac{1}{2^n} \begin{bmatrix} U_n(x) & -U_{n-1}(x) \\ -U_{n-1}(x) & U_n(x) \end{bmatrix}. \end{aligned}$$

Remark 6. The polynomials $\{U_n \mp U_{n-1}\}_{n=0}^{\infty}$ with $U_{-1} = 0$, which are orthogonal with respect to the measures $\frac{x\pm 1}{\sqrt{1-x^2}}$, are the well-known Chebyshev polynomials of the third and fourth kind, respectively. □

Remark 7. The symmetric structure of the MOPRL can be encoded in the equation

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \tilde{\hat{P}}_n(x) = \tilde{\hat{P}}_n(x) \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$ □

As in the Jacobi case, the new two scalar families of orthogonal polynomials are related through

$$\left((x \mp 1) \frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{2} \right) (U_n(x) \pm U_{n-1}(x)) = \left(n + \frac{1}{2} \right) \big(U_n(x) \mp U_{n-1}(x) \big). \tag{20}$$

This follows from

$$\begin{aligned} (x \mp 1)(U_n'(x) \pm U_{n-1}'(x)) &= (x \mp 1) \frac{\mathrm{d}}{\mathrm{d}x} \left(\frac{T_{n+1}(x) \mp T_n(x)}{x \mp 1} \right) \\ &= T_{n+1}'(x) \mp T'(x) - \frac{T_{n+1}(x) \mp T_n(x)}{x \mp 1} \\ &= (n+1) U_n(x) - n U_{n-1}(x) - U_n(x) \mp U_{n-1}(x). \end{aligned}$$

Here we have used that $T_n' = nU_{n-1}$.

Differential equation (20) can be written in matrix form

$$\begin{aligned} &\left(\begin{bmatrix} 0 & x-1 \\ x+1 & 0 \end{bmatrix} \frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \frac{1}{2} \\ \frac{1}{2} & a_2 \end{bmatrix} \right) \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} \\ &\quad = \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix} \begin{bmatrix} a_1 & n + \frac{1}{2} \\ n + \frac{1}{2} & a_2 \end{bmatrix} \end{aligned} \tag{21}$$

where $a_1, a_2 \in \mathbb{R}$ are arbitrary constants. Notice also that this matrix equation is invariant under multiplication on the right- and on the left-hand sides by arbitrary diagonal matrices $L = \operatorname{diag}(l_1, l_2)$ and $R = \operatorname{diag}(r_1, r_2)$,

$$\begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix}\left(\begin{bmatrix} 0 & x-1 \\ x+1 & 0 \end{bmatrix}\frac{\mathrm{d}}{\mathrm{d}x} + \begin{bmatrix} a_1 & \frac{1}{2} \\ \frac{1}{2} & a_2 \end{bmatrix}\right)\begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}\begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix}$$
$$= \begin{bmatrix} U_n(x) - U_{n-1}(x) & 0 \\ 0 & U_n(x) + U_{n-1}(x) \end{bmatrix}\begin{bmatrix} l_1 & 0 \\ 0 & l_2 \end{bmatrix}\begin{bmatrix} a & n+\frac{1}{2} \\ n+\frac{1}{2} & b \end{bmatrix}\begin{bmatrix} r_1 & 0 \\ 0 & r_2 \end{bmatrix}.$$

After the similarity transformation we find out that the orthogonal polynomial $\tilde{\hat{P}}_n$ satisfies

$$\left(A_1(x)\frac{\mathrm{d}}{\mathrm{d}x} + A_0\right)\tilde{\hat{P}}_n(x) = \tilde{\hat{P}}_n(x)\Lambda_n,$$

where

$$A_1(x) = \begin{bmatrix} -\delta_+ x + \delta_- & \delta_- x - \delta_+ \\ -\delta_- x + \delta_+ & \delta_+ x - \delta_- \end{bmatrix}, \qquad \Lambda_n = \begin{bmatrix} A_+ - \delta_+\left(n+\frac{1}{2}\right) & A_- + \delta_-\left(n+\frac{1}{2}\right) \\ A_- - \delta_-\left(n+\frac{1}{2}\right) & A_+ + \delta_+\left(n+\frac{1}{2}\right) \end{bmatrix},$$

and $A_0 = \Lambda_{n=0}$ with

$$\delta_\pm = l_1 r_2 \pm l_2 r_1, \qquad\qquad A_\pm = l_1 r_1 a_1 \pm l_2 r_2 a_2.$$

Equations (3.1) and (3.2) of Section 3 of [27] can be recovered choosing $\left(\delta_+ = A_+ = 0, \delta_- = 1, A_- = -\frac{1}{2}\right)$ and $\left(\delta_- = A_- = 0, \delta_+ = -1, A_+ = \frac{1}{2}\right)$, respectively.

However, they are all equivalent to (21), another form of writing (20). It is in fact this last equation (20) a quite interesting one. Indeed, we have two families of Darboux-transformed orthogonal polynomials interconnected by two first-order differential equations. Moreover, we conclude

$$\left((x^2-1)\frac{\mathrm{d}^2}{\mathrm{d}x^2} + (2x \mp 1)\frac{\mathrm{d}}{\mathrm{d}x} + \frac{1}{4}\right)\big(U_n(x) \mp U_{n-1}(x)\big) = \left(n + \frac{1}{2}\right)^2 \big(U_n(x) \mp U_{n-1}(x)\big).$$

Example 3

Here we comment on the matrix Gegenbauer matrix-valued polynomials discussed in [66]. In this case the matrix of weights is a symmetric matrix, $W^{(\nu)} : [-1,1] \to \mathbb{R}^{N\times N}$, with matrix coefficients of the form

$$(W^{(\nu)}(x))_{i,j} := (1-x^2)^{\nu-1/2} \sum_{k=\max(0,i+j+1-N)} \alpha_k^{(\nu)}(i,j) C_{i+j-2k}^{(\nu)}(x), \qquad i \geq j,$$

where $\alpha_t^{(\nu)}(i,j)$ are some coefficients and $C_n^{(\nu)}(x)$ stands for the Gegenbauer or ultraspherical polynomials. Erik Koelink and Pablo Román kindly communicated us a nice feature of the matrix Christoffel transformation discussed in this paper when acting on this reach family of MOPRL: two families of Gegenbauer MOPRL associated with matrices of weights $W^{(\nu_1)}(x)$ and $W^{(\nu_2)}(x)$, such that $\nu_1 - \nu_2 = m \in \mathbb{Z}$, are linked by a matrix Christoffel transformation. Now, the perturbing polynomial $W(x)$ has $\deg W = 2m$. These examples are, in general, reducible to two irreducible blocks of sizes $N/2$, for N even, and $(N+1)/2$ and $(N-1)/2$ for odd N. For a discussion on the orthogonal and nonorthogonal reducibility of these examples see [66, 67].

4 Singular Leading Coefficient Matrix Polynomial Perturbations

After studying some examples that the literature provides us with, one may realize that, even thought it is generic to assume the perturbing matrix polynomial $W(x)$ to have a nonsingular leading coefficient, many examples do have a singular matrix as its leading coefficient. This situation is a special feature of the matrix case setting since in the scalar case, having a singular leading term would mean that this coefficient is just zero (affecting, of course, to the degree of the polynomial). For this reason, when dealing with this kind of matrix polynomials talking about their degree should make no sense. The effect that this fact has on our reasoning is that since $\deg[\det W(x)] \leq Np$ the information encoded in the zeros (and corresponding adapted polynomials) of $\det W(x)$ is no longer enough to make the matrices Π_{kN} of the needed size. Therefore, there will be no way to express the perturbed polynomials just in terms of the initial ones evaluated at the zeros of $\det W(x)$ and the method to find a Christoffel type formula fails. However, the information that seems to be missing in these cases may actually not be necessary due to the singular character of the leading coefficient of the perturbing polynomial. Let us consider the following example to take a glimpse of this scenario.

Let us pick up some scalar measure $\mathrm{d}\mu(x)$ and its associated monic orthogonal polynomials $\{p_k(x)\}_{k=0}^{\infty}$ together with their norms and three-term recurrence relation

$$h_k\delta_{kj} := \langle p_k, p_j\rangle, \qquad xp_k(x) = J_{k,k-1}p_{k-1}(x) + J_{k,k}p_k(x) + p_{k+1}(x), \qquad J_{k,k-1} = \frac{h_k}{h_{k-1}} > 0.$$

Now, consider its $2q \times 2q$ matrix diagonal extension $\in \mathbb{R}^{2q\times 2q}[x]$

$$P_k(x) := p_k(x)I_{2q}, \qquad\qquad H_k := h_kI_{2q}.$$

Our aim is to consider the following matrix polynomial (with singular leading coefficient)

$$W(x) := \begin{bmatrix} I_q + AA^\top x^2 & Ax \\ A^\top x & I_q \end{bmatrix}, \qquad A \in \mathbb{R}^{q\times q},$$

which is inspired by the $q=1$ case $\begin{bmatrix} 1+a^2x^2 & ax \\ ax & 1 \end{bmatrix}$ (see [45], and references therein) and study the corresponding perturbations of our initial scalar measure; that is $\mathrm{d}\hat{\mu}(x) := W(x)\,\mathrm{d}\mu(x)$ in order to obtain the transformed matrix orthogonal polynomials

$$\hat{P}_k(x) := \begin{bmatrix} (\hat{P}_k)_{1,1} & (\hat{P}_k)_{1,2} \\ (\hat{P}_k)_{2,1} & (\hat{P}_k)_{2,2} \end{bmatrix}, \qquad \hat{P}_k(x) \in \mathbb{R}^{2q\times 2q}[x], (\hat{P}_k)_{i,j} \in \mathbb{R}^{q\times q}[x],$$

$$\langle \hat{P}_k, \hat{P}_j\rangle_W := \delta_{k,j}\hat{H}_k = \delta_{k,j}\begin{bmatrix} (\hat{H}_k)_{1,1} & (\hat{H}_k)_{1,2} \\ (\hat{H}_k)_{2,1} & (\hat{H}_k)_{2,2} \end{bmatrix}, \qquad \hat{H}_k \in \mathbb{R}^{2q\times 2q}, (\hat{H}_k)_{i,j} \in \mathbb{R}^{q\times q}.$$

We have split them up this way for computational purposes. Notice that since $W(x) = W(x)^\top$ we have $\hat{M} = \hat{M}^\top := \hat{S}^{-1}\hat{H}[\hat{S}^{-1}]^\top$ and, therefore, $\hat{P}^{[1]} = \hat{P}^{[2]} := \hat{P}$ and $\hat{H}_k = (\hat{H}_k)^\top$.

Let us point out that

$$W(x) = \mathcal{W}(x)\mathcal{W}(x)^\top, \qquad \mathcal{W} := \begin{bmatrix} I_q & Ax \\ 0 & I_q \end{bmatrix}, \qquad \mathcal{W}^{-1} = \begin{bmatrix} I_q & -Ax \\ 0 & I_q \end{bmatrix}.$$

This implies that $\det W = \det \mathcal{W} = 1$ and, consequently, there is no spectral analysis to perform as there are noneigenvalues at all. Thus, the relation between the original and perturbed measures and moment matrices is

$$[\mathcal{W}(x)]^{-1}\,\mathrm{d}\hat{\mu} = \mathrm{d}\mu[\mathcal{W}(x)]^\top, \qquad [\mathcal{W}(\Lambda)]^{-1}\hat{M} = M[\mathcal{W}(\Lambda)]^\top.$$

Definition 14. We introduce the resolvent or connection matrix

$$\omega := \hat{S}\mathcal{W}(\Lambda)S^{-1}. \qquad \square$$

Proposition 26. The matrix ω is block tridiagonal, having only its diagonal and first superdiagonal and subdiagonal non-zero, and satisfies

$$\omega^{-1} = H[\omega]^\top \hat{H}^{-1}.$$

Moreover, we have the important connection formula

$$\omega P = \hat{P}\mathcal{W}(x). \qquad \square$$

Proof. The first relation is a consequence of the LU factorization of the moment matrices and the connection formula is a straightforward consequence of the definition of ω. ■

Proposition 27.

(i) The matrices

$$\rho_{k+1} := \left(I_q + J_{k+1,k}A^\top A\right)^{-1}, \qquad k \in \{-1,0,1,\dots\},$$

exist.

(ii) The perturbed MOPRL can be written in terms of the original orthogonal polynomials as follows

$$\begin{aligned}\hat{P}_{1,k+1}(x)\mathcal{W}(x) = &-\begin{bmatrix} J_{k+1,k}J_{k+1,k+1}A\rho_{k+1}A^\top & 0 \\ J_{k+1,k}\rho_{k+1}A^\top & 0\end{bmatrix} p_k(x)\\ &+\begin{bmatrix} I_q & J_{k+1,k+1}A\rho_{k+1} \\ 0 & \rho_{k+1}\end{bmatrix} p_{k+1}(x) + \begin{bmatrix} 0 & A \\ 0 & 0\end{bmatrix} p_{k+2}(x),\end{aligned}$$

for $k \in \{-1,0,1,\dots\}$. □

Proof. From the $(k+1)$th row of the connection formula we have that

$$\omega_{k+1,k}p_k(x) + \omega_{k+1,k+1}p_{k+1}(x) + \omega_{k+1,k+2}p_{k+2}(x) = \hat{P}_{k+1}(x)\mathcal{W}(x),$$

but from the Definition 14 and Proposition 26 one realizes that the previous expression reads

$$\begin{aligned}\hat{P}_{k+1}(x)\mathcal{W}(x) = \hat{H}_{k+1}&\begin{bmatrix} 0 & -A \\ 0 & 0\end{bmatrix}^\top h_k^{-1}p_k(x)\\ &+\begin{bmatrix} (\omega_{k+1,k+2})_{11} & (\omega_{k+1,k+2})_{12} \\ (\omega_{k+1,k+2})_{21} & (\omega_{k+1,k+2})_{22}\end{bmatrix} p_{k+1}(x) + \begin{bmatrix} 0 & A \\ 0 & 0\end{bmatrix} p_{k+2}(x).\end{aligned}$$

Now, taking into account that both $(\hat{P}_{k+1})_{11}$, $(\hat{P}_{k+1})_{22}$ are monic $q\times q$ polynomials of degree $k+1$, while $(\hat{P}_{k+1})_{12}$, $(\hat{P}_{k+1})_{21}$ are $q\times q$ polynomials of degree less than $k+1$, it is not hard to see (after using the three-term recurrence relation of the initial polynomials) that

$$\begin{aligned}(\omega_{k+1,k+2})_{11} &= I_q, & (\omega_{k+1,k+2})_{12} &= J_{k+1,k+1}A - h_k^{-1}(\hat{H}_{k+1})_{12}A^\top A,\\ (\omega_{k+1,k+2})_{21} &= 0, & (\omega_{k+1,k+2})_{22} &= I_q - h_k^{-1}(\hat{H}_{k+1})_{22}A^\top A.\end{aligned}$$

Hence, we have every coefficient that appears in the connection formula in terms of the still unknown norms of the orthogonal polynomials. Therefore, we just need to compute the second block column of the following integral

$$\begin{aligned}
&\int \left[\omega_{k+1,k}p_k(x)+\omega_{k+1,k+1}p_{k+1}(x)+\omega_{k+1,k+2}p_{k+2}(x)\right]\left[(\mathcal{W}(x))^\top x^{k+1}\right]\mathrm{d}\mu(x)\\
&\quad=\int \hat{P}_{k+1}(x)\mathcal{W}(x)\left[(\mathcal{W}(x))^\top x^{k+1}\right]\mathrm{d}\mu(x)\\
&\quad=\int \hat{P}_{k+1}(x)\,\mathrm{d}\hat{\mu}(x)x^{k+1}\\
&\quad=\hat{H}_{k+1},
\end{aligned}$$

which yields

$$(\hat{H}_{k+1})_{12}=J_{k+1,k+1}h_{k+1}A\left(I_q+J_{k+1,k}A^\top A\right)^{-1}, \qquad (\hat{H}_{k+1})_{22}=h_{k+1}\left(I_q+J_{k+1,k}A^\top A\right)^{-1}.$$

Therefore,

$$(\omega_{k+1,k+2})_{12}=J_{k+1,k+1}A\left(I_q+J_{k+1,k}A^\top A\right)^{-1}, \qquad (\omega_{k+1,k+2})_{22}=\left(I_q+J_{k+1,k}A^\top A\right)^{-1},$$

and the stated result follows. ■

For $q=1$ and the classical measures we have, see [47],

Corollary 1. Starting from the classical measures

(i) Hermite monic polynomials $\left\{\mathcal{H}_k(x)\right\}_{k=0}^\infty$ with norm $h_k=\pi^{\frac{1}{2}}\frac{k!}{2^k}$

$$J_{k+1,k}=\frac{k+1}{2}, \qquad J_{k+1,k+1}=0, \qquad \rho_{k+1}:=\frac{2}{2+a^2(k+1)}.$$

(ii) Laguerre monic polynomials $\left\{\mathcal{L}_k^\alpha(x)\right\}_{k=0}^\infty$ with norm $h_k=k!\Gamma(k+1+\alpha)$

$$\begin{aligned}
&J_{k+1,k}=(k+1)(k+\alpha+2), \quad J_{k+1,k+1}=(2k+\alpha+3),\\
&\rho_{k+1}:=\frac{1}{1+a^2(k+1)(k+1+\alpha)}.
\end{aligned}$$

and perturbing them by the matrix polynomial

$$W(x)=\mathcal{W}(x)\left(\mathcal{W}(x)\right)^\top, \qquad \mathcal{W}(x)=\begin{pmatrix}1 & ax\\ 0 & 1\end{pmatrix},$$

one obtains the perturbed orthogonal polynomials related to the classical orthogonal polynomials as follows:

$$\hat{\mathcal{H}}_{k+1}(x)\mathcal{W}(x) = \begin{pmatrix} 0 & 0 \\ \frac{\rho_{k+1}-1}{a} & 0 \end{pmatrix} \mathcal{H}_k(x) + \begin{pmatrix} 1 & 0 \\ 0 & \rho_{k+1} \end{pmatrix} \mathcal{H}_{k+1}(x) + \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \mathcal{H}_{k+2}(x),$$

$$\begin{aligned}\hat{\mathcal{L}}^{\alpha}_{k+1}(x)\mathcal{W}(x) = {}& \begin{pmatrix} (\rho_{k+1}-1)(2k+3+\alpha) & 0 \\ \frac{\rho_{k+1}-1}{a} & 0 \end{pmatrix} \mathcal{L}^{\alpha}_k(x) + \begin{pmatrix} 1 & a(2k+3+\alpha)\rho_{k+1} \\ 0 & \rho_{k+1} \end{pmatrix} \mathcal{L}^{\alpha}_{k+1}(x) \\ & + \begin{pmatrix} 0 & a \\ 0 & 0 \end{pmatrix} \mathcal{L}^{\alpha}_{k+2}(x).\end{aligned}$$

□

5 Extension to non-Abelian 2D Toda Hierarchies

Matrix orthogonal polynomials are connected with non-Abelian Toda lattices, see [11, 80].

5.1 Block Hankel moment matrices versus multi-component Toda hierarchies

Let us take $M = (m_{i,j})_{i,j=0}^{\infty}$, $m_{i,j} \in \mathbb{R}^{p\times p}$ a semi-infinite block matrix having a Gaussian factorization

$$M = (S_1)^{-1} H (S_2)^{-\top},$$

where S_1, S_2 are lower uni-triangular block matrices and H is block diagonal. Notice that conditions for this factorization to hold were given in Proposition 10.

Definition 15. We introduce some continuous flows or perturbations of this semi-infinite matrix. For that aim we first consider the diagonal matrices

$$t_{i,j} = \operatorname{diag}(t_{i,j,1}, \dots, t_{i,j,p}) \in \mathbb{R}^{p\times p}, \qquad i = 1, 2, \qquad j \in \mathbb{Z}_+,$$

the semi-infinite undressed wave matrices

$$V_i^{(0)} := \exp\Big(\sum_{j=0}^{\infty} t_{i,j}\Lambda^j\Big), \qquad i = 1, 2,$$

and the perturbed matrix $M(t)$, $t = (t_1, t_2)$, $t_i = \{t_{i,j,a}\}_{\substack{j\in\mathbb{Z}_+ \\ a\in\{1,\dots,p\}}}$

$$M(t) = V_1^{(0)}(t_1) M \big(V_2^{(0)}(t_2)\big)^{-\top}.$$

□

Observe that we do not require any Hankel form for the matrix M, modelled by $\Lambda M = M\Lambda^\top$. However, if $M(0)$ is a Hankel matrix $M(t)$ is also a Hankel matrix taking into account $\Lambda M(t) = M(t)\Lambda^\top$. Hence, if $\mathrm{d}\mu(x)$ is the initial matrix of measures, then the new matrix of measures $\mathrm{d}\mu(x,t)$ will be

$$\mathrm{d}\mu(x,t) = \exp\left(\sum_{j=0}^{\infty} t_{1,j}x^j\right)\mathrm{d}\mu(x)\exp\left(-\sum_{j=0}^{\infty} t_{2,j}x^j\right).$$

Here $M(t)$ will be the moment matrix of the matrix of measures. Moreover, if at any time the matrix of measures is block Hankel then it was and it will be a Hankel block matrix at any time. If we assume that we can perform the Gaussian factorization again, we can write

$$M(t) = (S_1(t))^{-1}H(t)(S_2(t))^{-\top}.$$

As we know, for the block Hankel case we are dealing with bi-orthogonal or orthogonal polynomials with respect to the associated matrix of measures. What happens in the general case? Following [3] and [71] we can understand the Gaussian factorization also as a bi-orthogonality condition. The semi-infinite vectors of polynomials will be

$$P^{[1]}(x) := S_1(t)\chi(x), \qquad\qquad P^{[2]}(x) := S_2(t)\chi(x),$$

and we consider a sesquilinear form in $\mathbb{R}^{p\times p}[x]$, see Section 1.4, that for any couple of matrix polynomials $P = \sum_{k=0}^{\deg P} p_k x^k$ and $Q(x) = \sum_{l=0}^{\deg Q} q_l x^l$ is defined by

$$\langle P(x), Q(x)\rangle = \sum_{\substack{k=1,\dots,\deg P\\ l=1,\dots,\deg Q}} p_k M_{k,l}(q_l)^\top,$$

where we can interpret

$$M_{k,l} = \left\langle x^k 1_p, x^l 1_p\right\rangle$$

as the Gram matrix of the sesquilinear form. With respect to this sesquilinear form we have the bi-orthogonality condition

$$\left\langle P^{[1]}_k(x), P^{[2]}_l(x)\right\rangle = H_k\delta_{k,l}.$$

For a block Hankel initial condition, this sesquilinear form is just the sequilinear product associated with a linear functional of a measure. In [10], different examples are discussed for the matrix orthogonal polynomials scenario. For example, multigraded Hankel matrices M fulfilling

$$\left(\sum_{a=1}^{p} \Lambda^{n_a} E_{a,a}\right) M = M \left(\sum_{a=1}^{p} \left(\Lambda^{\top}\right)^{m_a} E_{a,a}\right),$$

where $n_1, \dots, n_p, m_1, \dots, m_p$ are positive integers, can be realized as

$$M_{k,l} = \int x^k \, \mathrm{d}\mu^{(l)}(x)$$

in terms of matrices of measures $\mathrm{d}\mu^{(l)}(x)$ which satisfy the following periodicity condition

$$\mathrm{d}\mu_{a,b}^{(l+m_a)}(x) = x^{n_a} \, \mathrm{d}\mu_{a,b}^{(l)}(x). \tag{22}$$

Therefore, given the measures $\mathrm{d}\mu_{a,b}^{(0)}, \dots, \mathrm{d}\mu_{a,b}^{(m_b-1)}$ we can recover all the others from (22). In this case, we have generalized orthogonality conditions like

$$\int P_k^{[1]}(x) \, \mathrm{d}\mu^{(l)}(x) = 0, \qquad l = 0, \dots, k-1.$$

Coming back to the Gaussian factorization, we consider the wave matrices

$$\begin{aligned} V_1(t) &:= S_1(t) V_1^{(0)}(t_1), \\ \tilde{V}_2(t) &:= \tilde{S}_2(t) (V_2^{(0)}(t_2))^{\top}, \end{aligned}$$

where $\tilde{S}_2(t) := H(t)(S_2(t))^{-\top}$.

Proposition 28. The wave matrices satisfy

$$\big(V_1(t)\big)^{-1} \tilde{V}_2(t) = M. \tag{23}$$

□

Proof. It is a consequence of the Gaussian factorization. ■

Given a semi-infinite matrix A we have unique splitting $A = A_+ + A_-$ where A_+ is an upper triangular block matrix while is A_- a strictly lower triangular block matrix.

Proposition 29. The following equations hold

$$\frac{\partial S_1}{\partial t_{1,j,a}}(S_1)^{-1} = -\Big(S_1 E_{a,a}\Lambda^j (S_1)^{-1}\Big)_-, \qquad \frac{\partial S_1}{\partial t_{2,j,a}}(S_1)^{-1} = \Big(\tilde{S}_2 E_{a,a}\big(\Lambda^\top\big)^j (\tilde{S}_2)^{-1}\Big)_-,$$
$$\frac{\partial \tilde{S}_2}{\partial t_{1,j,a}}(\tilde{S}_2)^{-1} = \Big(S_1 E_{a,a}\Lambda^j (S_1)^{-1}\Big)_+, \qquad \frac{\partial \tilde{S}_2}{\partial t_{2,j,a}}(\tilde{S}_2)^{-1} = -\Big(\tilde{S}_2 E_{a,a}\big(\Lambda^\top\big)^j (\tilde{S}_2)^{-1}\Big)_+. \qquad \square$$

Proof. Taking right derivatives of (23) yields

$$\frac{\partial V_1}{\partial t_{i,j,a}}(V_1)^{-1} = \frac{\partial \tilde{V}_2}{\partial t_{i,j,a}}(\tilde{V}_2)^{-1}, \qquad i \in \{1,2\}, \qquad j \in \mathbb{Z}_+,$$

where

$$\frac{\partial V_1}{\partial t_{1,j,a}}(V_1)^{-1} = \frac{\partial S_1}{\partial t_{1,j,a}}(S_1)^{-1} + S_1 E_{a,a}\Lambda^j (S_1)^{-1}, \quad \frac{\partial V_1}{\partial t_{2,j,a}}(V_1)^{-1} = \frac{\partial S_1}{\partial t_{2,j,a}}(S_1)^{-1},$$
$$\frac{\partial \tilde{V}_2}{\partial t_{1,j,a}}(\tilde{V}_2)^{-1} = \frac{\partial \tilde{S}_2}{\partial t_{1,j,a}}(\tilde{S}_2)^{-1}, \quad \frac{\partial \tilde{V}_2}{\partial t_{2,j,a}}(\tilde{V}_2)^{-1} = \frac{\partial \tilde{S}_2}{\partial t_{2,j,a}}(\tilde{S}_2)^{-1} + \tilde{S}_2 E_{a,a}\big(\Lambda^\top\big)^j (\tilde{S}_2)^{-1},$$

and the result follows immediately. ■

As a consequence, we derive

Proposition 30. The multicomponent 2D Toda lattice equations

$$\frac{\partial}{\partial t_{2,1,b}}\left(\frac{\partial H_k}{\partial t_{1,1,a}}(H_k)^{-1}\right) + E_{a,a}H_{k+1}E_{b,b}(H_k)^{-1} - H_k E_{b,b}(H_{k-1})^{-1}E_{a,a} = 0$$

hold. □

Proof. From Proposition 29 we get

$$\frac{\partial H_k}{\partial t_{1,1,a}}(H_k)^{-1} = \beta_k E_{a,a} - E_{a,a}\beta_{k+1}, \qquad \frac{\partial \beta_k}{\partial t_{2,1,b}} = H_k E_{b,b}(H_{k-1})^{-1},$$

where $\beta_k \in \mathbb{R}^{p\times p}$, $k = 1, 2, \dots$, are the first subdiagonal coefficients in S_1. ■

The multi-component Toda and KP hierarchies were introduced in [100]. In [70, 71] its relevance in integrable aspects of differential geometry was emphasized, and in [62] a representation approach was developed, while in [2, 11] it was used in relation with multiple orthogonality. A comprehensive approach to multi-component 2D Toda hierarchy with applications in dispersionless integrability or generalized orthogonal polynomials can be found in [10, 72, 73].

If we introduce the total flows given by the derivatives

$$\partial_{i,j} := \sum_{a=1}^{p} \frac{\partial}{\partial t_{i,j,a}},$$

we get the non-Abelian 2D Toda lattice

$$\partial_{2,1}\big(\partial_{1,1}(H_k)\cdot(H_k)^{-1}\big) + H_{k+1}(H_k)^{-1} - H_k(H_{k-1})^{-1} = 0.$$

The non-Abelian Toda lattice was introduced in the context of string theory by Polyakov, [86, 87], and then studied under the inverse spectral transform by Mikhailov [79] and Riemann surface theory by Krichever [69]. The Darboux transformations were considered in [90] and later in [84].

The non-Abelian 2D Toda lattice hierarchy is a reduction of the multicomponent hierarchy by taking the diagonal time matrices $t_{i,j} = \mathrm{diag}(t_{i,j,1},\dots,t_{i,j,p})$ proportional to the identity; that is

$$t_{i,j} \mapsto t_{i,j} I_p, \qquad\qquad t_{i,j} \in \mathbb{R}.$$

These equations are just the first members of an infinite set of nonlinear partial differential equations, an integrable hierarchy. Its elements are given by

Definition 16. The partial, Lax and Zakharov–Shabat matrices are given by

$$\begin{aligned} \Pi_{1,a} &:= S_1 E_{a,a}(S_1)^{-1}, & \Pi_{2,a} &:= \tilde{S}_2 E_{a,a}(\tilde{S}_2)^{-1},\\ L_1 &:= S_1 \Lambda (S_1)^{-1}, & L_2 &:= \tilde{S}_2 \Lambda^{\top}(\tilde{S}_2)^{-1},\\ B_{1,j,a} &:= \big(\Pi_{1,a}(L_1)^j\big)_+, & B_{2,j,a} &:= \big(\Pi_{2,a}(L_2)^j\big)_-. \end{aligned}$$

□

Proposition 31 (The integrable hierarchy). The wave matrices obey the evolutionary linear systems

$$\begin{aligned} \frac{\partial V_1}{\partial t_{1,j,a}} &= B_{1,j,a} V_1, & \frac{\partial V_1}{\partial t_{2,j,a}} &= B_{2,j,a} V_1,\\ \frac{\partial \tilde{V}_2}{\partial t_{1,j,a}} &= B_{1,j,a} \tilde{V}_2, & \frac{\partial \tilde{V}_2}{\partial t_{2,j,a}} &= B_{2,j,a} \tilde{V}_2, \end{aligned}$$

the partial and Lax matrices are subject to the following *Lax equations*

$$\frac{\partial \Pi_{i',a'}}{\partial t_{i,j,a}} = \Big[B_{i,j,a}, \Pi_{i',a'}\Big], \qquad\qquad \frac{\partial L_{i'}}{\partial t_{i,j,a}} = \Big[B_{i,j,a}, L_{i'}\Big],$$

and Zakharov–Sabat matrices fulfill the following *Zakharov–Shabat equations*

$$\frac{\partial B_{i',j',a'}}{\partial t_{i,j,a}} - \frac{\partial B_{i,j,a}}{\partial t_{i',j',a'}} + \left[B_{i,j,a}, B_{i',j',a'}\right] = 0. \qquad \square$$

Proof. Follows from Proposition 29. ∎

Given two semi-infinite block matrices A, B the notation $[A, B] = AB - BA$ stands for the usual commutator of matrices.

A crucial observation, regarding orthogonal polynomials, must be pointed out. When orthogonal polynomials are involved, and the matrices to factorize are block Hankel, equivalently $\Lambda M = M\Lambda^\top$, we get $L_1 = S_1 \Lambda S_1^{-1} = \tilde{S}_2 \Lambda^\top \tilde{S}_2^{-1} = L_2$. As the reader may have noticed the Lax matrices L_1 and L_2 are, by construction, lower and upper Hessenberg block matrices, respectively. However, when the Hankel property holds both Lax matrices are equal,

$$L_1 = L_2,$$

and, therefore, we are faced to a tridiagonal block matrix; that is a Jacobi block matrix. Moreover, this Hankel condition implies an invariance property under the flows introduced, as we have that $M(t) = V_1^{(0)}(t_1 - t_2)M$, that is there are only one type of flows. This condition also implies that for the total flows we have

$$\begin{aligned} (\partial_{1,j} + \partial_{2,j})V_1 &= V_1 \Lambda^j, & (\partial_{1,j} + \partial_{2,j})\tilde{V}_2 &= \tilde{V}_2(\Lambda^\top)^j, \\ (\partial_{1,j} + \partial_{2,j})L_1 &= 0, & (\partial_{1,j} + \partial_{2,j})L_2 &= 0. \end{aligned}$$

Therefore, in the block Hankel case we are dealing with the multicomponent 1D Toda hierarchy.

5.2 The Christoffel transformation for the non-Abelian 2D Toda hierarchy

The idea is to follow what we did in Section 2.1 and consider an initial condition $\hat{M}$ at $t = 0$, this is

$$\hat{M} = W(\Lambda)M$$

for a matrix polynomial $W(x) \in \mathbb{R}^{p\times p}[x]$. Observe that using the scalar times $t_{i,j} \in \mathbb{R}$ of the non-Abelian flows determined by

$$V_i^{(0)} := \exp\left(\sum_{j=0}^{\infty} t_{i,j}\Lambda^j\right), \qquad\qquad i = 1, 2,$$

the perturbed matrix is given by

$$\begin{aligned}\hat{M}(t) =& V_1^{(0)}(t_1)\hat{M}\big(V_2^{(0)}(t_2)\big)^{-\top}\\ =& W(\Lambda)M(t).\end{aligned}$$

Here we have used that $[W(\Lambda), V_1^{(0)}(t)] = 0$, $\forall t_{1,j} \in \mathbb{R}$. Let us stress that we could request only $t_{1,j}$ to be scalars and let $t_{2,j}$ to be diagonal matrices. Despite this is a more general situation, we prefer to show how the method works in this simpler scenario.

Assuming that the block Gauss factorization holds, we proceed as in Section 2.1 and introduce the resolvents

$$\omega^{[1]}(t) := \hat{S}_1(t)W(\Lambda)\big(S_1(t)\big)^{-1}, \qquad\qquad \omega^{[2]}(t) := (S_2(t)\big(\hat{S}_2(t)\big)^{-1})^{\top}.$$

From the LU factorization we get

$$\big(\hat{S}_1(t)\big)^{-1}\hat{H}(t)\big(\hat{S}_2(t)\big)^{-\top} = W(\Lambda)\big(S_1(t)\big)^{-1}H(t)\big(S_2(t)\big)^{-\top},$$

so that

$$\hat{H}(t)(S_2(t)\big(\hat{S}_2(t)\big)^{-1})^{\top} = \hat{S}_1(t)W(\Lambda)\big(S_1(t)\big)^{-1}H(t),$$

and, consequently,

$$\hat{H}(t)\omega^{[2]}(t) = \omega^{[1]}(t)H(t)$$

holds. Hence, as in the static case where the variable t does not appear, we have that this t-dependent resolvent matrix has a band block upper triangular structure

$$\omega^{[1]} = \begin{bmatrix} \omega^{[1]}_{0,0} & \omega^{[1]}_{0,1} & \omega^{[1]}_{0,2} & \dots & \omega^{[1]}_{0,N-1} & I_p & 0 & 0 & \dots \\ 0 & \omega^{[1]}_{1,1} & \omega^{[1]}_{1,2} & \dots & \omega^{[1]}_{1,N-1} & \omega^{[1]}_{1,N} & I_p & 0 & \dots \\ 0 & 0 & \omega^{[1]}_{2,2} & \dots & \omega^{[1]}_{2,N-1} & \omega^{[1]}_{2,N} & \omega^{[1]}_{2,N+1} & I_p & \ddots \\ & \ddots & \ddots & \ddots & & & & \ddots & \ddots \end{bmatrix}$$

with

$$\hat{H}_k(t) = \omega^{[1]}_{k,k}(t)H_k(t),$$

and the connection formulas described in Proposition 19 hold in this wider context.

Moreover, if $W(x)$ is a monic polynomial we can ensure that the Christoffel formula is also fulfilled for the non-Abelian 2D Toda and Theorem 2 remains valid also in this scenario. Formulas (10) and (12) hold directly and need no further explanation. However, (11) needs the following brief discussion. The Christoffel–Darboux kernel is defined exactly as we did in (4) but very probably there is no such a formula as the CD formula given in Proposition 16 is present in this scenario. However, as was shown in [12], there are cases, such as the multigraded reductions, where one has a generalized CD formula.

Funding

The Universidad Complutense de Madrid Program "Ayudas para Becas y Contratos Complutenses Predoctorales en España 2011" to Gerardo Ariznabarreta, the Spanish "Ministerio de Economía y Competitividad" research project [MTM2012-36732-C03-01], *Ortogonalidad y aproximación; teoría y aplicaciones* to Manuel Mañas and Francisco Marcellán, "Ministerio de Economía y Competitividad" research project [MTM2015-65888-C4-2], *Ortogonalidad, teoría de la aproximación y aplicaciones en física matemática* to Juan Carlos García-Ardila and Francisco Marcellán and "Ministerio de Economía y Competitividad" research project [MTM2015-65888-C4-3], *Ortogonalidad, teoría de la aproximación y aplicaciones en física matemática* to Carlos Alvárez-Fernández, Gerardo Ariznabarreta and Manuel Mañas

References

[1] Abadir, K. and J. Magnus. *Matrix Algebra,* Cambridge, Cambridge University Press, 2005.

[2] Adler, M. and P. van Moerbeke "Group factorization, moment matrices and Toda lattices." *International Mathematics Research Notices* 12 (1997): 556–72.

[3] Adler, M. and P. van Moerbeke "Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems." *Communications in Mathematical Physics* 207 (1999): 589–620.

[4] Adler, M. and P. van Moerbeke "Vertex operator solutions to the discrete KP hierarchy." *Communications in Mathematical Physics* 203 (1999): 185–210.

[5] Adler, M. and P. van Moerbekee "The spectrum of coupled random matrices." *Annals of Mathematics* 149 (1999): 921–76.

[6] Adler, M. and P. van Moerbeke "Hermitian, symmetric and symplectic random ensembles: PDEs for the distribution of the spectrum." *Annals of Mathematics* 153 (2001): 149–89.

[7] Adler, M. and P. van Moerbeke "Darboux transforms on band matrices, weights and associated polynomials." *International Mathematics Research Notices* 18 (2001): 935–84.

[8] Adler, M., P. van Moerbeke, and P. Vanhaecke "Moment matrices and multi-component KP, with applications to random matrix theory." *Communications in Mathematical Physics* 286 (2009): 1–38.

[9] Akhiezer, N. I. *Classical Moment Problem and Some Related Questions in Analysis.* New York, Translated by N. Kemmer Hafner Publishing Co. 1965.

[10] Álvarez-Fernández, C., M. Mañas, and U. Fidalgo Prieto "The multicomponent 2D Toda hierarchy: generalized matrix orthogonal polynomials, multiple orthogonal polynomials and Riemann–Hilbert problems." *Inverse Problems* 26 (2010): 055009 (15 pp).

[11] Álvarez-Fernández, C., U. Fidalgo Prieto, and M. Mañas "Multiple orthogonal polynomials of mixed type: Gauss-Borel factorization and the multi-component 2D Toda hierarchy." *Advances in Mathematics* 227 (2011): 1451–525.

[12] Álvarez-Fernández, C. and M. Mañas "On the Christoffel–Darboux formula for generalized matrix orthogonal polynomials." *Journal of Mathematical Analysis and Applications* 418 (2014): 238–47.

[13] Álvarez-Nodarse, R., A. J. Durán, and A. M. de los Ríos "Orthogonal matrix polynomials satisfying second-order difference equation." *Journal of Approximation Theory* 169 (2013): 40–55.

[14] Aptekarev, A. I. and E. M. Nikishin "The scattering problem for a discrete Sturm-Liouville operator." *Matematicheskii Sbornik* 121 (163) 327–58 (1983): Mathematics of the USSR. Sbornik 49 (1984): 325–55.

[15] Ariznabarreta, G. and M. Mañas "Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems." *Advances in Mathematics* 264 (2014): 396–463.

[16] Ariznabarreta, G. and M. Mañas "Multivariate orthogonal polynomials and integrable systems." arXiv:1409.0570.

[17] Ariznabarreta, G. and M. Mañas "Darboux transformations for multivariate orthogonal polynomials." arXiv:1503.04786.

[18] Ariznabarreta, G. and M. Mañas "Multivariate orthogonal Laurent polynomials and integrable systems." arXiv:1506.08708.

[19] Ariznabarreta, G. and M. Mañas "Linear spectral transformations for multivariate orthogonal polynomials and multispectral Toda hierarchies." arXiv:1511.09129.

[20] Berg, C. "Fibonacci numbers and orthogonal polynomials." *Journal of Computational and Applied Mathematics* 17 (2011): 75–88.

[21] Bergvelt, M. J. and A. P. E. ten Kroode "Tau-functions and zero-curvature equations of Toda-AKNS type." *Journal of Mathematical Physics* 29 (1988): 1308–20.

[22] Bergvelt, M. J. and A. P. E. ten Kroode "Partitions, Vertex Operators Constructions and Multi-Component KP Equations." *Pacific Journal of Mathematics* 171 (1995): 23–88.

[23] Brezinski, C. “Padé-Type Approximation and General Orthogonal Polynomials.” *International Series of Numerical Institute* 50: Basel-Boston, M: Birkhäuser, 1980.

[24] Bueno. M. I. and F. Marcellán “Darboux transformation and perturbation of linear functionals.” *Linear Algebra and its Applications* 384 (2004): 215–42.

[25] Bueno, M. I. and F. Marcellán “Polynomial perturbations of bilinear functionals and Hessenberg matrices.” *Linear Algebra and its Applications* 414 (2006): 64–83.

[26] Cantero, M. J., F. Marcellán, L. Moral, and L. Velázquez “Darboux transformations for CMV matrices.” arXiv:1503.05003.

[27] Castro, M. M. and F. A. Grünbaum “Orthogonal matrix polynomials satisfying first-order differential equations: a collection of instructive examples.” *Journal of Nonlinear Mathematical Physics* 4 (2005): 63–76.

[28] Chelyshkov, V. S. “Alternative Jacobi polynomials and orthogonal exponentials.” arXiv:1105.1838.

[29] Chihara, T. S. “emphAn Introduction to Orthogonal Polynomials.” Mathematics and its Applications Series, 13. Science Publishers, Gordon and Breach New York, 1978.

[30] Choque-Rivero, A. E. and L. E. Garza “Moment perturbation of matrix polynomials.” *Integral Transforms and Special Functions* 26 (2015): 177–91.

[31] Collar, A. R. “On the reciprocation of certain matrices.” *Proceedings of the Royal Society of Edinburgh Section A Mathematics* 59 (1939): 195–206.

[32] Christoffel, E. B. “Über die Gaussische Quadratur und eine Verallgemeinerung derselben.” *Journal für die reine und angewandte Mathematik* 55 (1858): 61–82.

[33] Damanik, D., A. Pushnitski, and B. Simon “The Analytic Theory of Matrix Orthogonal Polynomials.” *Surveys in Approximation Theory* 4 (2008): 1–85.

[34] Darboux, G. “Mémoire sur l’approximation des fonctions de très-grands nombres, et sur une classe étendue de développements en série.” *Journal des Mathématiques Pures et Appliquées* (3) 4 (1878): 5–56; 377–416.

[35] Darboux, G. “Sur une proposition relative aux équations linéaires.” *Comptes rendus hebdomadaires des séances de l’Académie des Sciences* 94 (1882): 1456–459.

[36] Date, E., M. Jimbo, M. Kashiwara, and T. Miwa “Operator approach to the Kadomtsev-Petviashvili equation. Transformation groups for soliton equations. III.” *Journal of the Physical Society of Japan* 50 (1981): 3806–12.

[37] Date, E., M. Jimbo, M. Kashiwara, and T. Miwa “Transformation groups for soliton equations. Euclidean Lie algebras and reduction of the KP hierarchy.” *Publications of the Research Institute for Mathematical Sciences* 18 (1982): 1077–10.

[38] Date, E., M. Jimbo, M. Kashiwara, and T. Miwa “Transformation groups for soliton equations” in *Nonlinear Integrable Systems-Classical Theory and Quantum Theory* M. Jimbo and T. Miwa (eds.). Singapore, World Scientific, 1983.

[39] Delgado, A. M., J. S. Geronimo, P. Iliev, and F. Marcellán “Two variable orthogonal polynomials and structured matrices.” *SIAM Journal on Matrix Analysis and Applications* 28 (2006): 118–47.

[40] Derevyagin, M., J. C. García-Ardila, and F. Marcellán "Multiple Geronimus transformations." *Linear Algebra and its Applications* 454 (2014): 158–83.

[41] Doliwa, A., P. M. Santini, and M. Mañas "Transformations of Quadrilateral Lattices." *Journal of Mathematical Physics* 41 (2000): 944–90.

[42] Durán, A. J. "Markov's theorem for orthogonal matrix polynomials." *Canadian Journal of Mathematics* 48 (1996): 1180–195.

[43] Durán, A. J. "On orthogonal polynomials with respect to a definite positive matrix of measures." *Canadian Journal of Mathematics* 47 (1995): 88–112.

[44] Durán, A. J. "Matrix inner product having a matrix symmetric second-order differential operator." *Rocky Mountain Journal of Mathematics* 27 (1997): 585–600.

[45] Durán, A. J. and F. A. Grünbaum "A survey on orthogonal matrix polynomials satisfying second-order differential equations." *Journal of Computational and Applied Mathematics* 178 (2005): 169–90.

[46] Durán, A. J. and F. A. Grünbaum "Structural formulas for orthogonal matrix polynomials satisfying second-order differential equations." *Constructive Approximation* 22 (2005): 255–71.

[47] Durán, A. J. and P. López-Rodríguez "Structural formulas for orthogonal matrix polynomials satisfying second-order differential equations II." *Constructive Approximation* 26 (2007): 29–47.

[48] Durán, A. J. and W. Van Assche "Orthogonal matrix polynomials and higher-order recurrence relations." *Linear Algebra and its Applications* 219 (1995): 261–80.

[49] Eisenhart, L. P. Transformations of Surfaces: Hamburg, Princeton University Press, 1923.

[50] Fuhrmann, P. A. "Orthogonal Matrix Polynomials and System Theory." *Rend. Sem. Mat. Univ. Politec. Torino* 1987, Special Issue, 68–124.

[51] Gautschi, W. "An algorithmic implementation of the generalized Christoffel theorem." in Numerical Integration, edited by G. Hämmerlin, *International series Numerical Mathematics* 57, 89–106. Basel: Birkhäuser, 1982.

[52] Gel'fand, I. M., S. Gel'fand, V. S. Retakh, and R. Wilson "Quasideterminants." *Advances in Mathematics* 193 (2005): 56–141.

[53] Geronimo, J. S. "Scattering theory and matrix orthogonal polynomials on the real line." *Circuits, Systems, and Signal Processing* 1 (1982): 471–95.

[54] Geronimus, J. "On polynomials orthogonal with regard to a given sequence of numbers." *Comm. Inst. Sci. Math. Mec. Univ.* Kharkoff [Zapiski Inst. Mat. Mech.] (4) 17 (1940): 3–18.

[55] Gohberg, I., P. Lancaster, and L. Rodman *Matrix Polynomials*, Computer Science and Applied Mathematics. [Harcourt Brace Jovanovich, Publishers], New York, Academic Press, Inc. 1982.

[56] Grünbaum, F. A. "The Darboux Process and a Noncommutative Bispectral Problem: Some Explorations and Challenges." Geometric Aspects of Analysis and Mechanics Progress in Mathematics 292 161–177. Progr. Math., 292. New York: Birkhäuser/Springer, 2011.

[57] Grünbaum, F. A. and L. Haine "Bispectral Darboux transformations: an extension of the Krall polynomials." *International Mathematics Research Notices* 8 (1997): 359–92.

[58] Grünbaum, F. A. and M. Yakimov "Discrete bispectral Darboux transformations from Jacobi operators." *Pacific Journal of Mathematics* 204 (2002): 395–431.

[59] Grünbaum, F. A., I. Pacharoni, and J. Tirao "Matrix valued orthogonal polynomials of the Jacobi type." *Indagationes Mathematicae-New Series* 14 (2003): 353–66.

[60] Grünbaum, F. A., I. Pacharoni, and J. Tirao "Matrix valued orthogonal polynomials of the Jacobi type: The role of group representation theory." *Ann. I. Fourier* 55 (2005): 2051–68.

[61] Horn, R. A. and C. R. Johnson *Matrix Analysis*, 2nd ed., Cambridge, Cambridge University Press, 2013.

[62] Kac, V. G. and J. W. van de Leur "The n-component KP hierarchy and representation theory." *Journal of Mathematical Physics* 44 (2003): 3245–293.

[63] Karlberg, L. and H. Wallin "Padé type approximants and orthogonal polynomials for Markov-Stieltjes functions." *Journal of Computational and Applied Mathematics* 32 (1990): 153–57.

[64] Koelink, E., M. van Pruijssen, and P. Román "Matrix valued orthogonal polynomials related to $(\mathrm{SU}(2)x\,\mathrm{SU}(2), \mathrm{diag})$." *International Mathematics Research Notices* no. 24 (2012): 5673–730.

[65] Koelink, E., M. van Pruijssen, and P. Román "Matrix valued orthogonal polynomials related to $(\mathrm{SU}(2)x\,\mathrm{SU}(2), \mathrm{diag})$ II." *Publications of the Research Institute for Mathematical Sciences* 49 (2013): 271–312.

[66] Koelink, E., A. M. de los Ríos, and P. Román Matrix-valued Gegenbauer polynomials. arXiv:1403.2938.

[67] Koelink, E. and P. Román "Orthogonal vs. non-orthogonal reducibility of matrix-valued measures." *SIGMA* 12 (2016): 008, 9 pages.

[68] Krein, M. G. "The fundamental propositions of the theory of representations of Hermitian operators with deficiency index (m, m)." (Russian) *Ukrainian Mathematical Journal* 1 (1949): 3–66.

[69] Krichever, I. M. "The periodic non-Abelian Toda chain and its two-dimensional generalization" (Appendix to the paper B A Dubrovin *Theta functions and non-linear equations*, *Russian Mathematical Surveys* + 36 (1981) 11–92).

[70] Mañas, M., L. Martínez Alonso, and E. Medina "Dressing methods for geometric nets: I. Conjugate nets." *Journal of Physics A: Mathematical & General* 33 (2000): 2871–894.

[71] Mañas, M., L. Martínez Alonso, and E. Medina "Dressing methods for geometric nets: II. Orthogonal and Egorov nets." *Journal of Physics A: Mathematical & General* 33 (2000): 7181–206.

[72] Mañas, M., L. Martínez Alonso, and C. Alvárez-Fernández "The multicomponent 2D Toda hierarchy: discrete flows and string equations." *Inverse Problems* 25 (2009): 0650077 (31 pp).

[73] Mañas, M. and L. Martínez Alonso "The multicomponent 2D Toda hierarchy: dispersionless limit." *Inverse Problems* 25 (2009): 115020 (22 pp).

[74] Marcellán, F. and G. Sansigre "On a class of matrix orthogonal polynomials on the real line." *Linear Algebra and its Applications* 181 (1993): 97–109.

[75] Markus, A. S. "Introduction to the spectral theory of polynomials operator pencil." Translated from the Russian by H. H. McFaden. Translation edited by Ben Silver. With an appendix by M. V. Keldysh. Translations of Mathematical Monographs, 71. Providence, RI, American Mathematical Society, 1988.

[76] Markus, A. S. and I. V. Mereuca "On the complete n-tuple of roots of the operator equation corresponding to a polynomial operator bundle." *Izvestiya Rossiiskoi Akademii Nauk. Seriya Matematicheskaya* 37 (1973): 1108–131.

[77] Matveev, V. B. and M. A. Salle "Differential-difference evolution equations. II: Darboux Transformation for the Toda lattice." *Letters in Mathematical Physics* 3 (1979): 425–29.

[78] Maroni, P. "Sur la suite de polynômes orthogonaux associée á la forme $u = \delta_c + \lambda(x-c)^{-1}$." *Periodica Mathematica Hungarica* 21 (1990): 223–48 (French).

[79] Mikhailov, A. V. "The reduction problem and the inverse scattering method." *Physica D* 3 (1981): 73–117.

[80] Miranian, L. "Matrix valued orthogonal polynomials on the real line: some extensions of the classical theory." *Journal of Physics A: Mathematical & General* 38 (2005): 5731–49.

[81] Miranian, L. "Matrix valued orthogonal polynomials in the unit circle: some extensions of the classical theory." *Canadian Mathematical Bulletin* 52 (2009): 95–104.

[82] Moutard, Th. F. "Sur la construction des équations de la forme $(1/z)\partial^2 z/\partial x \partial y = \lambda(x, y)$ qui admettenent une intégrale générale explicite." *Journal de l'École polytechnique - Mathématiques* 45 (1878): 1–11.

[83] Mulase, M. "Complete integrability of the Kadomtsev–Petviashvili equation." *Advances in Mathematics* 54 (1984): 57–66.

[84] Nimmo, J. J. C. and R. Willox "Darboux Transformations for the Two-Dimensional Toda System." *Proceedings of Royal Society* A453 (1997): 2497–525.

[85] Olver, P. J. "On Multivariate Interpolation." *Studies in Applied Mathematics* 116 (2006): 201–40.

[86] Polyakov, A. M. "Quantum geometry of bosonic strings." *Physics Letters B* 103 (1981): 207–10.

[87] Polyakov, A. M. "Quantum geometry of fermionic strings." *Physics Letters B* 103 (1981): 211–3.

[88] Rodman, L. *Orthogonal matrix polynomials,* in Orthogonal Polynomials (Columbus OH 1989), P. Nevai Editor, NATO Adv. Sci. Inst. Ser. C. Math. Phys. Sci. 294, 345–62. Dordrecht: Kluwer, 1990.

[89] Rosenberg, M. "The square integrability of matrix-valued functions with respect to a nonnegative Hermitian measure." *Duke Mathematical Journal* 31 (1964): 291–8.

[90] Salle, M. A. "Darboux transformations for non-Abelian and nonlocal equations of the Toda chain type." *Theoretical and Mathematical Physics* + 53 (1982): 1092–9.

[91] Sato, M. "Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds (Random Systems and Dynamical Systems)." *RIMS Kokyuroku* 439 (1981): 30–46.

[92] Sato, M. and Y. Sato "Soliton Equations as Dynamical Systems on Infinite-Dimensional Grassmann Manifold." Nonlinear Partial Differential Equations in Applied Science (Tokyo, 1982) Mathematical Studies 81, 259–71. Amsterdam, North-Holland, 1983.

[93] Sinap, A. and W. Van Assche "Polynomial interpolation and Gaussian quadrature for matrix-valued functions." *Linear Algebra and its Applications* 207 (1994): 71–114.

[94] Sinap, A. and W. Van Assche "Orthogonal matrix polynomials and applications." *Proceedings of the Sixth International Congress on Computational and Applied Mathematics (Leuven, 1994). Journal of Computational and Applied Mathematics* 66 (1996): 27–52.

[95] Szegő, G. "Orthogonal Polynomials." 4th ed., American Mathematical Society Colloquium Publication Series, vol 23. Providence RI: American Mathematical Society, 1975.

[96] Tirao, J. and I. Zurrián Reducibility of matrix weights. arXiv:1501.04059.

[97] Yoon, G. "Darboux transforms and orthogonal polynomials." *Bulletin of the Korean Mathematical Society* 39 (2002): 359–76.

[98] Ueno, K. and K. Takasaki "Toda lattice hierarchy. I." *Proceedings of the Japan Academy, Series A* 59 (1984): 167–70.

[99] Ueno, K. and K. Takasaki "Toda lattice hierarchy. II." *Proceedings of the Japan Academy, Series A* 59 (1984): 215–8.

[100] Ueno, K. and K. Takasaki "Toda lattice hierarchy." *Group Representations and Systems of Differential Equations* (Tokyo, 1982), Advances Studies in Pure Mathematics, 4, 1–95, Holland, Amsterdam, 1984.

[101] Uvarov, V. B. "The connection between systems of polynomials that are orthogonal with respect to different distribution functions." *USSR Computational Mathematics and Mathematical Physics* 9 (1969): 25–36.

[102] Zhedanov, A. "Rational spectral transformations and orthogonal polynomials." *Journal of Computational and Applied Mathematics* 85 (1997): 67–86.

2 Multivariate orthogonal polynomials and integrable systems

Advances in Mathematics 302 (2016) 628–739

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Multivariate orthogonal polynomials and integrable systems

Gerardo Ariznabarreta [1,2], Manuel Mañas [*,1]

Departamento de Física Teórica II (Métodos Matemáticos de la Física), Universidad Complutense de Madrid, Ciudad Universitaria, Plaza de Ciencias 1, 28040 Madrid, Spain

A R T I C L E   I N F O

Article history:
Received 26 September 2014
Accepted 25 June 2016
Available online xxxx
Communicated by Charles Fefferman

MSC:
15A23
33C45
37K10
37L60
42C05
46L55

Keywords:
Multivariate orthogonal polynomials
Borel–Gauss factorization
Quasi-determinants
Christoffel–Darboux kernels
Darboux transformations
Christoffel formula
Quasi-tau matrices
Kernel polynomials
Integrable hierarchies
Toda equations

A B S T R A C T

Multivariate orthogonal polynomials in D real dimensions are considered from the perspective of the Cholesky factorization of a moment matrix. The approach allows for the construction of corresponding multivariate orthogonal polynomials, associated second kind functions, Jacobi type matrices and associated three term relations and also Christoffel–Darboux formulae. The multivariate orthogonal polynomials, their second kind functions and the corresponding Christoffel–Darboux kernels are shown to be quasi-determinants—as well as Schur complements—of bordered truncations of the moment matrix; quasi-tau functions are introduced. It is proven that the second kind functions are multivariate Cauchy transforms of the multivariate orthogonal polynomials. Discrete and continuous deformations of the measure lead to Toda type integrable hierarchy, being the corresponding flows described through Lax and Zakharov–Shabat equations; bilinear equations are found. Varying size matrix nonlinear partial difference and differential equations of the 2D Toda lattice type are shown to be solved by matrix coefficients of the multivariate orthogonal polynomials. The discrete flows, which are shown to be connected with a Gauss–Borel factorization of the Jacobi type matrices and its quasi-determinants, lead to expressions for

* Corresponding author.
E-mail addresses: gariznab@ucm.es (G. Ariznabarreta), manuel.manas@ucm.es (M. Mañas).

[1] Thanks for financial support from the Spanish Ministerio de Economía y Competitividad research project [MTM2015-65888-C4-3-P], *Ortogonalidad, teoría de la aproximación y aplicaciones en física matemática.*
[2] Thanks for financial support from the Universidad Complutense de Madrid Program "Ayudas para Becas y Contratos Complutenses Predoctorales en España 2011".

http://dx.doi.org/10.1016/j.aim.2016.06.029

KP equations

the multivariate orthogonal polynomials and their second kind functions in terms of shifted quasi-tau matrices, which generalize to the multidimensional realm, those that relate the Baker and adjoint Baker functions to ratios of Miwa shifted τ-functions in the 1D scenario. In this context, the multivariate extension of the elementary Darboux transformation is given in terms of quasi-determinants of matrices built up by the evaluation, at a poised set of nodes lying in an appropriate hyperplane in $\mathbb{R}^D$, of the multivariate orthogonal polynomials. The multivariate Christoffel formula for the iteration of m elementary Darboux transformations is given as a quasi-determinant. It is shown, using congruences in the space of semi-infinite matrices, that the discrete and continuous flows are intimately connected and determine nonlinear partial difference–differential equations that involve only one site in the integrable lattice behaving as a Kadomstev–Petviashvili type system. Finally, a brief discussion of measures with a particular linear isometry invariance and some of its consequences for the corresponding multivariate polynomials is given. In particular, it is shown that the Toda times that preserve the invariance condition lay in a secant variety of the Veronese variety of the fixed point set of the linear isometry.

Contents

1. Introduction

This paper is devoted to the study of the interrelation between the theory of Multivariate Orthogonal Polynomials, or orthogonal polynomials on several variables, and the theory of Integrable Systems of Toda type. We perform this analysis with the aid of the Gauss–Borel factorization of the moment matrix, that in this case reduces to a Cholesky factorization. To understand better the situation we now proceed to give a brief description on the state of the art for multivariate orthogonal polynomials, then we recall some facts regarding Toda equations and integrable systems. As we use quasi-determinants in a number of places we have also included some comments regarding this subject. Finally, we describe the aims, results and the layout of the paper.

1.1. On multivariate orthogonal polynomials

Multivariate orthogonal polynomials have been a subject of study for many years, we refer the reader to the book by Charles F. Dunkl and Yuan Xu [44] where the authors of this paper enjoyed learning diverse aspects of multivariate orthogonality. The authors present in that book the general theory and emphasize the classical types of orthogonal polynomials whose weight functions are supported on standard domains such as the cube, the simplex, the sphere and the ball. It also focuses on those of Gaussian type, for which fairly explicit formulae exist. Another general source could be the lecture notes [124] which provide an introduction to orthogonal polynomials of several variables. It covers the basic theory but deals mostly with examples, paying special attention to those orthogonal polynomials associated with classical type weight functions supported on the standard domains, for which fairly explicit formulae exist.

The recurrence relation for orthogonal polynomials in several variables was studied by Xu in [119], while in [120] he linked multivariate orthogonal polynomials with a commutative family of self-adjoint operators and the spectral theorem was used to show the existence of a three term relation for the orthogonal polynomials. He discusses in [121] how the three term relation leads to the construction of multivariate orthogonal polynomials and cubature formulae. Xu considers in [125] polynomial subspaces that contain discrete multivariate orthogonal polynomials with respect to the bilinear form are identified and show that the discrete orthogonal polynomials still satisfy a three-term relation and that Favard's theorem holds. Explicit three term recurrence relations for the determination of multivariate orthogonal polynomials, which allow for the derivation of evaluation algorithms of finite series of these polynomials, were obtained [20]. Recursive three-term recurrence for the multivariate Jacobi polynomials on a simplex are explicitly given in [118]. In [102] several relations linking differences of bivariate discrete orthogonal polynomials and polynomials are given. We should also mention the work [41] where bivariate real valued polynomials orthogonal with respect to a positive linear functional are considered; interestingly the authors discuss orthogonal polynomials associated with positive definite block Hankel matrices whose entries are also Hankel and develop methods for constructing such matrices.

Multivariate Padé approximants cubature formulae were considered in [23]. The analysis of orthogonal polynomials and cubature formulae on the unit ball, the standard simplex, and the unit sphere [123] lead to conclude the strong connection of orthogonal structures and cubature formulae for these three regions. In [83] Tchebychev polynomials were obtained using symmetric and antisymmetric sums of exponentials and Gaussian cubatures were found, which exist very rarely in higher dimension. The paper [122] presents a systematic study of the common zeros of polynomials in several variables which are related to higher-dimensional quadrature. In [78] a description of polynomials orthogonal on the bicircle and polycircle and their relation to bounded analytic functions on the polydisk is given. Important in this work is a Christoffel–Darboux like formula which in the bivariate case can be related to stable polynomials, Bernstein–Szegő measures and gives a new proof of Ando theorem in operator theory.

Karlin and McGregor [76] and Milch [90] discussed interesting examples of multivariate Hahn and Krawtchouk polynomials related to growth of birth and death processes. A study of two-variable orthogonal polynomials associated with a moment functional satisfying the two-variable analogue of the Pearson differential equation and an extension of some of the usual characterizations of the classical orthogonal polynomials in one variable was found [50]. In [8] semiclassical orthogonal polynomials in two variables are defined as the orthogonal polynomials associated with a quasi-definite linear functional satisfying a matrix Pearson-type differential equation, semiclassical functionals are characterized by means of the analogue of the structure relation in one variable and nontrivial examples of semiclassical orthogonal polynomials in two variables where given. Iliev and Xu gave in [70] a characterization of all second order difference operators of several variables that have discrete orthogonal polynomials as eigenfunctions are given and under some mild assumptions, they give a complete solution of the problem.

In [51] the authors analyze a bilinear form obtained by adding a Dirac mass to a positive definite moment functional defined in the linear space of polynomials in several variables. A new proof of Gasper theorem on the positivity of sums of triple products on Jacobi polynomials was given in [30]; this theorem plays an important role in setting up a convolution structure for Jacobi polynomials, the correlation operator is an operator on the N-sphere looking for its eigenfunction expansion in various angular momentum sectors leads to Gasper's theorem and to the Koornwinder–Schwartz product formulae for the biangle which constitutes an extension of Gasper's theorem to the bivariate case. Xu discusses in [126] monomial orthogonal polynomials with respect to the weight function on the unit sphere as well as for the related weight functions on the unit ball and on the standard simplex getting explicit formulae for the L^2 norm and explicit expansions in terms of known orthonormal basis.

There are Maple libraries—MOPS—for Jack, Hermite, Laguerre, and Jacobi multivariate polynomials. These libraries also deal with eigenvalue statistics for the Hermite, Laguerre, and Jacobi ensembles of random matrix theory [43].

1.2. On the Toda equations

Sometimes the names given to equations or theorems do not correspond exactly to the original discoverers of the result. This is one of those cases.

The Toda equations can be traced back to the classical work *Leçons sur la Théorie Générale des Surfaces* published in 1915, by the French Mathematician Jean Gaston Darboux [37]; when he studies the Laplace method on reduction and invariance properties associated with the canonical hyperbolic equation $\Delta r = 0$ where Δ is a second order real hyperbolic operator. In the *Deuxième Partie. Livre IV. Chapitre II. La méthode de Laplace* if we go to number 336 we discover recursion (27) (page 30 of [37]) for the invariants h_k and h_{k-1}—of equations E_k in number 335—:

$$h_{k+1} + h_{k-1} = 2h_k - \frac{\partial^2 \log h_k}{\partial x \partial y},$$

that for the new dependent variable q_k given by

$$h_k = \mathrm{e}^{q_{k-1} - q_k}$$

reads as the 2D Toda equation

$$\frac{\partial^2 q_k}{\partial x \partial y} = \mathrm{e}^{q_k - q_{k+1}} - \mathrm{e}^{q_{k-1} - q_k},$$

that for the dimensional reduction $x = \pm y = t$ simplifies to the Toda equation

$$\frac{\partial^2 q_k}{\partial t^2} = \mathrm{e}^{q_k - q_{k+1}} - \mathrm{e}^{q_{k-1} - q_k}.$$

Then, more than half a century later the Japanese Physicist Morikazu Toda, introduced [112] a simple model, that he named as exponential lattice, for a one-dimensional crystal in solid state physics with a nearest neighbor interaction, with potential $\phi(r) = \frac{a}{b}\mathrm{e}^{-r} + ar + c$, $a, b > 0$, such that the particles are subject to

$$\begin{aligned}\frac{\mathrm{d}p_k}{\mathrm{d}t}(t) &= \mathrm{e}^{-(q_k(t)-q_{k-1}(t))} - \mathrm{e}^{-(q_{k+1}(t)-q_k(t))},\\ \frac{\mathrm{d}q_k}{\mathrm{d}t}(t) &= p_k(t),\end{aligned}$$

where q_k and p_k are the displacement of the k-th particle from its equilibrium position, and its momentum (here the mass is set equal to the unity). In [112] exact solutions where obtained in terms of the Jacobian elliptic functions, it was also shown that the system has N normal modes and the expansion due to vibration of the chain was discussed. Later on [113] relations between this nonlinear exponential lattice, the Boussinesq equation and the Korteweg–de Vries equation showed up and therefrom two-soliton solutions were given in each case for both the head-on and the overtaking collisions.

The Toda lattice is a completely integrable system *à la Liouville* as it was shown in 1974 first by Michel Henón [67] and then by Hermann Flaschka [53] in terms Flaschka's variables:

$$a_k(t) = \frac{1}{2}\mathrm{e}^{-\frac{q_{k+1}(t)-q_k(t)}{2}}, \qquad b_k(t) = -\frac{1}{2}p_k(t),$$

so that 1D Toda equations are written as follows

$$\dot{a}_k(t) = a_k(t)\big(b_{k+1}(t) - b_k(t)\big), \tag{1.2.1}$$

$$\dot{b}_k(t) = 2\big(a_k(t)^2 - a_{k-1}(t)^2\big). \tag{1.2.2}$$

These equations can be reformulated as the Lax equation $\dot{L}(t) = [P(t), L(t)]$; the *Lax pair*, L and P, are linear operators in the space $\ell^2(\mathbb{Z})$ of square summable sequences given by

$$\begin{aligned}(L(t)f)_k &= a_k(t)f_{k+1} + a_{k-1}(t)f_{k-1} + b_k(t)f_k,\\ L(t) &= \begin{pmatrix} b_0(t) & a_0(t) & 0 & 0 & \dots\\ a_0(t) & b_1(t) & a_1(t) & 0 & \dots\\ 0 & a_1(t) & b_2(t) & a_2(t) & \dots\\ & \ddots & \ddots & \ddots & \end{pmatrix},\end{aligned} \tag{1.2.3}$$

$$\begin{aligned}(P(t)f)_k &= a_k(t)f_{k+1} - a_{k-1}(t)f_{k-1},\\ P(t) &= \begin{pmatrix} 0 & a_0(t) & 0 & 0 & \dots\\ a_0(t) & 0 & a_1(t) & 0 & \dots\\ 0 & a_1(t) & 0 & a_2(t) & \dots\\ & \ddots & \ddots & \ddots & \end{pmatrix}.\end{aligned} \tag{1.2.4}$$

Observe that L is a Jacobi operator, with only the superdiagonal, diagonal and subdiagonal nonzero. The spectrum of $L(t)$ does not depend on time. These eigenvalues give a set of independent integrals of motion: the Toda lattice is completely integrable. In particular, the Toda lattice can be solved by virtue of the inverse scattering transform for the Jacobi operator L. For arbitrary and sufficiently fast decaying initial conditions asymptotically for large t the solution splits into a sum of solitons and a decaying dispersive part. The inverse scattering transform for this system was applied to find solutions in [84,52]. Also in 1975 Mark Kac and Pierre van Moerbeke published two articles in PNAS regarding the Toda Lattice. In [73] a discrete version of Floquet's theory was applied to a system of non-linear differential equations related to the periodic Toda lattice and some solutions found by Toda where shown to fit in the inverse scattering formalism, but more important was [74] where the motion of the periodic Toda lattice was explicitly determined in terms of Abelian integrals.

1.3. Gauss–Borel factorization in integrable systems and orthogonal polynomials

The seminal paper of Mikio Sato [104,105], and further developments performed by the Kyoto school through the use of the bilinear equation and the τ-function formalism [38–40], settled the basis for the Lie group theoretical description of integrable hierarchies, in this direction we have the relevant contribution by Motohico Mulase [95] in which the factorization problems, dressing procedure, and linear systems were the key for integrability. In this dressing setting the multicomponent integrable hierarchies of Toda type were analyzed in depth by Kimio Ueno and Kanehisa Takasaki [114–116]. See also the papers [24,25] and [72] on the multi-component KP hierarchy and [86] on the multi-component Toda lattice hierarchy. In a series of papers Mark Adler and Pierre van Moerbeke showed how the Gauss–Borel factorization problem appears in the theory of the 2D Toda hierarchy and what they called the discrete KP hierarchy [1–6]. These papers clearly established—from a group-theoretical setup—why standard orthogonality of polynomials and integrability of nonlinear equations of Toda type where so close. In fact, the Gauss–Borel factorization of the moment matrix may be understood as the Gauss–Borel factorization of the initial condition for the integrable hierarchy. To see the connection between the work of Mulase and that of Adler and van Moerbeke see [49]. Later on, in the recent paper [7], it is shown that the multiple orthogonal construction described in previous paragraphs was linked with the multi-component KP hierarchy.

In the Madrid group, based on the Gauss–Borel factorization, we have been searching further the deep links between the Theory of Orthogonal Polynomials and the Theory of Integrable Systems. In [13] we studied the generalized orthogonal polynomials [1] and their matrix extensions from the Gauss–Borel viewpoint. In [14] we gave a complete study in terms of factorization for multiple orthogonal polynomials of mixed type and characterized the integrable systems associated to them. Then, we studied Laurent orthogonal polynomials in the unit circle trough the CMV approach in [11] and find in [12] the Christoffel–Darboux formula for generalized orthogonal matrix polynomials. These methods where further extended, for example we gave an alternative Christoffel–Darboux formula for mixed multiple orthogonal polynomials [16] or developed the corresponding theory of matrix Laurent orthogonal polynomials in the unit circle and its associated Toda type hierarchy [15].

1.4. On quasi-determinants

We would like to make some comments on Schur complements and quasi-determinants. Besides its name observe that the Schur complement was not introduced by Issai Schur but by Emilie Haynsworth in 1968 in [65,66]. In fact, Haynsworth coined that named because the Schur determinant formula given in what today is known as Schur lemma in [107]. In the book [128] one can find an ample overview on Schur complement and many of its applications. The most easy examples of quasi-determinants are Schur complements. Israel Gel'fand and collaborators have made many important contributions to the subject and the survey article [54] is an excellent reference. In addition, we also recommend Peter Olver's paper on multivariate interpolation where in §3 the reader will find an alternative interesting approach to the subject. In the late 1920s Archibald Richardson [100,101], one of the two responsible of Littlewood–Richardson rule, and the famous logician Arend Heyting [68], founder of intuitionist logic, studied possible extensions of the determinant notion to division rings. Heyting defined the *designant* of a matrix with noncommutative entries, which for 2×2 matrices was the Schur complement, and generalized to larger dimensions by induction. Let us stress that both Richardson's and Heyting's *quasi-determinants* were generically rational functions of the matrix coefficients. Soon, in 1931, Oystein Ore [97] manifested his disgust with the rational character of the just introduced quasi-determinant and gave a polynomial proposal, the Ore's determinant. A definitive impulse to the modern theory was given by the Gel'fand's school [55,46,47, 56–58]. Quasi-determinants where defined over free division rings and it was early noticed that it is not an analog of the commutative determinant but rather of a ratio determinants. A cornerstone for quasi-determinants is the *heredity principle*, quasi-determinants of quasi-determinants are quasi-determinants; there is no analog of such a principle for determinants. However, many of the properties of determinants extend to this case, see the cited papers and also [80] for quasi-minors expansions. Let us mention that in the early 1990s the Gel'fand school [56] already noticed the role quasi-determinants for some integrable systems, see also [99] for some recent work in this direction regarding non-Abelian Toda and Painlevé II equations. Jon Nimmo and his collaborators, the Glasgow school, have studied the relation of quasi-determinants and integrable systems, in particular we can mention the papers [60,61,81,59,82]; in this direction see also [63,129,64]. All this paved the route, using the connection with orthogonal polynomials *à la Cholesky*, to the appearance of quasi-determinants in the multivariate orthogonality context. Later, in 2006 Peter Olver applied quasi-determinants to multivariate interpolation [96]. This is the approach we apply in this paper. As in [96] the blocks have different sizes, and so multiplication of blocks is only allowed if they are *compatible*. In general, the (noncommutative) multiplication makes sense if the number of columns and rows of the blocks involved fit well. Moreover, we are only permitted to invert diagonal entries that in general makes the minors expansions by columns or rows not applicable [80] but allows for other result, like the Sylvester's theorem, to hold in this wider scenario.

1.5. Aims, results and structure of the paper

The question was posed to us by Jeff Geronimo: What about the integrable systems associated with multivariate orthogonal polynomials? To answer this question we consulted [44] and we readily noticed the ubiquity of the Gauss–Borel factorization in the subject, and therefore the opportunity to link it with the theory of integrable systems. Once this fact was realized we applied the factorization technology of the moment matrix to reproduce the general theory presented in [44].

The main difference with the case of orthogonal polynomials in the real line (OPRL) is that now the moment matrix is a block matrix, with its elements being rectangular matrices of varying size. We had come across matrix blocks before when we studied matrix orthogonal polynomials, but there the size of each block was fixed, now is variable. This intrinsic fact, leagued with the multivariate character, leads on the one hand to the appearance of Schur complements and quasi-determinants and, on the other hand, to multivariate Cauchy integrals and integrals along the Shilov border of polydisks—that is, to be faced with some basic facts of complex analysis in several variables. The Schur complement already appeared in the study of matrix orthogonal polynomials, see for example [15,31,32], but we did not understood yet in [15] the important role played in the theory by quasi-determinants; now we do. We also get across the symmetric algebra [48,77] which is isomorphic to the set of multivariate polynomials. All the necessary material regarding these issues can be found in the appendices.

1.5.1. Results

In the first place we recover a number of classical results from the multivariate orthogonality general theory, see for example [44], using a Cholesky factorization[3] of a symmetric moment matrix. We got the multivariate orthogonal polynomials associated with a given Borel measure and the corresponding second kind functions, that happen to be multivariate Cauchy transforms of the polynomials. All these objects have *quasi-determinantal expressions in terms of bordered truncated moment matrices.* Then, the shift matrices allow to get the three term relation and also Jacobi type matrices and Christoffel–Darboux formulae.

Once we have been able to reproduce, with a Cholesky flavor, classical results for multivariate orthogonal polynomials, we begun the quest of discrete and continuous deformations of the measure which lead to equations of the Toda type. We found both partial difference and partial differential nonlinear equations for the varying size matrices.

[3] A Gauss–Borel factorization for the symmetric case.

Moreover, we introduce *quasi-tau matrices* and find the analogous, in this multi-variable scenario, to the 1D expressions of the orthogonal polynomials and their second kind functions as ratios of Miwa shifted tau functions. Besides these achievements we noticed that the discrete flows allow for the finding of the multivariate extension of the elementary Darboux transformations via what we named as the *sample matrix trick*. These allow not only to express the kernel polynomials, but also their second kind functions, the quasi-tau matrices and some other important coefficients as quasi-determinants of the original data. The sample matrix trick allows also for the study of iterated Darboux transformations and the finding of the *multivariate version of the Christoffel formula*. Many relevant elements of Toda integrable theory, as linear systems, Lax equations, Zakharov–Shabat equations and bilinear equations, are found. An *asymptotic module* or *asymptotic congruence* arguments permit for another perspective of the hierarchy, and we find KP type equations for this multivariate case. Finally, a *linear isometry invariance* of the measure is assumed and we get, through the Cholesky factorization, the consequences for the multivariate orthogonality and the corresponding integrable systems.

1.5.2. The layout of the paper

After this introduction we discuss in §2 the general theory of multivariate orthogonal polynomials by using the Cholesky factorization of a moment matrix. We describe the monomials and order them, according to the reserve lexicographic order, so that we can analyze the conditions for the Cholesky factorization to hold and find the multivariate orthogonal polynomials and their associated second kind functions and their integral representation. The shift matrices are introduced and the three term relations are recovered. The Christoffel–Darboux formulae is deduced in this context.

In §3 we introduce discrete Toda deformations of the measure, we find the corresponding integrable discrete flows, wave matrices, lattice resolvents and Lax (or Jacobi type matrices) pairs are given; a quasi-determinantal expression in terms of the Jacobi matrix for the lattice resolvent is found. Discrete Lax and Zakharov–Shabat equations and corresponding discrete Toda type equations for the varying size quasi-tau functions are described. Then, we find some interesting expressions for the multivariate orthogonal polynomials and their second kind functions in terms of quasi-tau functions and their shifts. For the orthogonal polynomials we need to use the Moore–Penrose pseudo-inverse of a matrix given in terms of the shift matrices and for the second kind functions we need to use a composed, or total, translation. In the 1D scenario these formulae are the well known expressions for these objects in terms of quotients of tau functions and their Miwa shifts (which happen to be discrete flows). We observe that these discrete transformations are elementary Darboux (or Christoffel) transformations and we are able, introducing the sample matrix trick, to give an explicit expression for the transformed polynomials in terms quasi-determinants of the original ones. The iteration of these multivariate elementary Darboux leads to a multivariate Christoffel formulae expressing the new orthogonal polynomials $\tilde{P}_{[k]}(\boldsymbol{x})$ in terms of quasi-determinants of the original ones $P_{[k]}(\boldsymbol{x}),\dots,P_{[k+n]}(\boldsymbol{x})$ evaluated at some appropriate nodes. This approach leads to the finding of quasi-determinantal expressions for the kernel polynomials in terms of the evaluation of the Christoffel–Darboux kernels.

Continuous Toda deformations of the measure are discussed in §4. We introduce Baker and adjoint Baker functions in terms of multivariate orthogonal polynomials and their multivariate Cauchy transforms, we find the corresponding Lax and Zakharov–Shabat equations and write a continuous Toda type equation for the quasi-tau matrices. The discrete flows are identified with Miwa shifts and the bilinear equations, with integrals along tori—Shilov borders of appropriate polydisks—are given. Next, in §5 we apply the congruence technique to find KP type equations, nonlinear equations that relate through nonlinear partial differential–difference equations coefficients of the polynomials but for the same k, not involving, as it does happen in the Toda scenario, near neighbors $k+1$ and $k-1$. Using this method we connect discrete and continuous flows. Then, we present linear equations and corresponding nonlinear partial differential equations for the second order flows. We end the section by exploring the linear equations for the third order flows and giving some hints for higher order flows. Finally, in §6 we study some linear isometry type symmetries of the measure and its consequences on the multivariate orthogonal polynomials; we discuss also what discrete or continuous flows preserve this symmetry.

In the appendices we present some necessary material for reading of the paper. In particular, compositions, multisets and symmetric algebras are briefly treated in Appendix A. Then, in Appendix B we recall some aspects of pseudo-inverses, Schur complements and quasi-determinants and, in Appendix C, we give some notations and results that appear in the analysis in several complex variables. For the sake of clarity some of the proofs of propositions and theorems have been collected in Appendix D.

1.5.3. Further lines

From the submission of this paper to its acceptance for publication we have publish three more papers on transformations for multivariate orthogonal polynomials. In [17] we considered the theory of Christoffel transformations (here called Darboux transformations) in full generality. Notice that the treatment given in the present article, composing elementary degree one Darboux transformations, do not cover the general situation, as in the multivariate case the irreducible polynomials could have any degree and, consequently, an arbitrary polynomial possibly will not factor in terms of degree one polynomials. Then, in [18], we considered more general transformations, including linear functionals and the multiplication by multivariate rational functions. In [19] we extended these constructions to the multidimensional torus and multivariate Laurent polynomials.

Finally, let us comment that, regarding the theory on transformations for matrix orthogonal polynomials, recently two papers [9,10] have treated that subject using factorization techniques and the spectral theory of matrix polynomials. In the first one only the Christoffel transformation is considered, while in the second one the theory is developed in full generality for perturbation of sesquilinear forms and rational transformations with masses.

2. Multivariate orthogonality à la Cholesky

We study multivariate orthogonal polynomials in a D-dimensional real space (MVOPR) in terms of a Cholesky factorization of a semi-infinite moment matrix. We consider D independent real variables $\boldsymbol{x} = (x_1, x_2, \dots, x_D)^\top \in \Omega \subseteq \mathbb{R}^D$ varying in the domain Ω together with a Borel measure $\mathrm{d}\mu(x) \in \mathcal{B}(\Omega)$. The inner product of two real valued functions $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ is defined by

$$\langle f, g\rangle := \int_\Omega f(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})g(\boldsymbol{x}).$$

2.1. Ordering the monomials

Given a multi-index $\alpha = (\alpha_1, \dots, \alpha_D)^\top \in \mathbb{Z}_+^D$ of non-negative integers we write $\boldsymbol{x}^\alpha = x_1^{\alpha_1} \cdots x_D^{\alpha_D}$; the length[4] of α is $|\alpha| := \sum_{a=1}^D \alpha_a$. This length induces the total ordering of monomials, $\boldsymbol{x}^\alpha < \boldsymbol{x}^{\alpha'} \Leftrightarrow |\alpha| < |\alpha'|$, that we will use to arrange the monomials. For each non-negative integer $k \in \mathbb{Z}_+$ we introduce the set

$$[k] := \{\alpha \in \mathbb{Z}_+^D : |\alpha| = k\},$$

built up with those vectors in the lattice $\mathbb{Z}_+^D$ with a given length k.

We will use the graded reversed lexicographic order; i.e., for $\alpha_1, \alpha_2 \in [k]$

$$\begin{aligned}\alpha_1 > \alpha_2 \Leftrightarrow \exists p \in \mathbb{Z}_+ \text{ with } p < D \text{ such that } \alpha_{1,1} = \alpha_{2,1}, \dots, \alpha_{1,p} = \alpha_{2,p}\\ \text{and } \alpha_{1,p+1} < \alpha_{2,p+1},\end{aligned}$$

and if $\alpha^{(k)} \in [k]$ and $\alpha^{(\ell)} \in [\ell]$, with $k < \ell$ then $\alpha^{(k)} < \alpha^{(\ell)}$. Given the set of integer vectors of length k we use the reversed lexicographic order and write

$$[k] = \left\{\alpha_1^{(k)}, \alpha_2^{(k)}, \dots, \alpha_{|[k]|}^{(k)}\right\} \text{ with } \alpha_a^{(k)} < \alpha_{a+1}^{(k)}.$$

Here $|[k]|$ is the cardinality of the set $[k]$, i.e., the number of elements in the set. Observe that $|[0]| = 1$, $|[1]| = D$ and $|[2]| = \frac{(D+1)D}{2}$.

We introduce the vectors of monomials

[4] Also known as absolute value, order or norm.

$$
\chi := \begin{pmatrix} \chi_{[0]} \\ \chi_{[1]} \\ \vdots \\ \chi_{[k]} \\ \vdots \end{pmatrix} \qquad \text{where} \qquad \chi_{[k]} := \begin{pmatrix} \boldsymbol{x}^{\alpha_1} \\ \boldsymbol{x}^{\alpha_2} \\ \vdots \\ \boldsymbol{x}^{\alpha_{|[k]|}} \end{pmatrix},
$$

$$
\chi^* := \Big(\prod_{a=1}^{D} x_a^{-1}\Big)\chi(x_1^{-1},\dots,x_D^{-1});
$$

for example $\chi_{[0]} = 1$, $\chi^*_{[0]} = \prod_{a=1}^{D} x_a^{-1}$ and

$$
\chi_{[1]} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_D \end{pmatrix}, \qquad \chi^*_{[1]} = \Big(\prod_{a=1}^{D} x_a^{-1}\Big)\begin{pmatrix} x_1^{-1} \\ x_2^{-1} \\ \vdots \\ x_D^{-1} \end{pmatrix},
$$

$$
\chi_{[2]} = \begin{pmatrix} x_1^2 \\ x_1x_2 \\ \vdots \\ x_1x_D \\ x_2^2 \\ x_2x_3 \\ \vdots \\ x_2x_D \\ x_3^2 \\ \vdots \\ x_D^2 \end{pmatrix}, \qquad \chi^*_{[2]} = \Big(\prod_{a=1}^{D} x_a^{-1}\Big)\begin{pmatrix} x_1^{-2} \\ x_1^{-1}x_2^{-1} \\ \vdots \\ x_1^{-1}x_D^{-1} \\ x_2^{-2} \\ x_2^{-1}x_3^{-1} \\ \vdots \\ x_2^{-1}x_D^{-1} \\ x_3^{-2} \\ \vdots \\ x_D^{-2} \end{pmatrix}.
$$

Observe that for $k = 1$ we have that the vectors $\alpha_a^{(1)} = \boldsymbol{e}_a$ for $a \in \{1,\dots,D\}$ form the canonical basis of $\mathbb{R}^D$, and for any $\alpha_j \in [k]$ we have $\alpha_j = \sum_{a=1}^{D} \alpha_j^a \boldsymbol{e}_a$. For the sake of simplicity unless needed otherwise we will drop off the super-index and write α_j instead of $\alpha_j^{(k)}$, as is understood that $|\alpha_j| = k$. Notice that

$$
\chi^\top(\boldsymbol{y})\chi^*(\boldsymbol{x}) = \prod_{a=1}^{D} \frac{1}{x_a - y_a}, \quad \forall \boldsymbol{x},\boldsymbol{y} \in \mathbb{C}^D \text{ such that } |x_a| > |y_a|. \tag{2.1.1}
$$

2.2. Monomials and symmetric tensor powers

The dual space of the symmetric tensor powers, see Appendix A, happens to be isomorphic to the set of symmetric multilinear functionals on $\mathbb{R}^D$, $\big(\operatorname{Sym}^k(\mathbb{R}^D)\big)^* \cong S((\mathbb{R}^D)^k, \mathbb{R})$. Hence, homogeneous polynomials of a given total degree can be identified with symmetric tensor powers. Each multi-index $\boldsymbol{\alpha} \in [k]$[5] can be thought of as a weak D-composition of k (or weak composition in D parts), $k = \alpha_1 + \cdots + \alpha_D$. Notice that these weak compositions may be considered as multisets and that, given a linear basis $\{\boldsymbol{e}_a\}_{a=1}^D$ of $\mathbb{R}^D$, as we know from Appendix A, we have the linear basis $\{\boldsymbol{e}_{a_1} \odot \cdots \odot \boldsymbol{e}_{a_k}\}_{\substack{1\leq a_1\leq\cdots\leq a_k\leq D\\ k\in\mathbb{Z}_+}}$ for the symmetric power $\mathrm{S}^k(\mathbb{R}^D)$, where we are using multisets $1 \leq a_1 \leq \cdots \leq a_k \leq D$. In particular, see Appendix A.2.2, the vectors of this basis $\boldsymbol{e}_{a_1}^{\odot M(a_1)} \odot \cdots \odot \boldsymbol{e}_{a_p}^{\odot M(a_p)}$, or better its duals $(\boldsymbol{e}^*_{a_1})^{\odot M(a_1)} \odot \cdots \odot (\boldsymbol{e}^*_{a_p})^{\odot M(a_p)}$ are in bijection with monomials of the form $x_{a_1}^{M(a_1)} \cdots x_{a_p}^{M(a_p)}$. Therefore, either counting weak compositions or multisets we are led to the following conclusion: the cardinality of $[k]$ is $|[k]| = \left(\!\binom{D}{k}\!\right) = \binom{D+k-1}{k}$.

The monomials can be nicely expressed in terms of symmetric products and the multinomial matrix, see Appendix A. The reverse lexicographic order can be applied to $\big(\mathbb{R}^D\big)^{\odot k} \cong \mathbb{R}^{|[k]|}$, we then take a linear basis of $\mathrm{S}^k(\mathbb{R}^D)$ as the ordered set $B_c = \{\boldsymbol{e}^{\alpha_1}, \ldots, \boldsymbol{e}^{\alpha_{|[k]|}}\}$ with $\boldsymbol{e}^{\alpha_j} := \boldsymbol{e}_1^{\odot\alpha_j^1} \odot \cdots \odot \boldsymbol{e}_D^{\odot\alpha_j^D}$ so that $\chi_{[k]}(\boldsymbol{x}) = \sum_{i=1}^{|[k]|} \boldsymbol{x}^{\alpha_j} \boldsymbol{e}^{\alpha_j}$. This means that in this canonical basis the column matrix representing $\chi_{[k]}$ is $[\chi_{[k]}]_{B_c} = \begin{pmatrix} \boldsymbol{x}^{\alpha_1} \\ \vdots \\ \boldsymbol{x}^{\alpha_{|[k]|}} \end{pmatrix}$. We will identify $\chi_{[k]}$ with $[\chi_{[k]}]_{B_c}$.

Proposition 2.2.1. *If $[\boldsymbol{x}^{\odot k}]_{B_c}$ is the column matrix representing $\boldsymbol{x}^{\odot k}$ in the canonical basis B_c we have*

$$\chi_{[k]}(\boldsymbol{x}) = \big(\mathcal{M}_{[k]}\big)^{-1} [\boldsymbol{x}^{\odot k}]_{B_c}. \tag{2.2.1}$$

Proof. It is a consequence of the multinomial theorem for symmetric powers

$$\begin{aligned} \boldsymbol{x}^{\odot k} &= (x_1\boldsymbol{e}_1 + \cdots + x_D\boldsymbol{e}_D)^{\odot k} \\ &= \sum_{j=1}^{|[k]|} \binom{k}{\alpha_j} \boldsymbol{x}^{\alpha_j} \boldsymbol{e}_1^{\odot\alpha_j^1} \odot \cdots \odot \boldsymbol{e}_D^{\odot\alpha_j^D} \\ &= \mathcal{M}_{[k]}\chi_{[k]}(\boldsymbol{x}). \qquad \square \end{aligned}$$

2.3. Cholesky factorization of the moment matrix

In this paper we will consider semi-infinite matrices A with a block or partitioned structure induced by the graded reversed lexicographic order

$$A = \begin{pmatrix} A_{[0],[0]} & A_{[0],[1]} & \cdots \\ A_{[1],[0]} & A_{[1],[1]} & \cdots \\ \vdots & \vdots & \end{pmatrix}, \quad A_{[k],[\ell]} = \begin{pmatrix} A_{\alpha_1^{(k)},\alpha_1^{(\ell)}} & \ldots & A_{\alpha_1^{(k)},\alpha_{|[\ell]|}^{(\ell)}} \\ \vdots & & \vdots \\ A_{\alpha_{|[k]|}^{(k)},\alpha_1^{(\ell)}} & \ldots & A_{\alpha_{|[k]|}^{(k)},\alpha_{|[\ell]|}^{(\ell)}} \end{pmatrix} \in \mathbb{R}^{|[k]|\times|[\ell]|}.$$

[5] Observe that in [48] we have diverse notation $[k] \equiv \Xi(D,k)$.

We use the notation $0_{[k],[\ell]} \in \mathbb{R}^{|[k]|\times|[\ell]|}$ for the rectangular zero matrix, $0_{[k]} \in \mathbb{R}^{|[k]|}$ for the zero vector, and $\mathbb{I}_{[k]} \in \mathbb{R}^{|[k]|\times|[k]|}$ for the identity matrix. For the sake of simplicity we normally just write 0 or $\mathbb{I}$ for the zero or identity matrices, and we implicitly assume that the sizes of these matrices are the ones indicated by their position in the partitioned matrix.

Definition 2.3.1. Associated with the measure $\mathrm{d}\mu$ we have the following moment matrix

$$G := \int_\Omega \chi(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})\chi(\boldsymbol{x})^\top.$$

We write the moment matrix in block form

$$G = \begin{pmatrix} G_{[0],[0]} & G_{[0],[1]} & \cdots \\ G_{[1],[0]} & G_{[1],[1]} & \cdots \\ \vdots & \vdots & \end{pmatrix}$$

with each entry being a rectangular matrix with real coefficients

$$\begin{aligned} G_{[k],[\ell]} &:= \int_\Omega \chi_{[k]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})(\chi_{[\ell]}(\boldsymbol{x}))^\top, \qquad k,\ell = 0,1,\dots, \qquad (2.3.1) \\ &= \begin{pmatrix} G_{\alpha_1^{(k)},\alpha_1^{(\ell)}} & \cdots & G_{\alpha_1^{(k)},\alpha_{|[\ell]|}^{(\ell)}} \\ \vdots & & \vdots \\ G_{\alpha_{|[k]|}^{(k)},\alpha_1^{(\ell)}} & \cdots & G_{\alpha_{|[k]|}^{(k)},\alpha_{|[\ell]|}^{(\ell)}} \end{pmatrix} \in \mathbb{R}^{|[k]|\times|[\ell]|}, \\ G_{\alpha_i^{(k)},\alpha_j^{(\ell)}} &:= \int_\Omega \boldsymbol{x}^{\alpha_i^{(k)}+\alpha_j^{(\ell)}}\mathrm{d}\mu(\boldsymbol{x}) \in \mathbb{R}. \end{aligned}$$

Truncated moment matrices are given by

$$G^{[\ell]} := \begin{pmatrix} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} \\ \vdots & & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} \end{pmatrix},$$

and for $k \geq \ell$ we will also use the following bordered truncated moment matrix

$$G_k^{[\ell+1]} := \left(\begin{array}{ccc} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} \\ \vdots & & \vdots \\ G_{[\ell-2],[0]} & \cdots & G_{[\ell-2],[\ell-1]} \\ \hline G_{[k],[0]} & \dots & G_{[k],[\ell-1]} \end{array}\right)$$

where we have replaced the last row of blocks, $\begin{pmatrix} G_{[\ell-1],[0]} & \dots & G_{[\ell-1],[\ell-1]} \end{pmatrix}$, of the truncated moment matrix $G^{[\ell+1]}$ by the row of blocks $\begin{pmatrix} G_{[k],[0]} & \dots & G_{[k],[\ell-1]} \end{pmatrix}$.

Notice that from the above definition we know that the moment matrix is a symmetric matrix, $G = G^\top$, which implies that a Gauss–Borel factorization of it, in terms of unitriangular lower[6] and upper triangular matrices, is a Cholesky factorization. We describe now when and how the Cholesky factorization of the moment can be performed. The result and its proof uses Schur complements, see Appendix B.2.

Proposition 2.3.1.

(1) *If* $\det G^{[\ell]} \neq 0$ *for all* $\ell = 0, 1, \dots$ *then* G *admits the following Cholesky type factorization*

$$G = S^{-1} H \left(S^{-1}\right)^\top, \tag{2.3.2}$$

with

$$S^{-1} = \begin{pmatrix} \mathbb{I} & 0 & 0 & \cdots \\ (S^{-1})_{[1],[0]} & \mathbb{I} & 0 & \cdots \\ (S^{-1})_{[2],[0]} & (S^{-1})_{[2],[1]} & \mathbb{I} & \\ \vdots & \vdots & & \ddots \end{pmatrix},$$

$$H = \begin{pmatrix} H_{[0]} & 0 & 0 & \\ 0 & H_{[1]} & 0 & \cdots \\ 0 & 0 & H_{[2]} & \\ \vdots & \vdots & & \ddots \end{pmatrix}.$$

(2) *When* $\det G^{[\ell]} \neq 0$ *the Cholesky type factorization holds and*

$$\det G^{[\ell]} = \prod_{k=0}^{\ell-1} \det H_{[k]} \neq 0$$

so that all $H_{[k]}$ *are invertible,* $k = 0, 1, \dots$.

Proof. See Appendix D.1. □

A quasi-determinant version, see Appendix B, of the above result can be given

Proposition 2.3.2. *If the quasi-determinants of the truncated moment matrices are invertible*

$$\det \Theta_*(G^{[k+1]}) \neq 0, \qquad k = 0, 1, \dots,$$

[6] Lower triangular with the block diagonal populated by identity matrices.

the Cholesky factorization (2.3.2) can be performed where

$$H_{[k]} = \Theta_*(G^{[k+1]}), \qquad (S^{-1})_{[k],[\ell]} = \Theta_*(G^{[\ell+1]}_k)\Theta_*(G^{[\ell+1]})^{-1}.$$

Proof. It is just a consequence of Theorem 3 of [96], see Appendix B. □

2.3.1. On quasi-tau functions

In the 1D scenario the tau functions can be introduced as the determinant of a truncated moment matrix

$$\tau_k := \det G^{[k]}, \qquad k = 1, 2, \dots$$

and $\tau_0 = 1$, so that

$$H_k = \frac{\tau_{k+1}}{\tau_k}, \qquad k = 0, 1, 2, \dots \tag{2.3.3}$$

and

$$\tau_{k+1} = H_k H_{k-1} \cdots H_0.$$

Moreover, observe that (2.3.3) can be written as a quasi-determinant

$$H_k = \det(G^{[k+1]}/G^{[k]}) = \Theta_*(G^{[k+1]}).$$

Thus, in the 1D scenario the described analogy suggests that the squared norms H_k can be considered as quasi-tau functions, being the tau functions $\tau_k = \det G^{[k]}$ determinants of the truncated moment matrix and the quasi-tau functions $H_k = \Theta_*(G^{[k+1]})$ quasi-determinants of the truncated moment matrix. This extends to the multivariant setting and now we have $H_{[k]} = \Theta_*(G^{[k+1]})$, motivating us to refer to these matrices as quasi-tau matrices. Let us mention that other authors have introduced similar concepts before, for example in [92] a matrix valued tau function is considered for the case of matrix orthogonal polynomials. However, the motivation of the author did not come from the quasi-determinant expressions in terms of the moment matrix but from formulae from integrable systems.

2.4. MVOPR

With the aid of the Cholesky factorization we are ready to introduce the MVOPR

Definition 2.4.1. The MVOPR associated to the measure $\mathrm{d}\mu$ are

$$P = S\chi = \begin{pmatrix} P_{[0]} \\ P_{[1]} \\ \vdots \end{pmatrix}, \qquad P_{[k]}(\boldsymbol{x}) = \sum_{\ell=0}^{k} S_{[k],[\ell]}\chi_{[\ell]}(\boldsymbol{x}) = \begin{pmatrix} P_{\alpha_1^{(k)}} \\ \vdots \\ P_{\alpha^{(k)}_{|[k]|}} \end{pmatrix},$$

$$P_{\alpha_i^{(k)}}=\sum_{\ell=0}^{k}\sum_{j=1}^{|[\ell]|}S_{\alpha_i^{(k)},\alpha_j^{(\ell)}}\boldsymbol{x}^{\alpha_j^{(\ell)}}. \tag{2.4.1}$$

We introduce the coefficients

$$\beta_{[k]}:=S_{[k],[k-1]},\qquad k\geq 1,$$

which take values in the linear space of rectangular matrices $\mathbb{R}^{|[k]|\times|[k-1]|}$ and also define the subdiagonal matrix

$$\beta=\begin{pmatrix}0&0&0&0&\cdots\\ \beta_{[1]}&0&0&0&\cdots\\ 0&\beta_{[2]}&0&0&\cdots\\ 0&0&\beta_{[3]}&0&\cdots\\ \vdots&\vdots&\ddots&\ddots&\ddots\end{pmatrix}.$$

An immediate consequence of Proposition 2.3.2 is

Proposition 2.4.1. *The following quasi-determinantal expression holds true*

$$\beta_{[k]}=-\Theta_*(G_k^{[k]})\Theta_*(G^{[k]})^{-1}.$$

Observe that $P_{[k]}=\chi_{[k]}(\boldsymbol{x})+\beta_{[k]}\chi_{[k-1]}(\boldsymbol{x})+\cdots$ is a vector constructed with the polynomials $P_{\alpha_i}(\boldsymbol{x})$ of degree k, each of which has only one monomial of degree k; i.e., we can write $P_{\alpha_i}(\boldsymbol{x})=\boldsymbol{x}^{\alpha_i}+Q_{\alpha_i}(\boldsymbol{x})$, with $\deg Q_{\alpha_i}<k$.

Proposition 2.4.2. *The MVOPR satisfy*

$$\int_\Omega P_{[k]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})(P_{[\ell]}(\boldsymbol{x}))^\top=\int_\Omega P_{[k]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})(\chi_{[\ell]}(\boldsymbol{x}))^\top=0,$$
$$\ell=0,1,\dots,k-1, \tag{2.4.2}$$
$$\int_\Omega P_{[k]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})(P_{[k]}(\boldsymbol{x}))^\top=\int_\Omega P_{[k]}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})(\chi_{[k]}(\boldsymbol{x}))^\top=H_{[k]}. \tag{2.4.3}$$

Therefore, we have the following orthogonality conditions

$$\int_\Omega P_{\alpha_i^{(k)}}(\boldsymbol{x})P_{\alpha_j^{(\ell)}}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})=\int_\Omega P_{\alpha_i^{(k)}}(\boldsymbol{x})\boldsymbol{x}^{\alpha_j^{(\ell)}}\mathrm{d}\mu(\boldsymbol{x})=0,$$
$$\ell=0,1,\dots,k-1,\quad i=1,\dots,|[k]|,\quad j=1,\dots,|[\ell]|,$$

with the normalization conditions

$$\int_\Omega P_{\alpha_i}(\boldsymbol{x})P_{\alpha_j}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})=\int_\Omega P_{\alpha_i}(\boldsymbol{x})\boldsymbol{x}^{\alpha_j}\mathrm{d}\mu(\boldsymbol{x})=H_{\alpha_i,\alpha_j},\quad i,j=1,\dots,|[k]|.$$

Here we use Dunkl and Xu's notation, see [44]. Despite the MVOPR are orthogonal for different k and ℓ, $P_{\alpha_i^{(k)}} \perp P_{\alpha_j^{(\ell)}}$, for $k=\ell$ the value of $\langle P_{\alpha_i}, P_{\alpha_j}\rangle = H_{\alpha_i,\alpha_j}$ is not zero in general, and the set of polynomials given by coefficients of the vector $P_{[k]}$ are not orthogonal among them. Observe that (2.4.3) implies that the matrices $H_{[k]}$ are Grammian matrices and that, the measure being positive definite, we can write $H_{[k]} = M_{[k]}^\top h_{[k]} M_{[k]}$, for some orthogonal matrix $M_{[k]} \in \mathrm{O}(\mathbb{R}^{|[k]|})$ and diagonal matrix $h_{[k]} = \operatorname{diag}(h_{[k],1}, \dots, h_{[k],|[k]|})$ with $h_{[k],j} > 0$ for $j \in \{1, \dots, |[k]|\}$. With the new vector polynomials $\tilde{P}_{[k]} = M_{[k]} P_{[k]}$ we do have

$$\int_\Omega \tilde{P}_{\alpha_i}(\boldsymbol{x}) \tilde{P}_{\alpha_j}(\boldsymbol{x}) \mathrm{d}\mu(\boldsymbol{x}) = \delta_{i,j} h_{[k],j}, \quad i,j = 1, \dots, |[k]|,$$

$h_{[k],j}$ being the squared norms of the polynomials. Now, instead of a block Cholesky factorization we have a standard Cholesky factorization $G = \tilde{S}^{-1} h (\tilde{S}^{-1})^\top$, with $\tilde{S} = \operatorname{diag}(M_{[0]}^\top, M_{[1]}^\top, \dots) S$ and $h = \operatorname{diag}(h_{[0]}, h_{[1]}, \dots)$. However, this scalar Cholesky factorization does not help much in the understanding of MVOPR, the reason will become clear in §2.6, where the three term relations or the Christoffel–Darboux formulae are deduced from the block Cholesky factorization. The clue is that the shift matrices, for which we have the symmetry (2.6.5) of the moment matrix, are naturally written in block form.

Also notice that $H_{[0]} = \int_\Omega \mathrm{d}\mu(\boldsymbol{x})$ is just the measure of the support.

Proposition 2.4.3. *The MVOPR can be expressed as Schur complements of bordered truncated moment matrices*

$$P_{[\ell]}(\boldsymbol{x}) = \mathrm{SC}\left(\begin{array}{ccc|c} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]}(\boldsymbol{x}) \\ \vdots & & \vdots & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]}(\boldsymbol{x}) \\ \hline G_{[\ell],[0]} & \cdots & G_{[\ell],[\ell-1]} & \chi_{[\ell]}(\boldsymbol{x}) \end{array}\right),$$

or as quasi-determinants

$$P_{[\ell]} = \Theta_* \begin{pmatrix} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]}(\boldsymbol{x}) \\ \vdots & & \vdots & \vdots \\ G_{[\ell],[0]} & \cdots & G_{[\ell],[\ell-1]} & \chi_{[\ell]}(\boldsymbol{x}) \end{pmatrix}.$$

Proof. Any semi-infinite block matrix $M = (M_{i,j})$ can be written in block form $M = \left(\begin{smallmatrix} M^{[\ell]} & M^{[\ell],[\geq \ell]} \\ M^{[\geq \ell][,\ell]} & M^{[\geq \ell]} \end{smallmatrix}\right)$, where $M^{[\ell]} = (M_{i,j})_{\substack{0 \leq i < \ell \\ 0 \leq j < \ell}}$ is the standard truncation, $M^{[\ell],[\geq \ell]} = (M_{i,j})_{\substack{0 \leq i < \ell \\ j \geq \ell}}$, $M^{[\geq \ell],[\ell]} = (M_{i,j})_{\substack{i \geq \ell \\ 0 \leq j < \ell}}$ and $M^{[\geq \ell]} = (M_{i,j})_{\substack{i \geq \ell \\ j \geq \ell}}$.

From the factorization of the moment matrix

$$
\begin{aligned}
SG = H(S^{-1})^\top \quad &\Longrightarrow \quad 0 = S^{[\geq \ell],[\ell]} G^{[\ell]} + S^{[\geq \ell]} G^{[\geq \ell],[\ell]} \\
&\Longrightarrow \quad S^{[\geq \ell],[\ell]} = -S^{[\geq \ell]} G^{[\geq \ell],[\ell]} \big(G^{[\ell]}\big)^{-1}.
\end{aligned}
$$

With this we can rewrite $P_{[\ell]} = \sum_{k=0}^{\ell} S_{[\ell],[k]} \chi_{[k]}$ as

$$
P_{[\ell]} = \chi_{[\ell]} - \begin{pmatrix} G_{[\ell],[0]} & G_{[\ell],[1]} & \cdots & G_{[\ell],[\ell-1]} \end{pmatrix} \Big(G^{[\ell]}\Big)^{-1} \chi^{[\ell]}.
$$

To get the stated result we need only to fix our attention in any of the rows of this matrix. □

In terms of ratios of determinant we get for the components

$$
P_{\alpha_j^{(\ell)}} = \mathrm{SC}\left(\begin{array}{ccc|c}
G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]} \\
\vdots & & \vdots & \vdots \\
G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]} \\
\hline
G_{\alpha_j^{(\ell)},[0]} & \cdots & G_{\alpha_j^{(\ell)},[\ell-1]} & \boldsymbol{x}^{\alpha_j^{(\ell)}}
\end{array}\right) = \frac{\begin{vmatrix}
G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]} \\
\vdots & & \vdots & \vdots \\
G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]} \\
G_{\alpha_j^{(\ell)},[0]} & \cdots & G_{\alpha_j^{(\ell)},[\ell-1]} & \boldsymbol{x}^{\alpha_j^{(\ell)}}
\end{vmatrix}}{\begin{vmatrix}
G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} \\
\vdots & & \vdots \\
G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]}
\end{vmatrix}}.
$$

2.5. Functions of the second kind

In this subsection we need some material regarding several complex variables analysis and we refer the reader to Appendix C. Complementary to the vector P of multivariate polynomials we introduce

Definition 2.5.1. Second kind functions are given by the coefficients of

$$
C := H(S^{-1})^\top \chi^* = \begin{pmatrix} C_{[0]} \\ C_{[1]} \\ \vdots \end{pmatrix}, \qquad C_{[k]} := \begin{pmatrix} C_{\alpha_1} \\ \vdots \\ C_{\alpha_{|[k]|}} \end{pmatrix}. \tag{2.5.1}
$$

Observe that for $\boldsymbol{z} = (z_1, \dots, z_D)^\top \in \mathbb{C}^D$ we have

$$
C_{[\ell]}(\boldsymbol{z}) = H_{[\ell]} \sum_{k=\ell}^{\infty} \big((S^{-1})_{[k],[\ell]}\big)^\top \chi^*_{[k]}(\boldsymbol{z}),
$$

which is a vector with each of its components $C_{\boldsymbol{\ell}_a}(\boldsymbol{z})$, $a = 1, \dots, |[\ell]|$, a D-fold Laurent series. This is just not the case for the definition of P, see (2.4.1), where we had finite sums instead of infinite series. In the case of $C_{\alpha_i}(\boldsymbol{z})$, which has domain of convergence $\mathscr{D}_{\alpha_i}$, we can introduce $\boldsymbol{w} = \boldsymbol{z}^{-1} := (z_1^{-1}, \dots, z_D^{-1})^\top$—i.e., $\boldsymbol{z} = \boldsymbol{w}^{-1}$—and notice that $C_{\alpha_i}(\boldsymbol{z}(\boldsymbol{w})) = \sum_{\boldsymbol{\beta} \in \mathbb{Z}_+^D} c_{\boldsymbol{\beta}} \boldsymbol{w}^{\boldsymbol{\beta}}$ is a power series in $\boldsymbol{w}$ and consequently converges in a complete Reinhardt domain $\mathscr{D}$. Therefore, its domain of convergence is the union of polydisks, and in each of them the convergence is absolute and uniform. In particular, the polydisk of convergence $\Delta(\boldsymbol{r}) \subset \mathscr{D}$ satisfies the extended Cauchy–Hadamard formula $\limsup_{|\boldsymbol{\beta}|\to\infty} \sqrt[|\boldsymbol{\beta}|]{|c_{\boldsymbol{\beta}}|\boldsymbol{r}^{\boldsymbol{\beta}}} = 1$. The domain of convergence of C_{α_i} contains a polyannulus of convergence with polyradii given by $\boldsymbol{r} = \boldsymbol{0}$ and $\boldsymbol{R} = (r_1^{-1}, \dots, r_D^{-1})$.

Let us show that the second kind functions can be expressed as multivariate Cauchy transforms of the MVOPR.

Proposition 2.5.1. *The second kind functions satisfy*

$$C_{\alpha_i}(\boldsymbol{z}) = \int_\Omega \frac{P_{\alpha_i}(\boldsymbol{y})}{(z_1 - y_1)\cdots(z_D - y_D)} \mathrm{d}\mu(\boldsymbol{y}), \quad \forall \boldsymbol{z} \in \mathscr{D}_{\alpha_i} \setminus \operatorname{supp}(\mathrm{d}\mu), \quad i = 1, \dots, |[k]|.$$

Proof. See Appendix D.2. □

We introduce

$$\Gamma := G\chi^* = \begin{pmatrix} \Gamma_{[0]} \\ \Gamma_{[1]} \\ \vdots \end{pmatrix}, \qquad \Gamma_{[k]} := \begin{pmatrix} \Gamma_{\alpha_1} \\ \vdots \\ \Gamma_{\alpha_{|[k]|}} \end{pmatrix}.$$

Proposition 2.5.2. *The coefficients Γ_{α_i} are the multivariate Cauchy transform of the monomials $\boldsymbol{x}^{\alpha_i}$*

$$\Gamma_{\alpha_i}(\boldsymbol{x}) = \int_\Omega \frac{\boldsymbol{y}^{\alpha_i}}{(x_1 - y_1)\cdots(x_D - y_D)} \mathrm{d}\mu(\boldsymbol{y}).$$

Proof. The proof is a byproduct of the proof of Proposition 2.5.1. □

Proposition 2.5.3. *We have $C = S\Gamma$.*

Proposition 2.5.4. *In terms of Schur complements or quasi-determinants of bordered truncated moment matrices the functions of the second kind are*

$$C_{[\ell]} = \operatorname{SC}\left(\begin{array}{ccc|c} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \Gamma_{[0]} \\ \vdots & & \vdots & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \Gamma_{[\ell-1]} \\ \hline G_{[\ell],[0]} & \cdots & G_{[\ell],[\ell-1]} & \Gamma_{[\ell]} \end{array}\right)$$
$$= \Theta_*\begin{pmatrix} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \Gamma_{[0]} \\ \vdots & & \vdots & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \Gamma_{[\ell-1]} \\ G_{[\ell],[0]} & \cdots & G_{[\ell],[\ell-1]} & \Gamma_{[\ell]} \end{pmatrix}.$$

In terms of determinant ratios the components are

$$C_{\alpha_j^{(\ell)}} = \mathrm{SC}\left(\begin{array}{ccc|c} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \Gamma_{[0]} \\ \vdots & & \vdots & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \Gamma_{[\ell-1]} \\ \hline G_{\alpha_j^{(\ell)},[0]} & \cdots & G_{\alpha_j^{(\ell)},[\ell-1]} & \Gamma_{\alpha_j^{(\ell)}} \end{array}\right) = \frac{\begin{vmatrix} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \Gamma_{[0]} \\ \vdots & & \vdots & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \Gamma_{[\ell-1]} \\ G_{\alpha_j^{(\ell)},[0]} & \cdots & G_{\alpha_j^{(\ell)},[\ell-1]} & \Gamma_{\alpha_j^{(\ell)}} \end{vmatrix}}{\begin{vmatrix} G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} \\ \vdots & & \vdots \\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} \end{vmatrix}}.$$

Definition 2.5.2. Given k distinct labels $a_1, \dots, a_k$ in $\{1, \dots, D\}$ we introduce the reduced second kind functions

$$\widehat{C}_{a_1,\dots,a_k} = \lim_{x_{a_1}\to\infty} \cdots \lim_{x_{a_k}\to\infty} \Big(\Big[\prod_{i=1}^{k} x_{a_i}\Big] C\Big). \tag{2.5.2}$$

Observe that the labels in the reduced second kind functions indicate precisely those independent variables on which they do not depend; therefore, when $k = D$ we have a constant. For the reduced second kind functions we find

Proposition 2.5.5. *When* $\operatorname{supp}(\mathrm{d}\mu)$ *is a bounded set the reduced second kind functions fulfill*

$$\widehat{C}_{\alpha_i,a_1,\dots,a_k} = \int_\Omega \frac{P_{\alpha_i}(\boldsymbol{y})}{\prod_{i=1}^{D-k}(z_{b_i} - y_{b_i})} \mathrm{d}\mu(\boldsymbol{y}), \quad a = 1, \dots, |[\ell]|, \quad \boldsymbol{z} \in \mathscr{D}_{\alpha_i} \setminus \operatorname{supp}(\mathrm{d}\mu) \tag{2.5.3}$$

where $\{a_1, \dots, a_k\} \cup \{b_1, \dots, b_{D-k}\} = \{1, \dots, D\}$.

Proof. The Lebesgue dominated convergence theorem ensures that we can interchange the limit with the integral[7]

$$\lim_{z_{a_1}\to\infty} \cdots \lim_{z_{a_k}\to\infty} \Big(\Big[\prod_{i=1}^{k} z_{a_i}\Big] C\Big)$$

[7] The control or dominating function can be taken to be $g_{\boldsymbol{z}}(\boldsymbol{y}) = \frac{P_{\alpha_i}(\boldsymbol{y})}{\prod_{i=1}^{D-k}(z_{b_i}-y_{b_i})}$.

$$
\begin{aligned}
&= \int\limits_{\Omega} \lim_{z_{a_1}\to\infty} \cdots \lim_{z_{a_k}\to\infty} \Big(\Big[\prod_{i=1}^{k} z_{a_i}\Big) \frac{P(\boldsymbol{y})}{\prod_{a=1}^{D}(z_a - y_a)} \mathrm{d}\mu(\boldsymbol{y}) \\
&= \int\limits_{\Omega} \Big[\prod_{i=1}^{k} \lim_{z_{a_i}\to\infty} \Big(1 - \frac{y_{a_i}}{z_{a_i}}\Big)^{-1}\Big] \frac{P(\boldsymbol{y})}{\prod_{i=1}^{D-k}(z_{b_i} - y_{b_i})} \mathrm{d}\mu(\boldsymbol{y}). \qquad \square
\end{aligned}
$$

From this result we infer that

$$
\widehat{C}_{[k],1,\dots,D} = H_{[0]}\delta_{0,k}.
$$

2.6. The shift matrices

In this section we are going to discuss three term relations that extend the recursion relations existing in $D = 1$ to the multivariate case. For this aim we need to introduce a set of D shift matrices $\{\Lambda_1, \dots, \Lambda_D\}$ that play a very important role, they model the action of increasing by one the degree of the monomials.

Definition 2.6.1. The shift matrices are given by

$$
\Lambda_a = \begin{pmatrix}
0 & (\Lambda_a)_{[0],[1]} & 0 & 0 & \cdots \\
0 & 0 & (\Lambda_a)_{[1],[2]} & 0 & \cdots \\
0 & 0 & 0 & (\Lambda_a)_{[2],[3]} & \\
0 & 0 & 0 & 0 & \ddots \\
\vdots & \vdots & \vdots & \vdots &
\end{pmatrix}
$$

where the entries in the nonzero blocks are given by

$$
(\Lambda_a)_{\alpha_i^{(k)},\alpha_j^{(k+1)}} = \delta_{\alpha_i^{(k)}+\boldsymbol{e}_a,\alpha_j^{(k+1)}}, \quad a = 1, \dots, D, \quad i = 1, \dots, |[k]|, \quad j = 1, \dots, |[k+1]|.
$$

Related to these shift matrices we further introduce

Definition 2.6.2.

(1) Given any $\boldsymbol{k} = \sum_{a=1}^{D} k_a \boldsymbol{e}_a \in \mathbb{Z}_+^D$ we define

$$
\Lambda_{\boldsymbol{k}} := \Lambda_1^{k_1} \cdots \Lambda_D^{k_D} \qquad \Pi_{\boldsymbol{k}} := \left(\Lambda_1^\top\right)^{k_1} (\Lambda_1)^{k_1} \cdots \left(\Lambda_D^\top\right)^{k_D} (\Lambda_D)^{k_D}.
$$

(2) For $\boldsymbol{k} = n\boldsymbol{e}_a$, $a \in \{1, \dots, D\}$ and $n \in \mathbb{Z}_+$ we use the notation

$$
\Pi_{a,n} := \Pi_{n\boldsymbol{e}_a} = \left(\Lambda_a^\top\right)^n (\Lambda_a)^n.
$$

If $n = 1$ we use $\Pi_{a,1} = \Pi_a = \Lambda_a^\top \Lambda_a$.

Notice that $\boldsymbol{k}_b+\boldsymbol{e}_a\in[k+1]$.

Proposition 2.6.1.

(1) *The matrices $\Pi_{a,n}$ are projections, $(\Pi_{a,n})^2=\Pi_{a,n}$ and $\Pi_{a,n}=\Pi_{a,n}^\top$. Moreover they are diagonal matrices whose nonvanishing coefficients are the unity. The ones are located precisely in the entries of the diagonal corresponding to the entries (monomials) in χ which contain $(x_a)^m$ with $m\geq n$ among its factors. We can write*

$$\mathbb{I}=\Pi_{a,n}+\Pi_{a,n}^\perp,$$

where $\Pi_{a,n}^\perp$ is a diagonal matrix with its nonvanishing coefficients, which are equal to the unity, located in those entries of χ which contain $(x_a)^m$ with $m<n$ among its factors. We also have

$$\Pi_a^\perp\chi(\boldsymbol{x})=\chi(\boldsymbol{x})|_{x_a=0},\qquad \Pi_a^\perp\chi^*(\boldsymbol{x})=x_a^{-1}\lim_{x_a\to\infty}\big(x_a\chi^*(\boldsymbol{x})\big). \tag{2.6.1}$$

(2) *The shift matrices fulfill the following properties for all $\boldsymbol{k},\boldsymbol{\ell}\in\mathbb{Z}_+^D$*

$$\Lambda_{\boldsymbol{k}}\Lambda_{\boldsymbol{\ell}}=\Lambda_{\boldsymbol{k}+\boldsymbol{\ell}}=\Lambda_{\boldsymbol{\ell}}\Lambda_{\boldsymbol{k}},\qquad \Lambda_{\boldsymbol{k}}(\Lambda_{\boldsymbol{k}})^\top=\mathbb{I}.$$

(3) *When $a\neq b$ we have the commutation relations*

$$\Lambda_b^\top\Lambda_a=\Lambda_a\Lambda_b^\top,\qquad \Pi_a\Pi_b=\Pi_b\Pi_a=\Pi_{\boldsymbol{e}_a+\boldsymbol{e}_b}.$$

(4) *We also have the "eigenvalue" type properties*

$$\Lambda_{\boldsymbol{k}}\chi(\boldsymbol{x})=\boldsymbol{x}^{\boldsymbol{k}}\chi(\boldsymbol{x}),\qquad \Lambda_{\boldsymbol{k}}\chi^*(\boldsymbol{x})=\boldsymbol{x}^{-\boldsymbol{k}}\chi^*(\boldsymbol{x}), \tag{2.6.2}$$

$$\Lambda_{\boldsymbol{k}}^\top\chi(\boldsymbol{x})=\boldsymbol{x}^{-\boldsymbol{k}}\Pi_{\boldsymbol{k}}\chi,\qquad \Lambda_{\boldsymbol{k}}^\top\chi^*(\boldsymbol{x})=\boldsymbol{x}^{\boldsymbol{k}}\Pi_{\boldsymbol{k}}\chi^*(\boldsymbol{x}). \tag{2.6.3}$$

Proposition 2.6.2. *For k distinct labels $a_1,\dots,a_k\in\{1,\dots,D\}$, $a_i\neq a_j$ for $i\neq j$, and k complex numbers $q_{a_1},\dots,q_{a_k}\in\mathbb{C}$ we have*

$$\begin{aligned}\Big[\prod_{i=1}^k(\Lambda_{a_i}^\top-q_{a_i})\Big]\chi^*&=\Big[\prod_{i=1}^k(x_{a_i}-q_{a_i})\Big]\chi^*+(-1)^k\lim_{x_{a_1}\to\infty}\cdots\lim_{x_{a_k}\to\infty}\Big(\Big[\prod_{i=1}^k x_{a_i}\Big]\chi^*\Big),\\ &+\sum_{j=1}^{k-1}\frac{(-1)^j}{(k-j)!j!}\sum_{\sigma\in\mathfrak{S}_k}\Big(\Big[\prod_{i=j+1}^k\big(x_{a_{\sigma(i)}}-q_{a_{\sigma(i)}}\big)\Big]\lim_{x_{a_{\sigma(1)}}\to\infty}\cdots\lim_{x_{a_{\sigma(j)}}\to\infty}\Big(\Big[\prod_{i=1}^j x_{a_{\sigma(i)}}\Big]\chi^*\Big)\Big),\end{aligned} \tag{2.6.4}$$

where $\mathfrak{S}_k$ denotes the symmetric group of k letters.

Proof. See Appendix D.3. □

Proposition 2.6.3. *The moment matrix G satisfies*

$$\Lambda_{\boldsymbol{k}} G = G\big(\Lambda_{\boldsymbol{k}}\big)^\top, \quad \forall \boldsymbol{k} \in \mathbb{Z}_+^D. \tag{2.6.5}$$

Proof. It follows from Definition 2.3.1 of the moment matrix G and the eigenvalue property in Proposition 2.6.1 for $\Lambda_{\boldsymbol{k}}$. □

2.7. Jacobi matrices and three term relations

Once the shift matrices have been introduced we are ready to discuss its *dressing*, that leads to the Jacobi matrices which are extremely important not only for the general theory of MVOPR but also for exploring its connection with the Toda theory.

Definition 2.7.1. We introduce the following Jacobi type matrices

$$J_{\boldsymbol{k}} := S\Lambda_{\boldsymbol{k}} S^{-1}, \quad \forall \boldsymbol{k} \in \mathbb{Z}_+^D, \tag{2.7.1}$$

and the basic Jacobi matrices

$$J_a := J_{\boldsymbol{e}_a}.$$

Proposition 2.7.1. *The Jacobi type matrices satisfy*

$$J_{\boldsymbol{k}}^\top = H^{-1} J_{\boldsymbol{k}} H, \qquad J_{\boldsymbol{k}} P = \boldsymbol{x}^{\boldsymbol{k}} P, \quad \forall \boldsymbol{k} \in \mathbb{Z}_+^D. \tag{2.7.2}$$

Proof. Using the Cholesky factorization of the moment matrix G we get

$$S\Lambda_{\boldsymbol{k}} S^{-1} = H\left(S\Lambda_{\boldsymbol{k}} S^{-1}\right)^\top H^{-1}.$$

The eigenvalue property is obvious from (2.6.2). □

For the second kind functions we have

Proposition 2.7.2. *For k distinct labels $a_1, \dots, a_k \in \{1, \dots, D\}$, $a_i \neq a_j$ for $i \neq j$, and k complex numbers $q_{a_1}, \dots, q_{a_k} \in \mathbb{C}$ we have*

$$\begin{aligned}\Big[\prod_{i=1}^{k}(J_{a_i} - q_{a_i})\Big]C := \Big[\prod_{i=1}^{k}(x_{a_i} - q_{a_i})\Big]C + (-1)^k \widehat{C}_{a_1,\dots,a_k} \\ + \sum_{j=1}^{k-1} \frac{(-1)^j}{(k-j)!j!} \sum_{\sigma\in\mathfrak{S}_k} \Big(\Big[\prod_{i=j+1}^{k}(x_{a_{\sigma(i)}} - q_{a_{\sigma(i)}})\Big]\widehat{C}_{a_{\sigma(1)},\dots,a_{\sigma(j)}}\Big),\end{aligned}$$

Proof. See Appendix D.4. □

From these properties we easily conclude that

Proposition 2.7.3. *The explicit form of the basic Jacobi matrices is*

$$J_a = \begin{pmatrix} (J_a)_{[0],[0]} & (J_a)_{[0],[1]} & 0 & 0 & 0 & \cdots \\ (J_a)_{[1],[0]} & (J_a)_{[1],[1]} & (J_a)_{[1],[2]} & 0 & 0 & \cdots \\ 0 & (J_a)_{[2],[1]} & (J_a)_{[2],[2]} & (J_a)_{[2],[3]} & 0 & \cdots \\ 0 & 0 & (J_a)_{[3],[2]} & (J_a)_{[3],[3]} & (J_a)_{[3],[4]} & \\ \vdots & \vdots & & \ddots & \ddots & \ddots \end{pmatrix},$$

where

$$(J_a)_{[0],[0]} = -(\Lambda_a)_{[0],[1]}\beta_{[1]} \qquad (J_a)_{[0],[1]} = (\Lambda_a)_{[0],[1]}$$

and

$$\begin{aligned} (J_a)_{[k],[k-1]} &= H_{[k]}\left[(\Lambda_a)_{[k-1],[k]}\right]^\top \left(H_{[k-1]}\right)^{-1}, \\ (J_a)_{[k],[k]} &= \beta_{[k]}(\Lambda_a)_{[k-1],[k]} - (\Lambda_a)_{[k],[k+1]}\beta_{[k+1]}, \\ (J_a)_{[k],[k+1]} &= (\Lambda_a)_{[k],[k+1]}. \end{aligned}$$

From (2.7.1) and (2.7.2) it is easy to see that the $J_{\boldsymbol{k}}$ only have $(2|\boldsymbol{k}|+1)$ block diagonals that do not vanish. These are the $|\boldsymbol{k}|$ first block superdiagonals, the diagonal itself and $|\boldsymbol{k}|$ first block subdiagonals.

Definition 2.7.2. We introduce the following objects

$$\boldsymbol{\Lambda} := (\Lambda_1, \dots, \Lambda_D)^\top, \qquad \boldsymbol{J} := (J_1, \dots, J_D)^\top, \qquad \widehat{\boldsymbol{C}} := (\widehat{C}_1, \dots, \widehat{C}_D)^\top,$$

and for given any vector $\boldsymbol{n} = (n_1, \dots, n_D)^\top \in \mathbb{R}^D$ we define the following *dot products*

$$\boldsymbol{n} \cdot \boldsymbol{\Lambda} := \sum_{a=1}^{D} n_a \Lambda_a, \qquad \boldsymbol{n} \cdot \boldsymbol{J} := \sum_{a=1}^{D} n_a J_a.$$

The celebrated three term relations [44] in the multivariate context are

Proposition 2.7.4. *The MVOPR satisfy the following three term relations*[8]

$$\begin{aligned} (\boldsymbol{n}\cdot\boldsymbol{x})P_{[k]} = H_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})^\top_{[k-1],[k]}H^{-1}_{[k-1]}P_{[k-1]} + \big(\beta_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} \\ - (\Lambda_a)_{[k],[k+1]}\beta_{[k+1]}\big)P_{[k]} + (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}P_{[k+1]}, \end{aligned} \tag{2.7.3}$$

[8] Observe that for $k = 0$ we get

$$\boldsymbol{n}\cdot\boldsymbol{x} + (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[0],[1]}\beta_{[1]} = (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[0],[1]}P_{[1]}(\boldsymbol{x}),$$

which agrees with the definition of P in terms of the factorization matrix, that is, $P_{[1]}(\boldsymbol{x})^\top = (x_1, \dots, x_D) + \beta^\top_{[1]}$.

for $k = 1, 2, \dots$. The second kind functions satisfy[9]

$$(\boldsymbol{n}\cdot\boldsymbol{x})C_{[k]} = H_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})^{\top}_{[k-1],[k]}H^{-1}_{[k-1]}C_{[k-1]} + \big(\beta_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\beta_{[k+1]}\big)C_{[k]} + (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}C_{[k+1]} - \boldsymbol{n}\cdot\widehat{\boldsymbol{C}}_{[k]}, \tag{2.7.4}$$

for $k = 1, 2, \dots$.

Proof. These three term relations follow from (2.6.2) and Propositions 2.7.2, for $k = 1$, and 2.7.3. $\square$

2.8. Christoffel–Darboux formulae

Definition 2.8.1. The Christoffel–Darboux kernels are

$$K^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) := \left(\chi^{[\ell]}(\boldsymbol{x})\right)^{\top}\left(G^{[\ell]}\right)^{-1}\chi^{[\ell]}(\boldsymbol{y}) = \sum_{k=0}^{\ell-1}\left(P_{[k]}(\boldsymbol{x})\right)^{\top}\left(H_{[k]}\right)^{-1}P_{[k]}(\boldsymbol{y}),$$

while the second kind Christoffel–Darboux kernels are given by

$$Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) := \left((\chi^*)^{[\ell]}(\boldsymbol{x})\right)^{\top}\left(G^{[\ell]}\right)(\chi^*)^{[\ell]}(\boldsymbol{y}) = \sum_{k=0}^{\ell-1}\left(C_{[k]}(\boldsymbol{x})\right)^{\top}\left(H_{[k]}\right)^{-1}C_{[k]}(\boldsymbol{y}).$$

The Christoffel–Darboux kernel $K^{(\ell+1)}(\boldsymbol{x},\boldsymbol{y})$ gives the projection, $S_\ell : L^2(\mathbb{R}^D,\mu) \to L^2(\mathbb{R}^D,\mu)$, on the set of MVOPR of degree ℓ or less, given any vector in the real Hilbert space $f \in L^2(\mathbb{R}^D,\mu)$ its orthogonal projection is given by the truncated Fourier series $S_\ell(f)(\boldsymbol{x}) = \int_\Omega K^{(\ell+1)}(\boldsymbol{x},\boldsymbol{y})f(\boldsymbol{y})\mathrm{d}\mu(\boldsymbol{y})$. In fact, if P is a MVOPR of degree ℓ or less then $P = S_\ell(P)$ and $S_\ell \circ S_\ell = S_\ell$; these kernels are subject to the *reproducing property* $K^{(\ell+1)}(\boldsymbol{x},\boldsymbol{y}) = \int_\Omega K^{(\ell+1)}(\boldsymbol{x},\boldsymbol{z})\mathrm{d}\mu(\boldsymbol{z})K^{(\ell+1)}(\boldsymbol{z},\boldsymbol{y})$. This projection is the best approximation to f with MVOPR of degree ℓ or less, in the sense that the mean square distance, $\int_\Omega (f(\boldsymbol{x}) - S_\ell(f)(\boldsymbol{x}))^2\mathrm{d}\mu(\boldsymbol{x})$, is minimized and all the other polynomials of degree ℓ or less have a bigger mean square distance to f. When the space of MVOPR is dense in $L^2(\mathbb{R}^D,\mu)$ then the Fourier series converges in the mean square distance to f, $\lim_{\ell\to\infty}\int_\Omega (f(\boldsymbol{x}) - S_\ell(f)(\boldsymbol{x}))^2\mathrm{d}\mu(\boldsymbol{x}) = 0$. For more information see §3.5 of [44].

From Definition 2.8.1 it directly follows that

Proposition 2.8.1. *In terms of Schur complements or quasi-determinants of bordered truncated moment matrices the Christoffel–Darboux kernel is expressed as*

[9] For $k = 0$ we get

$$(x_a + (\Lambda_a)_{[0],[1]}\beta_{[1]})C_{[0]}(\boldsymbol{x}) = (\Lambda_a)_{[0],[1]}C_{[1]}(\boldsymbol{x}) + \hat{C}_{[0],a}(\boldsymbol{x}).$$

$$K^{[\ell]}(\boldsymbol{x},\boldsymbol{y})=\operatorname{SC}\left(\begin{array}{ccc|c}G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]}(\boldsymbol{y})\\ \hline \chi_{[0]}^{\top}(\boldsymbol{x}) & \cdots & \chi_{[\ell-1]}^{\top}(\boldsymbol{x}) & 0\end{array}\right)$$

$$=\Theta_*\begin{pmatrix}G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]}(\boldsymbol{y})\\ \chi_{[0]}^{\top}(\boldsymbol{x}) & \cdots & \chi_{[\ell-1]}^{\top}(\boldsymbol{x}) & 0\end{pmatrix},$$

while the second kind Christoffel–Darboux kernel has the following one

$$Q^{[\ell]}(\boldsymbol{x},\boldsymbol{y})=\operatorname{SC}\left(\begin{array}{ccc|c}(G^{-1})_{[0],[0]} & \cdots & (G^{-1})_{[0],[\ell-1]} & \chi^*_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ (G^{-1})_{[\ell-1],[0]} & \cdots & (G^{-1})_{[\ell-1],[\ell-1]} & \chi^*_{[\ell-1]}(\boldsymbol{y})\\ \hline (\chi^*_{[0]})^{\top}(\boldsymbol{x}) & \cdots & (\chi^*_{[\ell-1]})^{\top}(\boldsymbol{x}) & 0\end{array}\right)$$

$$=\Theta_*\begin{pmatrix}(G^{-1})_{[0],[0]} & \cdots & (G^{-1})_{[0],[\ell-1]} & \chi^*_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ (G^{-1})_{[\ell-1],[0]} & \cdots & (G^{-1})_{[\ell-1],[\ell-1]} & \chi^*_{[\ell-1]}(\boldsymbol{y})\\ (\chi^*_{[0]})^{\top}(\boldsymbol{x}) & \cdots & (\chi^*_{[\ell-1]})^{\top}(\boldsymbol{x}) & 0\end{pmatrix}.$$

In terms of determinants we have

$$K^{[\ell]}(\boldsymbol{x},\boldsymbol{y})=\frac{\begin{vmatrix}G_{[0],[0]} & \cdots & G_{[0],[\ell-1]} & \chi_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]} & \chi_{[\ell-1]}(\boldsymbol{y})\\ \chi_{[0]}^{\top}(\boldsymbol{x}) & \cdots & \chi_{[\ell-1]}^{\top}(\boldsymbol{x}) & 0\end{vmatrix}}{\begin{vmatrix}G_{[0],[0]} & \cdots & G_{[0],[\ell-1]}\\ \vdots & & \vdots\\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]}\end{vmatrix}},$$

$$Q^{[\ell]}(\boldsymbol{x},\boldsymbol{y})=\begin{vmatrix}(G^{-1})_{[0],[0]} & \cdots & (G^{-1})_{[0],[\ell-1]} & \chi^*_{[0]}(\boldsymbol{y})\\ \vdots & & \vdots & \vdots\\ (G^{-1})_{[\ell-1],[0]} & \cdots & (G^{-1})_{[\ell-1],[\ell-1]} & \chi^*_{[\ell-1]}(\boldsymbol{y})\\ (\chi^*_{[0]})^{\top}(\boldsymbol{x}) & \cdots & (\chi^*_{[\ell-1]})^{\top}(\boldsymbol{x}) & 0\end{vmatrix}$$

$$\cdot\begin{vmatrix}G_{[0],[0]} & \cdots & G_{[0],[\ell-1]}\\ \vdots & & \vdots\\ G_{[\ell-1],[0]} & \cdots & G_{[\ell-1],[\ell-1]}\end{vmatrix}.$$

Theorem 2.8.1. *The following Christoffel–Darboux formula holds*

$$K^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) = \frac{\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}P_{[\ell]}(\boldsymbol{x})\big)^\top\big(H_{[\ell-1]}\big)^{-1}P_{[\ell-1]}(\boldsymbol{y}) - P_{[\ell-1]}(\boldsymbol{x})^\top\big(H_{[\ell-1]}\big)^{-1}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}P_{[\ell]}(\boldsymbol{y})}{\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{y})}, \tag{2.8.1}$$

while for the second kind kernel we have

$$\begin{aligned} Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) &= \frac{\big[(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{x})\big]^\top\big(H_{[\ell-1]}\big)^{-1}\big[C_{[\ell-1]}(\boldsymbol{y})\big] - \big[C_{[\ell-1]}(\boldsymbol{x})\big]^\top\big(H_{[\ell-1]}\big)^{-1}\big[(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{y})\big]}{\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{y})} \\ &+ \frac{\big[\boldsymbol{n}\cdot\widehat{\boldsymbol{C}}^{[\ell]}(\boldsymbol{x})\big]^\top\big(H^{[\ell]}\big)^{-1}C^{[\ell]}(\boldsymbol{y}) - \big[C^{[\ell]}(\boldsymbol{x})\big]^\top\big(H^{[\ell]}\big)^{-1}\boldsymbol{n}\cdot\widehat{\boldsymbol{C}}^{[\ell]}(\boldsymbol{y})}{\boldsymbol{n}\cdot(\boldsymbol{x}-\boldsymbol{y})}. \end{aligned} \tag{2.8.2}$$

Proof. In the first place, for the polynomials, on the one hand we have

$$\begin{aligned} &(P^{[\ell]}(\boldsymbol{x}))^\top(H^{-1}\boldsymbol{n}\cdot\boldsymbol{J})^{[\ell]}P^{[\ell]}(\boldsymbol{y}) \\ &= (\boldsymbol{n}\cdot\boldsymbol{y})(P^{[\ell]}(\boldsymbol{x}))^\top(H^{[\ell]})^{-1}P^{[\ell]}(\boldsymbol{y}) - P_{[l-1]}(\boldsymbol{x})^\top\big(H_{[\ell-1]}\big)^{-1}(\boldsymbol{n}\cdot\boldsymbol{J})_{[\ell-1],[\ell]}P_{[l]}(\boldsymbol{y}), \end{aligned}$$

and on the other hand

$$\begin{aligned} &(P^{[\ell]}(\boldsymbol{x}))^\top(H^{-1}\boldsymbol{n}\cdot\boldsymbol{J})^{[\ell]}P^{[\ell]}(\boldsymbol{y}) \\ &= (\boldsymbol{n}\cdot\boldsymbol{x})(P^{[\ell]}(\boldsymbol{x}))^\top(H^{[\ell]})^{-1}P^{[\ell]}(\boldsymbol{y}) - \big((\boldsymbol{n}\cdot\boldsymbol{J})_{[\ell-1][\ell]}P_{[\ell]}(\boldsymbol{x})\big)^\top(H_{[\ell-1]})^{-1}P_{[\ell-1]}(\boldsymbol{y}). \end{aligned}$$

For the second kind functions we proceed similarly. However, we must take care of the appearing of the reduced second kind functions. In this case the two possibilities are

$$\begin{aligned} &(C^{[\ell]}(\boldsymbol{x}))^\top(H^{-1}\boldsymbol{n}\cdot\boldsymbol{J})^{[\ell]}C^{[\ell]}(y) \\ &\quad= (\boldsymbol{n}\cdot\boldsymbol{y})(C^{[\ell]}(\boldsymbol{x}))^\top(H^{[\ell]})^{-1}C^{[\ell]}(\boldsymbol{y}) - (C^{[\ell]}(\boldsymbol{x}))^\top(H^{[l]})^{-1}(\boldsymbol{n}\cdot\hat{\boldsymbol{C}}^{[\ell]}(\boldsymbol{y})) \\ &\qquad - C_{[\ell-1]}(\boldsymbol{x})^\top(H_{[\ell-1]})^{-1}(\boldsymbol{n}\cdot\boldsymbol{J})_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{y}) \\ &(C^{[\ell]}(\boldsymbol{x}))^\top(H^{-1}\boldsymbol{n}\cdot\boldsymbol{J})^{[\ell]}C^{[\ell]}(y) \\ &\quad= (\boldsymbol{n}\cdot\boldsymbol{x})(C^{[\ell]}(\boldsymbol{x}))^\top(H^{[\ell]})^{-1}C^{[\ell]}(\boldsymbol{y}) - (\boldsymbol{n}\cdot\hat{\boldsymbol{C}}^{[\ell]}(\boldsymbol{x}))^\top(H^{[\ell]})^{-1}C^{[\ell]}(\boldsymbol{y}) \\ &\qquad - ((\boldsymbol{n}\cdot\boldsymbol{J})_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{x}))^\top(H_{[l-1]})^{-1}C_{[\ell-1]}(\boldsymbol{y}). \qquad \square \end{aligned}$$

Observe also that (2.8.2) is not a standard Christoffel–Darboux formula because the last term involves all the reduced second kind functions. These terms are absent in the scalar case but in this multivariant scenario show up.

3. On discrete Toda and MVOPR

In this section we discuss the connection between MVOPR associated with different measures; this collection of measures can be considered as a lattice of measures. The set of transformations which we are about to introduce is not the more general one but capture the essential facts.

3.1. The discrete flows

For the construction of D discrete flows we consider an invertible matrix

$$N = (n_{a,b})_{a,b=1,\dots,D} \in \mathrm{GL}(\mathbb{R}^D),$$

and therefore D linearly independent vectors $\boldsymbol{n}_a = (n_{a,1}, \dots, n_{a,D})^\top$, and a vector $\boldsymbol{q} = (q_1, \dots, q_D)^\top \in \mathbb{R}^D$, where $q_a \neq 0$, $a = \{1, \dots, D\}$. For a given measure $\mathrm{d}\mu$ and each multi-index $\boldsymbol{m} = (m_1, \dots, m_D)^\top \in \mathbb{Z}^D$ we consider the measure

$$\mathrm{d}\mu_{\boldsymbol{m}}(\boldsymbol{x}) = \Big[\prod_{a=1}^{D} \big(\boldsymbol{n}_a \cdot \boldsymbol{x} - q_a\big)^{m_a}\Big] \mathrm{d}\mu(\boldsymbol{x}).$$

Associated with this deformed measure we introduce the set

$$R := \{\boldsymbol{x} \in \mathbb{R}^D : |\boldsymbol{n}_1 \cdot \boldsymbol{x}| < |q_1|, \dots, |\boldsymbol{n}_D \cdot \boldsymbol{x}| < |q_D|\}, \tag{3.1.1}$$

and the related sets $R_a := \{\boldsymbol{x} \in \mathbb{R}^D : -|q_a| < \boldsymbol{n}_a \cdot \boldsymbol{x} < |q_a|\}$, $a \in \{1, \dots, D\}$. Observe that $R = \cap_{a=1}^{D} R_a$ is a bounded open convex polytope included in the ball centered at the origin of radius $\max_{a\in\{1,\dots,D\}} |q_a|$. We see that the border of R_a is $\partial R_a = \pi_a^+ \cup \pi_a^-$ in terms of the hyperplanes $\pi_a^\pm := \{\boldsymbol{x} \in \mathbb{R}^D : \boldsymbol{n}_a \cdot \boldsymbol{x} = \pm q_a\}$. The measure $\mathrm{d}\mu_{\boldsymbol{m}}$ has a definite sign in $R \cap \operatorname{supp}(\mu)$ since the hyperplane $\pi_a^+ : \boldsymbol{n}_a \cdot \boldsymbol{x} = q_a$ belongs to the border and therefore is unreachable in R.

As we will see later on, §3.4, these discrete flows are built up in terms of Darboux transformations. Sometimes the flows described by $m_a \to m_a + 1$ are known as Christoffel transformations and those associated with $m_a \to m_a - 1$ as Geronimus transformations.

A natural question arises here: Which are the corresponding moment matrices? and the answer, given in terms of shift matrices, is fairly nice.

Proposition 3.1.1. *For a given Borel measure μ let us assume that* $\operatorname{supp}\mu \subset R$*, with R given in (3.1.1), then the moment matrices $G(\boldsymbol{m})$ of the measures $\mathrm{d}\mu_{\boldsymbol{m}}$ satisfy*

$$G(\boldsymbol{m}) = \Big(\prod_{a=1}^{D} (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{m_a}\Big) G = G\Big(\prod_{a=1}^{D} (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda}^\top - q_a)^{m_a}\Big). \tag{3.1.2}$$

Proof. We need to be especially careful when m_a is a negative integer because, in that case we are dealing with powers of the inverse matrix of $(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)$, $a \in \{1, \dots, D\}$. The request $q_a \neq 0$ ensures that $(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{-1}$ can be formally given as the following upper triangular matrix

$$(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{-1} = -q_a^{-1} - q_a^{-2}(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})^2 - q_a^{-3}(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})^2 - \cdots .$$

This series is a matrix organized by superdiagonals with $\big((\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{-1}\big)_{[k],[k+j]} = -q_a^{-(j+1)}((\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})^j)_{[k],[k+j]}$ the j-th block in the k-th superdiagonal, and no series is involved for a given block; i.e., the expression is not only formal but it is well defined. However, we should also tackle the more subtle problem of the domain of these matrices. In particular, its action on χ gives $(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})\chi(\boldsymbol{x}) = (\boldsymbol{n}_a \cdot \boldsymbol{x})\chi(\boldsymbol{x})$ and corresponding series is $(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{-1}\chi = -q_a^{-1} - q_a^{-2}(\boldsymbol{n}_a \cdot \boldsymbol{x}) - q_a^{-3}(\boldsymbol{n}_a \cdot \boldsymbol{x})^2 - \cdots$ which converges for $|\boldsymbol{n}_a \cdot \boldsymbol{x}| < |q_a|$. Recalling Proposition 2.6.1 the result follows. □

Definition 3.1.1.

(1) We introduce

$$W_0(\boldsymbol{m}) := \prod_{a=1}^{D} (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)^{m_a}.$$

(2) The action of the translation T_a on any function f on $\mathbb{Z}^D$ is defined by

$$(T_a f)(m_1, \dots, m_a, \dots, m_D) = f(m_1, \dots, m_a + 1, \dots, m_D),$$

and the partial difference operator is given by

$$\Delta_a := T_a - 1.$$

These translations depend on $N, \boldsymbol{q}$ and when needed we use the notation $T_a^{(N,\boldsymbol{q})}$ or $T_a^{(\boldsymbol{q})}$ to indicate it.

(3) The MVOPR and the corresponding second kind functions associated with $\mathrm{d}\mu_{\boldsymbol{m}}(\boldsymbol{x})$ will be denoted by $P(\boldsymbol{m}, \boldsymbol{x})$ and $C(\boldsymbol{m}, \boldsymbol{x})$.

(4) Assuming the block Cholesky factorization for[10] $G(\boldsymbol{m})$, $G(\boldsymbol{m}) = (S(\boldsymbol{m}))^{-1} H(\boldsymbol{m}) \times \big(S(\boldsymbol{m})^{-1}\big)^\top$, for each $\boldsymbol{m} \in \mathbb{Z}^D$ we introduce the following semi-infinite matrices

$$M_a(\boldsymbol{m}) := S(\boldsymbol{m})\big((T_a S)(\boldsymbol{m})\big)^{-1}. \tag{3.1.3}$$

[10] With $S(\boldsymbol{m})$ block lower triangular with the block diagonal populated by identities, and $H(\boldsymbol{m})$ block diagonal.

For the sake of simplicity, from hereon and when not needed we will omit writing the $\boldsymbol{m}$-dependence and it will be implicitly assumed.

Proposition 3.1.2. *The moment matrix satisfies*

$$T_aG = (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)G = G((\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})^\top - q_a).$$

Proof. We observe that (3.1.2) could be written as

$$G(\boldsymbol{m}) = W_0(\boldsymbol{m})G = G\big(W_0(\boldsymbol{m})\big)^\top. \qquad \square$$

Proposition 3.1.3. *The matrix $M_a = S(T_aS)^{-1}$ fulfills*

$$M_a = H\big((T_aS)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a)S^{-1}\big)^\top (T_aH)^{-1}. \tag{3.1.4}$$

Moreover, M_a is a block lower unitriangular matrix with only the first subdiagonal different from zero; i.e., $M_a = \mathbb{I} + \rho_a$ with

$$\begin{aligned}\rho_a &= H(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})^\top (T_aH)^{-1} &(3.1.5)\\ &= -\Delta_a\beta. &(3.1.6)\end{aligned}$$

Proof. For (3.1.4) introduce the Cholesky factorization in the second equality in (3.1.2), this equation has as direct consequence (3.1.5). The relation (3.1.6) follows from (3.1.3) and (3.1.4). $\square$

Componentwise we have

$$\rho_a := \begin{pmatrix} 0 & 0 & 0 & \cdots \\ \rho_{a,[1]} & 0 & 0 & \cdots \\ 0 & \rho_{a,[2]} & 0 & \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

with

$$\begin{aligned}\rho_{a,[k]} &:= H_{[k]}[(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]}]^\top \big(T_aH_{[k-1]}\big)^{-1} &(3.1.7)\\ &= -\Delta_a\beta_{[k]}. &(3.1.8)\end{aligned}$$

Definition 3.1.2. We introduce the wave matrices[11]

$$W_1 := SW_0, \qquad W_2 := H\big(S^{-1}\big)^\top,$$

and the lattice resolvents

[11] These definitions are motivated by the relation $W_1G = W_2$.

$$\omega_a := (T_a H) M_a^\top H^{-1}, \quad a \in \{1, \dots, D\}. \tag{3.1.9}$$

Proposition 3.1.4. *The evolved wave functions W_1 and W_2 satisfy the following*

$$G = W_1(\boldsymbol{m})^{-1} W_2(\boldsymbol{m}). \tag{3.1.10}$$

Proof. From the Cholesky factorization we deduce that

$$\begin{aligned} G &= W_0(\boldsymbol{m})^{-1} (S(\boldsymbol{m}))^{-1} H(\boldsymbol{m}) \big((S(\boldsymbol{m}))^{-1}\big)^\top \\ &= W_1(\boldsymbol{m})^{-1} W_2(\boldsymbol{m}). \qquad \square \end{aligned} \tag{3.1.11}$$

For the lattice resolvent we have

Proposition 3.1.5. *The lattice resolvent can be expressed as*

$$\omega_a = (T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1}. \tag{3.1.12}$$

Moreover, we have the explicit form

$$\omega_a = \alpha_a + \boldsymbol{n}_a \cdot \boldsymbol{\Lambda}$$

where the diagonal terms have the following alternative expressions

$$\alpha_a = (T_a H) H^{-1} \tag{3.1.13}$$

$$= (T_a \beta)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda}) - (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})\beta - q_a. \tag{3.1.14}$$

Componentwise we have

$$\alpha_{a,[k]} = (T_a H_{[k]}) H_{[k]}^{-1} \tag{3.1.15}$$

$$= (T_a \beta_{[k]})(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k],[k+1]} \beta_{[k+1]} - q_a. \tag{3.1.16}$$

Proof. The first relation is a consequence of (3.1.4) and (3.1.9), then (3.1.12) and (3.1.13) are consequences of (3.1.12) and (3.1.9). Finally, from (3.1.12) and (3.1.13) we infer (3.1.14). $\square$

As byproduct we get

Proposition 3.1.6. *The following equations hold*

$$(T_a H_{[k]}) H_{[k]}^{-1} = (T_a \beta_{[k]})(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k],[k+1]} \beta_{[k+1]} - q_a,$$

$$H_{[k]} [(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]}]^\top \big(T_a H_{[k-1]}\big)^{-1} = -\Delta_a \beta_{[k]}.$$

A set of Toda type equations can be derived.

Theorem 3.1.1. *The quasi-tau matrices $H_{[k]}$ are subject to the following equations*

$$\Delta_b\big((\Delta_a H_{[k]})H_{[k]}^{-1}\big) = (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k],[k+1]}H_{[k+1]}[(\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k],[k+1]}]^\top \big(T_b H_{[k]}\big)^{-1} \\ - (T_a H_{[k]})[(\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k-1],[k]}]^\top \big(T_a T_b H_{[k-1]}\big)^{-1}(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]}.$$

The matrices $\beta_{[k]}$ fulfill

$$\big((T_b\beta_{[k]})(\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k],[k+1]}\beta_{[k+1]} - q_b\big)(\Delta_a\beta_{[k]}) \\ = T_b\Delta_a\beta_{[k]}\big((T_aT_b\beta_{[k]})(\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}_b \cdot \boldsymbol{\Lambda})_{[k],[k+1]}(T_a\beta_{[k+1]}) - q_b\big).$$

Proof. An immediate consequence of Proposition 3.1.6. □

Let us stress that the matrices in the nonlinear lattice have increasing sizes.

Proposition 3.1.7. *If T_aG and G admit Cholesky decompositions then for each $a \in \{1,\dots,D\}$ we have the following LU factorization*

$$\boldsymbol{n}_a \cdot \boldsymbol{J} - q_a = M_a\omega_a, \tag{3.1.17}$$

and the UL factorization

$$T_a(\boldsymbol{n}_a \cdot \boldsymbol{J}) - q_a = \omega_a M_a. \tag{3.1.18}$$

Proof. From Proposition 3.1.2 we get

$$\begin{aligned} T_aG &= (T_aS)^{-1}(T_aH)\big((T_aS)^{-1}\big)^\top && \text{Cholesky factorization of } (T_aG)\\ &= (\boldsymbol{n}_a \cdot \boldsymbol{J} - q_a)G && \text{see Proposition 3.1.2}\\ &= (\boldsymbol{n}_a \cdot \boldsymbol{J} - q_a)S^{-1}H\big(S^{-1}\big)^\top && \text{Cholesky factorization of } G \end{aligned}$$

and therefore we have the Cholesky factorization (3.1.17). To prove (3.1.18) we observe that the Lax equations (3.1.20) lead to

$$\begin{aligned} T_a(\boldsymbol{n}_a \cdot \boldsymbol{J} - q_a)\omega_a &= \omega_a(\boldsymbol{n}_a \cdot \boldsymbol{J} - q_a)\\ &= \omega_a M_a\omega_a \end{aligned}$$

which imply the result. □

Thus, for each given direction these translations reproduce the behavior of the classical elementary Darboux transformations which imply the interchange or intertwining of the lower triangular and upper triangular factors in the Gauss–Borel decomposition. Later on we will discuss the explicit form of these Darboux transformations for the MVOPR, quasi-tau and β matrices.

Before we derived (3.1.7) and (3.1.15) where we expressed $\rho_{a,[k]}$, (3.1.8), (3.1.16) and $\alpha_{a,[k]}$ in terms of H and β matrices and its discrete time translations. Now, we show an alternative form of writing these functions, with no discrete time translations involved, in terms of quasi-determinants of truncated Jacobi matrices.

In (3.1.7) and (3.1.15) we expressed $\rho_{a,[k]}$ and $\alpha_{a,[k]}$ in terms of H's and its discrete time translations, and in (3.1.8) and (3.1.16) we gave alternative expressions in terms of β's and its discrete time translations. Now, we find expressions, in terms of quasi-determinants of truncated Jacobi matrices, that do not require of discrete time translations.

Theorem 3.1.2. *In terms of quasi-determinants of the Jacobi matrices we have the following formulae*

$$\begin{aligned}\rho_{a,[k]} &= (\boldsymbol{n}_a\cdot\boldsymbol{J}_{[k],[k-1]})\big(\Theta_*(\boldsymbol{n}_a\cdot\boldsymbol{J}^{[k]}-q_a\mathbb{I}^{[k]})\big)^{-1},\\ \alpha_{a,[k]} &= \Theta_*(\boldsymbol{n}_a\cdot\boldsymbol{J}^{[k+1]}-q_a\mathbb{I}^{[k+1]}).\end{aligned}$$

Proof. It is a direct consequence of Proposition 3.1.7 and Theorem 3.4 of [96]. □

In particular we deduce that

$$\rho_{a,[k]} = H_{[k]}(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}_{[k-1],[k]})^\top H_{[k-1]}\big(\Theta_*(\boldsymbol{n}_a\cdot\boldsymbol{J}^{[k]}-q_a\mathbb{I}^{[k]})\big)^{-1}.$$

Next, we are ready to collect the integrable system structure for these discrete flows giving the classical elements: linear systems, Lax and Zakharov–Shabat equations in its discrete version.

Proposition 3.1.8.

(1) *For each $a\in\{1,\dots,D\}$ the wave matrices W_1 and W_2 are solutions of the following linear system*

$$T_aW=\omega_aW. \tag{3.1.19}$$

(2) *The Jacobi type matrices $\boldsymbol{n}_a\cdot\boldsymbol{J}$, $a\in\{1,\dots,D\}$, satisfy the discrete Lax equations*

$$T_b(\boldsymbol{n}_a\cdot\boldsymbol{J})\omega_b=\omega_b\,(\boldsymbol{n}_a\cdot\boldsymbol{J}),\quad M_b\,T_b(\boldsymbol{n}_a\cdot\boldsymbol{J})=(\boldsymbol{n}_a\cdot\boldsymbol{J})M_b,\quad \forall a,b\in\{1,\dots,D\}. \tag{3.1.20}$$

(3) *The lattice resolvent ω_a and the M_a, $a\in\{1,\dots,D\}$, are subject to the discrete Zakharov–Shabat equations*

$$(T_a\omega_b)\omega_a=(T_b\omega_a)\omega_b,\quad M_a(T_aM_b)=M_b(T_bM_a),\quad \forall a,b\in\{1,\dots,D\}. \tag{3.1.21}$$

Proof. See Appendix D.5. $\square$

3.2. Quasi-tau functions formulae for MVOPR

3.2.1. Expressing the MVOPR in terms of quasi-tau matrices

We discuss here certain expressions for the MVOPR, in terms of the quasi-tau matrices $H_{[k]}$, that extend to the multidimensional situation the τ type formulae of the 1D scenario.

Proposition 3.2.1. *The MVOPR satisfy*

$$\rho_{a,[k]}(T_aP)_{[k-1]}+(T_aP)_{[k]}=P_{[k]}, \tag{3.2.1}$$

$$\alpha_{a,[k-1]}P_{[k-1]}+(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[k-1],[k]}P_{[k]}=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aP)_{[k-1]}, \tag{3.2.2}$$

Proof. It is just a consequence of $P=M_a(T_aP)$ and $\omega_aP=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aP)$. $\square$

Proposition 3.2.2. *When $\boldsymbol{p}\in\pi_a^+$, i.e. $\boldsymbol{n}_a\cdot\boldsymbol{p}=q_a$, the following relation holds*

$$(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[k-1],[k]}P_{[k]}(\boldsymbol{p})=-\alpha_{a,[k-1]}P_{[k-1]}(\boldsymbol{p}). \tag{3.2.3}$$

Proof. Set $\boldsymbol{n}_a\cdot\boldsymbol{x}=q_a$ in (3.2.2). $\square$

Definition 3.2.1. Together with $\boldsymbol{q}:=(q_1,\dots,q_d)$ we consider the following two rectangular matrices

$$[N\boldsymbol{\Lambda}]_k:=\begin{pmatrix}(\boldsymbol{n}_1\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\\ \vdots\\ (\boldsymbol{n}_D\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\end{pmatrix}\in\mathbb{R}^{D|[k]|\times|[k+1]|},$$

$$[\boldsymbol{T}^{(\boldsymbol{q})}H]_k:=\begin{pmatrix}T_1^{(\boldsymbol{q})}H_{[k]}\\ \vdots\\ T_D^{(\boldsymbol{q})}H_{[k]}\end{pmatrix}\in\mathbb{R}^{D|[k]|\times|[k]|}.$$

Observe that, putting together as rows the blocks $(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[k],[k+1]}$, for $a\in\{1,\dots,D\}$, we get a full column rank matrix $[N\boldsymbol{\Lambda}]_k$. Hence, the correlation matrix $[N\boldsymbol{\Lambda}]_k^\top[N\boldsymbol{\Lambda}]_k\in\mathbb{R}^{|[k+1]|\times|[k+1]|}$ is invertible and the pseudo-inverse is

$$[N\boldsymbol{\Lambda}]_k^+=\big([N\boldsymbol{\Lambda}]_k^\top[N\boldsymbol{\Lambda}]_k\big)^{-1}[N\boldsymbol{\Lambda}]_k^\top,$$

which happens to be a left inverse, see Appendix B.1.

We are now ready for

Theorem 3.2.1. *The MVOPR can be expressed in terms of quasi-tau matrices H and its discrete time translations as follows*

$$P_{[k]}(\boldsymbol{q}) = (-1)^k[N\boldsymbol{\Lambda}]^+_{k-1}[\boldsymbol{T}^{(N\boldsymbol{q})}H]_{k-1}(H_{[k-1]})^{-1}[N\boldsymbol{\Lambda}]^+_{k-2}[\boldsymbol{T}^{(N\boldsymbol{q})}H]_{k-2}$$
$$\times (H_{[k-2]})^{-1}\cdots[N\boldsymbol{\Lambda}]^+_0[\boldsymbol{T}^{(N\boldsymbol{q})}H]_0H^{-1}_{[0]}.$$

Proof. Proposition 3.2.2 takes an interesting form if we choose $\boldsymbol{p} \in \cap_{a=1}^D \pi_a^+$; i.e., $\boldsymbol{n}_a \cdot \boldsymbol{p} = q_a$, which means that $N\boldsymbol{p} = \boldsymbol{q}$ and as N is invertible we have $\boldsymbol{p} = N^{-1}\boldsymbol{q}$. In this case (3.2.3) takes the form

$$[N\boldsymbol{\Lambda}]_{k-1}P_{[k]}(N^{-1}\boldsymbol{q}) = -[\boldsymbol{T}^{(\boldsymbol{q})}H]_{k-1}H^{-1}_{[k-1]}P_{[k-1]}(N^{-1}\boldsymbol{q}).$$

But, since $[N\boldsymbol{\Lambda}]_{k-1}$ has full column rank k we can take its Moore–Penrose pseudo-inverse $[N\boldsymbol{\Lambda}]^+_{k-1}$ to get

$$P_{[k]}(N^{-1}\boldsymbol{q}) = -[N\boldsymbol{\Lambda}]^+_{(k-1)}[\boldsymbol{T}^{(\boldsymbol{q})}H]_{k-1}H^{-1}_{[k-1]}P_{[k-1]}(N^{-1}\boldsymbol{q}).$$

Iterating this relation we get the desired result. □

In the one-dimensional case we only have one component and the block matrices are just numbers and we should replace $(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k-1],[k]}$ by 1. Thus, for $D = 1$ the Theorem 3.2.1 gives

$$P_k(q) = (-1)^kTH_{k-1}(H_{k-1})^{-1}TH_{k-2}(H_{k-2})^{-1}\cdots TH_0H_0^{-1},$$

so that

$$\begin{aligned} P_k(q) &= (-1)^k\frac{TH_{k-1}TH_{k-2}\cdots TH_0}{H_{k-1}H_{k-2}\cdots H_0} \\ &= (-1)^k\frac{T\tau_k}{\tau_k}, \qquad\qquad \tau_k := \det G^{[k-1]} = H_{k-1}H_{k-2}\cdots H_0. \end{aligned}$$

This is a well known expression in terms of Miwa shifts of τ-functions, see §4.3, where the τ-function is the determinant of the OPRL moment matrix

3.2.2. Quasi-tau matrix expressions for the second kind functions

Proposition 3.2.3. *The second kind functions satisfy*

$$\rho_{a,[k]}(T_aC)_{[k-1]} + (T_aC)_{[k]} = (\boldsymbol{n}_a \cdot \boldsymbol{x} - q_a)C_{[k]} - \boldsymbol{n}_a \cdot \widehat{\boldsymbol{C}}_{[k]}, \tag{3.2.4}$$
$$\alpha_{a,[k]}C_{[k]} + (\boldsymbol{n}_a \cdot \boldsymbol{\Lambda})_{[k],[k+1]}C_{[k+1]} = (T_aC)_{[k]}. \tag{3.2.5}$$

Proof. See Appendix D.6. □

From now on in this subsection we take $N = \mathbb{I}_D$.

Definition 3.2.2. Let us introduce the composed or total translation

$$T=\prod_{a=1}^{D}T_a, \qquad M:=S\big(TS\big)^{-1}.$$

Observe that $P=M(TP)$. From Cholesky factorization (2.3.2) we get

Proposition 3.2.4. *The following relation holds*

$$M=(H)\big(S^{-1}\big)^{\top}\Big[\prod_{a=1}^{D}(\Lambda_a-q_a)\Big]^{\top}(TS)^{\top}(TH)^{-1}. \tag{3.2.6}$$

This lower unitriangular banded (with D subdiagonals) block matrix can be decomposed into the product of D lower unitriangular with only the diagonal and first subdiagonal different from zero

Proposition 3.2.5. *For each permutation σ in the symmetric group $\mathfrak{S}_D$ the following decomposition holds true*

$$M=M_{\sigma(1)}\big(T_{\sigma 1}M_{\sigma 2}\big)\cdots\big(T_{\sigma 1}\cdots T_{\sigma(D-1)}M_{\sigma D}\big). \tag{3.2.7}$$

Proof. From Definition 3.1.3 for each factor in the RHS of (3.2.7) we have

$$\begin{aligned}
M_{\sigma 1}&=S\big(T_{\sigma 1}S\big)^{-1},\\
T_{\sigma 1}M_{\sigma 2}&=\big(T_{\sigma 1}S\big)\big(T_{\sigma 1}T_{\sigma 2}S\big)^{-1},\\
&\vdots\\
T_{\sigma(1)}\cdots T_{\sigma(D-1)}M_{\sigma(D)}&=\big(T_{\sigma 1}\cdots T_{\sigma(D-1)}S\big)\big(T_{\sigma 1}\cdots T_{\sigma D}S\big)^{-1},
\end{aligned}$$

and the result follows. □

Definition 3.2.3. We introduce

$$\rho^{(a)}_{[k]}:=\Big(\prod_{j=a}^{D}T_j^{-1}\Big)\big(\rho_{a,[k]}\big)\in\mathbb{R}^{k\times(k-1)}. \tag{3.2.8}$$

Matrices can be expressed as

Proposition 3.2.6. *We have*

$$\rho^{(a)}_{[k]}:=\Big(\prod_{j=a}^{D}T_j^{-1}H_{[k]}\Big)(\Lambda_a)^{\top}_{[k-1],[k]}\Big(\prod_{j=a+1}^{D}T_j^{-1}H_{[k-1]}\Big)^{-1}. \tag{3.2.9}$$

Proof. It follows immediately when we substitute (3.1.7) into (3.2.8). □

For second kind functions, the analogous result to Theorem 3.2.1 is

Theorem 3.2.2. *The second kind functions can be written as*

$$C_{[k]}(\boldsymbol{q}) = (-1)^{k+D} \sum_{1\leq a_1\leq\cdots\leq a_k\leq D} \rho^{(a_k)}_{[k]}\cdots\rho^{(a_1)}_{[1]} T^{-1}H_{[0]}$$

or in terms of quasi-tau matrices and their inverse translations as follows

$$\begin{aligned}C_{[k]}(\boldsymbol{q}) = (-1)^{k+D} \sum_{1\leq a_1\leq\cdots\leq a_k\leq D} &\Big(\prod_{j=a_k}^{D} T_j^{-1}H_{[k]}\Big)(\Lambda_{a_k})^{\top}_{[k-1],[k]}\Big(\prod_{j=a_k+1}^{D} T_j^{-1}H_{[k-1]}\Big)^{-1}\cdots\\ &\times\Big(\prod_{j=a_1}^{D} T_j^{-1}H_{[1]}\Big)(\Lambda_{a_1})^{\top}_{[0],[1]}\Big(\prod_{j=a_1+1}^{D} T_j^{-1}H_{[0]}\Big)^{-1}T^{-1}H_{[0]}.\end{aligned} \tag{3.2.10}$$

Proof. See Appendix D.7. □

Observe that for $D=1$ the formula (3.2.10) reads

$$\begin{aligned}C_{[k]}(q) &= (-1)^{k+1}(T^{-1}H_{[k]})H_{[k-1]})^{-1}\cdots(T^{-1}H_{[1]})H_{[0]})^{-1}T^{-1}H_{[0]}\\ &= (-1)^{k+1}\frac{T^{-1}(H_{[k]}\cdots H_{[0]})}{H_{[k-1]}\cdots H_{[0]}}\\ &= (-1)^{k+1}\frac{T^{-1}\tau_k}{\tau_{k-1}},\end{aligned}$$

for $\tau_k = \det G^{[k]}$, which is the well known formula that expresses the adjoint Baker functions in terms of τ-functions and Miwa shifts. For $D=2$ we have $T=T_1T_2$ and $(T^{-1}M)^{-1} = (T_2^{-1}M_2)^{-1}(T^{-1}M_1)^{-1}$

$$\begin{aligned}C_{[k]}(\boldsymbol{q}) = (-1)^k\Big(&(T_2^{-1}H_{[k]})(\Lambda_1)^{\top}_{[k-1],[k]}(H_{[k-1]})^{-1}\cdots(T_2^{-1}H_{[1]})(\Lambda_1)^{\top}_{[0],[1]}(H_{[0]})^{-1}\\ &+(T_2^{-1}H_{[k]})(\Lambda_1)^{\top}_{[k-1],[k]}(H_{[k-1]})^{-1}\\ &\cdots(T_2^{-1}H_{[2]})(\Lambda_1)^{\top}_{[1],[2]}(H_{[1]})^{-1}(T^{-1}H_{[1]})(\Lambda_2)^{\top}_{[0],[1]}(T_2^{-1}H_{[0]})^{-1}\\ &+(T_2^{-1}H_{[k]})(\Lambda_1)^{\top}_{[k-1],[k]}(H_{[k-1]})^{-1}\\ &\cdots(T^{-1}H_{[2]})(\Lambda_2)^{\top}_{[1],[2]}(T_2^{-1}H_{[1]})^{-1}(T^{-1}H_{[1]})(\Lambda_2)^{\top}_{[0],[1]}(T_2^{-1}H_{[0]})^{-1}\\ &+(T^{-1}H_{[k]})(\Lambda_2)^{\top}_{[k-1],[k]}(T_2^{-1}H_{[k-1]})^{-1}\\ &\cdots(T^{-1}H_{[2]})(\Lambda_2)^{\top}_{[1],[2]}(T_2^{-1}H_{[1]})^{-1}(T^{-1}H_{[1]})(\Lambda_2)^{\top}_{[0],[1]}(T_2^{-1}H_{[0]})^{-1}\Big)T^{-1}H_{[0]}.\end{aligned}$$

3.3. Transforming the Christoffel–Darboux kernels and kernel polynomials

We give here some relations among translated and nontranslated Christoffel–Darboux kernels; we begin with the following result for the kernels $K^{(\ell)}(\boldsymbol{x},\boldsymbol{y})$ and the MVOPR.

Theorem 3.3.1. *The translated and nontranslated Christoffel–Darboux kernels are connected by*

$$K^{(\ell)}(\boldsymbol{x},\boldsymbol{y})=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aK)^{(\ell-1)}(\boldsymbol{x},\boldsymbol{y})+P_{[\ell-1]}(\boldsymbol{x})^\top(H_{[\ell-1]})^{-1}(T_aP)_{[\ell-1]}(\boldsymbol{y}). \tag{3.3.1}$$

Proof. See Appendix D.8. □

For the second kind kernels $Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y})$ and second kind functions $C(\boldsymbol{x})$ we have

Proposition 3.3.1. *The transformed and initial second kind Christoffel–Darboux kernels are connected by*

$$\begin{aligned}(T_aQ)^{(\ell)}&(\boldsymbol{x},\boldsymbol{y})-(T_aQ)^{(\ell)}(\boldsymbol{x},\boldsymbol{y}_a)\\&=(y_a-q_a)Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y})+\left[C_{[\ell]}(\boldsymbol{x})\right]^\top\left[(\Lambda_a)_{[\ell-1][\ell]}\right]^\top(T_aH)^{-1}_{[\ell-1]}\\&\quad\times\left[(T_aC)_{[\ell-1]}(\boldsymbol{y})-(T_aC)_{[\ell-1]}(\boldsymbol{y}_a)\right].\end{aligned} \tag{3.3.2}$$

Moreover, these kernels fulfill

$$\begin{aligned}&Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y})+(T_aQ)^{(\ell)}(\boldsymbol{x}_a,\boldsymbol{y})-(T_aQ)^{(\ell)}(\boldsymbol{x},\boldsymbol{y}_a)=\\&=\frac{\left[(\Lambda_a)_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{x})\right]^\top\left[\left(H_{[\ell-1]}\right)^{-1}[C_{[\ell-1]}(\boldsymbol{y})]-\left((T_aH)_{[\ell-1]}\right)^{-1}[(T_aC)_{[\ell-1]}(\boldsymbol{y}_a)]\right]}{x_a-y_a}\\&\quad-\frac{\left[\left(H_{[\ell-1]}\right)^{-1}C_{[\ell-1]}(\boldsymbol{x})-\left((T_aH)_{[\ell-1]}\right)^{-1}(T_aC)_{[\ell-1]}(\boldsymbol{x}_a)\right]^\top[(\Lambda_a)_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{y})]}{x_a-y_a}.\end{aligned}$$

Proof. Apply T_a to

$$\frac{(T_aC)(\boldsymbol{x})-(T_aC)(\boldsymbol{x}_a)}{x_a-q_a}=\left(M_a\right)^{-1}C(\boldsymbol{x}) \tag{3.3.3}$$

and use (2.8.2) in

$$\begin{aligned}&Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y})\\&=\frac{[(\Lambda_a)_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{x})]^\top\left(H_{[\ell-1]}\right)^{-1}[C_{[\ell-1]}(\boldsymbol{y})]-[C_{[\ell-1]}(\boldsymbol{x})]^\top\left(H_{[\ell-1]}\right)^{-1}[(\Lambda_a)_{[\ell-1],[\ell]}C_{[\ell]}(\boldsymbol{y})]}{x_a-y_a}\\&\quad+\frac{[-M_a(T_aC)^{[\ell]}(\boldsymbol{x}_a)]^\top\left(H^{[\ell]}\right)^{-1}C^{[\ell]}(\boldsymbol{y})-\left[C^{[\ell]}(\boldsymbol{x})\right]^\top\left(H^{[\ell]}\right)^{-1}[-M_a(T_aC)^{[\ell]}(\boldsymbol{y}_a)]}{x_a-y_a},\end{aligned}$$

but now we can let the M_a act on the $C^{[\ell]}$ instead of on the $(T_a C)^{[\ell]}$ and this way we obtain the result. □

3.4. Elementary Darboux transformations and the sample matrix trick

Darboux transformations were introduced in [36] in the context of the Sturm–Liouville theory and since then have been applied in several problems. It was in [89], a paper where explicit solutions of the Toda lattice where found, where this covariant transformation was given the name of *Darboux*. It has been used in the 1D realm of orthogonal polynomials quite successfully, see for example [127,28,29,88]. In Geometry the theory of transformations of surfaces preserving some given properties conforms a classical subject, in the list of such transformations given in the classical treatise by Eisenhart [45] we find the Levy transformation, which later on was named as elementary Darboux transformation and known in the orthogonal polynomials context as Christoffel transformation [127,109]; in this paper we have denoted it by T. The adjoint elementary Darboux or adjoint Levy transformation T^{-1} is also relevant [89,42] and is sometimes referred to as a Geronimus transformation [127], and in the notation of this paper corresponds to T^{-1}. For further information see [103,62]. In order to extend it to the multivariate realm let us recall some basic facts about the 1D case and then extend it to an arbitrary number of dimensions.

3.4.1. The 1D context. Elementary Darboux transform

For $D = 1$ (3.2.3) reads

$$P_k(q) = -\alpha_k P_{k-1}(q)$$

and as we are dealing with numbers we deduce

$$\alpha_k = -\frac{P_k(q)}{P_{k-1}(q)}$$

that reintroduced in the $D = 1$ version of (3.2.2) gives the so-called kernel polynomials [127],

$$TP_{k-1}(x) = \frac{P_k(x)P_{k-1}(q) - P_k(q)P_{k-1}(x)}{x-q}\frac{1}{P_{k-1}(q)} = K^{(k)}(x,q)\frac{H_k}{P_{k-1}(q)} \tag{3.4.1}$$

which is the standard elementary Darboux transformation for the OPRL. From $\alpha_k = (TH_k)H_k^{-1}$ we get

$$(T_k H_k)P_{k-1}(q) = -P_k(q)H_k.$$

Notice that we can recover this relation directly from the $D = 1$ version of (3.3.1) evaluated at $x = q$

$$K^{(\ell)}(q, y) = -P_\ell(q)(TH_{\ell-1})^{-1}(TP)_{\ell-1}(y).$$

That is, according to (3.4.1), the transformed polynomials are intimately related to the Christoffel–Darboux kernel; this motivates that the polynomials TP_k are sometimes known as kernel polynomials [127].

3.4.2. The multivariate elementary Darboux transformation

A nice property of (3.4.1) is that the Darboux transformed OPRL are expressed explicitly in terms of objects related to the OPRL associated with the original measure. This is apparently lost in the multivariate situation as, despite equation (3.2.2) gives new MVOPR associated with the shifted measure $T_a \mathrm{d}\mu_{\boldsymbol{m}}$ in terms of the MVOPR for the measure $\mathrm{d}\mu_{\boldsymbol{m}}$, now the equivalent relation (3.2.3) does not allow to express $\alpha_{a,[k]}$ in terms of MVOPR for the original measure solely. We could use Theorem 3.1.2 which involves no translations, however it is expressed in terms of quasi-determinants of the Jacobi type matrix and not in terms of the MVOPR. We will show a way to overcome this problem.

We begin with the elementary multivariate Darboux transformation associated with $\boldsymbol{n} \in \mathbb{R}^D$ and $q \in \mathbb{R}$. For that aim we are going to describe what we call the sample matrix trick.

Definition 3.4.1. Given the set $\{\boldsymbol{p}_1, \dots \boldsymbol{p}_{|[k]|}\} \subset \pi^+ = \{\boldsymbol{x} \in \mathbb{R}^D : \boldsymbol{x} \cdot \boldsymbol{n} = q\} \subset \mathbb{R}^D$, whose elements are known as *nodes*, we consider the *sample matrices*

$$\Sigma^k_{[\ell]} = \begin{pmatrix} P_{[\ell]}(\boldsymbol{p}_1) & \dots & P_{[\ell]}(\boldsymbol{p}_{|[k]|}) \end{pmatrix} \in \mathbb{R}^{|[\ell]| \times |[k]|}.$$

The set $\{\boldsymbol{p}_1, \dots \boldsymbol{p}_{|[k]|}\}$ of nodes is said to be a *poised set* for the *interpolation polynomials* $\{P_{\boldsymbol{k}_a}\}_{a=1}^{|[k]|}$ if the sample matrix $\Sigma^k_{[k]}$ is invertible, i.e. $\det \Sigma^k_{[k]} \neq 0$.

We now consider the transformation generated by the discrete flow $T\mathrm{d}\mu(x) = (\boldsymbol{n} \cdot \boldsymbol{x} - q)\mathrm{d}\mu(x)$. An important observation is that the matrix $\alpha_{[k]}$ can be expressed in terms of sample matrices of MVOPR, this is the *sample matrix trick*.

Proposition 3.4.1. *For a poised set $\{\boldsymbol{p}_1, \dots, \boldsymbol{p}_{|[k]|}\} \subset \pi^+ \subset \mathbb{R}^D$ of nodes we can write*

$$\alpha_{[k]} = -(\boldsymbol{n} \cdot \boldsymbol{\Lambda})_{[k],[k+1]} \Sigma^k_{[k+1]} \big(\Sigma^k_{[k]}\big)^{-1}. \tag{3.4.2}$$

Proof. From (3.2.3) we get

$$\alpha_{[k]} P_{[k]}(\boldsymbol{p}_i) = -(\boldsymbol{n} \cdot \boldsymbol{\Lambda})_{[k],[k+1]} P_{[k+1]}(\boldsymbol{p}_i), \qquad i = 1, \dots, |[k]|,$$

or

$$\begin{aligned}&\alpha_{[k]}\left(P_{[k]}(\boldsymbol{p}_1)\quad\dots\quad P_{[k]}(\boldsymbol{p}_{|[k]|})\right)\\&\quad=-\left((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}P_{[k+1]}(\boldsymbol{p}_1)\quad\dots\quad(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}P_{[k+1]}(\boldsymbol{p}_{|[k]|})\right),\end{aligned}$$

and as we are dealing with a poised set of nodes we get the result. □

Hence, we have a set of nodes $\{\boldsymbol{p}_1,\dots,\boldsymbol{p}_{|[k]|}\}\subset\pi^+\subset\mathbb{R}^D$ and a set of interpolation data, $-(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]}$, so that the linear combination $\phi(\boldsymbol{x})=\alpha_{[k]}P_{[k]}(\boldsymbol{x})$, the interpolation function, passes through the interpolation points; i.e., $\phi(\boldsymbol{p}_j)=-(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}P_{[k+1]}(\boldsymbol{p}_j)$.

Now, we are ready to give the elementary multivariate Darboux transformations for MVOPR

Theorem 3.4.1. *Given a poised set $\{\boldsymbol{p}_1,\dots\boldsymbol{p}_{|[k]|}\}\subset\pi^+\subset\mathbb{R}^D$ of nodes we have the following expressions of the elementary Darboux transformed MVOPR, the kernel polynomials $TP(\boldsymbol{x})$ associated with $(\boldsymbol{n}\cdot\boldsymbol{x}-q)\mathrm{d}\mu(\boldsymbol{x})$, in terms of quasi-determinants of the original MVOPR*

$$(TP)_{[k]}(\boldsymbol{x})=(\boldsymbol{n}\cdot\boldsymbol{x}-q)^{-1}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Theta_*\begin{pmatrix}\Sigma^k_{[k]} & P_{[k]}(\boldsymbol{x})\\ \Sigma^k_{[k+1]} & P_{[k+1]}(\boldsymbol{x})\end{pmatrix}. \tag{3.4.3}$$

For the second kind functions analogous relations hold

$$(TC)_{[k]}(\boldsymbol{x})=(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Theta_*\begin{pmatrix}\Sigma^k_{[k]} & C_{[k]}(\boldsymbol{x})\\ \Sigma^k_{[k+1]} & C_{[k+1]}(\boldsymbol{x})\end{pmatrix}. \tag{3.4.4}$$

Proof. To prove (3.4.3) introduce in (3.2.2) the expressions given in (3.4.2) to get the kernel polynomials

$$\begin{aligned}(TP)_{[k]}(\boldsymbol{x})&=(\boldsymbol{n}\cdot\boldsymbol{x}-q)^{-1}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\left(P_{[k+1]}(\boldsymbol{x})-\Sigma^k_{[k+1]}\left(\Sigma^k_{[k]}\right)^{-1}P_{[k]}(\boldsymbol{x})\right)\\&=(\boldsymbol{n}\cdot\boldsymbol{x}-q)^{-1}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\,\mathrm{SC}\begin{pmatrix}\Sigma^k_{[k]} & P_{[k]}(\boldsymbol{x})\\ \Sigma^k_{[k+1]} & P_{[k+1]}(\boldsymbol{x})\end{pmatrix},\end{aligned} \tag{3.4.5}$$

from where the result follows. For (3.4.4) we recall (3.2.5) and use (3.4.2). □

We remark that we are using the notation of [96] for the quasi-determinants, and in fact in this case the lower right corner is a $|[k]|$-th dimensional vector, not even a square matrix, therefore we are dealing with an extended quasi-determinant with some of elements not in a ring. Notice that this result extends to the multivariate situation the well known 1D situation described by (3.4.1). For the transformed quasi-tau functions H's and the coefficients β we have

Proposition 3.4.2. *When the conditions specified in Theorem 3.4.1 are satisfied the elementary Darboux transformations of the matrices $H_{[k]}$ and $\beta_{[k]}$ are given by the following quasi-determinantal formulae*

$$(TH)_{[k]} = (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Theta_*\begin{pmatrix}\Sigma^k_{[k]} & H_{[k]}\\ \Sigma^k_{[k+1]} & 0\end{pmatrix}$$

and we have the relation

$$(T\beta)_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} = q + (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Theta_*\begin{pmatrix}\Sigma^k_{[k]} & \mathbb{I}_{[k]}\\ \Sigma^k_{[k+1]} & \beta_{[k+1]}\end{pmatrix}.$$

Proof. The first relation is an immediate consequence (3.4.2) and (3.1.15) so that

$$(TH)_{[k]} = -(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]}\big(\Sigma^k_{[k]}\big)^{-1}H_{[k]}.$$

From (3.4.2) and (3.1.16) we get

$$(T\beta)_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} - (\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\beta_{[k+1]} - q = -(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]}\big(\Sigma^k_{[k]}\big)^{-1}$$

that is

$$(T\beta)_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} = q + (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\big(\beta_{[k+1]} - (\Sigma^k_{[k+1]}\big(\Sigma^k_{[k]}\big)^{-1}\big). \qquad \square$$

Observe that we can diagrammatically write *à la Gel'fand*

$$(TH)_{[k]} = \begin{vmatrix}\Sigma^k_{[k]} & H_{[k]}\\ (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]} & \boxed{0}\end{vmatrix},$$

$$(T\beta)_{[k]}(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]} = q + \begin{vmatrix}\Sigma^k_{[k]} & \mathbb{I}_{[k]}\\ (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]} & \boxed{(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\beta_{[k+1]}}\end{vmatrix}.$$

We can give a more explicit expression for β using the Moore–Penrose pseudo-inverse. In fact, using Proposition 6.2.3 and the multinomial matrix $\mathcal{M}_{[k]}$ (A.2.3) we can write

$$\begin{aligned}(T\beta)_{[k]} &= q\mathcal{M}_{[k]}^{-1/2}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}\big)^+\\ &+\begin{vmatrix}\Sigma^k_{[k]} & \mathcal{M}_{[k]}^{-1/2}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}\big)^+\\ (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\Sigma^k_{[k+1]} & \boxed{(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\beta_{[k+1]}\mathcal{M}_{[k]}^{-1/2}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}\big)^+}\end{vmatrix}.\end{aligned}$$

3.5. Multivariate Christoffel formula and quasi-determinants

For $D = 1$ there is a well known formula for the orthogonal polynomials $\{q_n(x)\}$ associated to a measure of the form $c(x-q_1)\cdots(x-q_m)\mathrm{d}\mu(x)$ in terms of the orthogonal polynomials $\{p_n(x)\}$ of the measure $\mathrm{d}\mu(x)$, see §2.5 of [109], as

$$q_n(x) = \frac{1}{c(x-q_1)\cdots(x-q_m)}\begin{vmatrix} p_n(x) & \dots & p_{n+m}(x) \\ p_n(q_1) & \dots & p_{n+m}(q_1) \\ \vdots & & \vdots \\ p_n(q_l) & \dots & p_{n+m}(q_l)\end{vmatrix}.$$

The Hungarian Mathematician Gabor Szegő, who considers the proof very easy, points out that it was proven for $\mathrm{d}\mu = \mathrm{d}x$ by Elwin Bruno Christoffel in [34]. This fact was rediscovered in the Toda context, see for example the formula (5.1.11) in [89] for $W_n^+(N)$.

In this section we will construct an analog to the Christoffel formula in the multivariate context. We use the sample matrix trick and quasi-determinants.

3.5.1. Iterating two elementary Darboux transformations

First, for a better understanding we discuss the iteration of two elementary Darboux transformations

$$\mathrm{d}\mu(\boldsymbol{x}) \to (\boldsymbol{n}^{(1)}\cdot\boldsymbol{x} - q^{(1)})\mathrm{d}\mu(\boldsymbol{x}) \to (\boldsymbol{n}^{(2)}\cdot\boldsymbol{x} - q^{(2)})(\boldsymbol{n}^{(1)}\cdot\boldsymbol{x} - q^{(1)})\mathrm{d}\mu(\boldsymbol{x})$$

or, equivalently, $\mathrm{d}\mu \to T^{(1)}T^{(2)}\mathrm{d}\mu$.

Given the corresponding lattice resolvents

$$\omega^{(a)} = (T^{(a)}S)(\boldsymbol{n}^{(a)}\cdot\boldsymbol{\Lambda} - q^{(a)})S^{-1}, \quad a \in \{1,2\}, \tag{3.5.1}$$

we introduce

Definition 3.5.1. The second iterated resolvent matrix is

$$\omega := \big(T^{(2)}\omega^{(1)}\big)\omega^{(2)}. \tag{3.5.2}$$

A first result regarding the two step Darboux transformation is

Proposition 3.5.1. *The MVOPR satisfy*

$$(\boldsymbol{n}^{(2)}\cdot\boldsymbol{x} - q^{(2)})(\boldsymbol{n}^{(1)}\cdot\boldsymbol{x} - q^{(1)})(T^{(2)}T^{(1)}P)(\boldsymbol{x}) = \omega P(\boldsymbol{x}). \tag{3.5.3}$$

Proof. It is a consequence of $(\boldsymbol{n}^{(a)}\cdot\boldsymbol{x} - q^{(a)})T^{(a)}P = \omega^{(a)}P$ with $a \in \{1,2\}$. □

Regarding the matrix structure of ω if we define

$$\boldsymbol{n} := q^{(1)}\boldsymbol{n}^{(2)} + q^{(2)}\boldsymbol{n}^{(1)} \in \mathbb{R}^D$$

we quickly find that

Proposition 3.5.2. *The second iterated resolvent ω decomposes in diagonals as follows*

$$\begin{aligned}\omega = &\underbrace{(\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})}_{\text{second superdiagonal}}\\ &+\underbrace{(T^{(1)}T^{(2)}\beta)(\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda}) - (\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\beta - \boldsymbol{n}\cdot\boldsymbol{\Lambda}}_{\text{first superdiagonal}}\\ &+\underbrace{(T^{(1)}T^{(2)}H)H^{-1}}_{\text{diagonal}}\end{aligned} \tag{3.5.4}$$

Proof. From (3.5.1) and (3.5.2) we get

$$\omega = (T^{(1)}T^{(2)}S)\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda}) - \boldsymbol{n}\cdot\boldsymbol{\Lambda} + q^{(1)}q^{(2)}\big)S^{-1}$$

and the two superdiagonal terms follow immediately. Now, from $\omega^{(a)} = \boldsymbol{n}^{(a)}\cdot\boldsymbol{\Lambda} + (T^{(a)}H)H^{-1}$ we get the diagonal part $\big(T^{(2)}((T^{(1)}H)H^{-1})\big)(T^{(2)}H)H^{-1}$. □

Notice that the Zakharov–Shabat or compatibility equations (3.1.21), which can be written as the symmetry condition $\omega = \big(T^{(2)}\omega^{(1)}\big)\omega^{(2)} = \big(T^{(1)}\omega^{(2)}\big)\omega^{(1)}$, are an immediate consequence of the previous result.

Proposition 3.5.3. *Relations (3.5.4) can be written componentwise as follows*

$$\begin{aligned}(\omega)_{[k],[k+2]} &= \big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]},\\ (\omega)_{[k],[k+1]} &= (T^{(1)}T^{(2)}\beta_{[k]})\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k-1],[k+1]}\\ &\quad - \big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]}\beta_{[k+2]} - (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\\ (\omega)_{[k],[k]} &= (T^{(1)}T^{(2)}H_{[k]})H_{[k]}^{-1}.\end{aligned}$$

Again we need to look at certain hyperplanes $\pi^{(a,+)} = \{\boldsymbol{x}\in\mathbb{R}^D : \boldsymbol{n}^{(a)}\cdot\boldsymbol{x} = q^{(a)}\}$, for $a\in\{1,2\}$.

Proposition 3.5.4. *For any $\boldsymbol{p}\in\pi^{1,+}\cup\pi^{2,+}$, i.e., either $\boldsymbol{n}^{(1)}\cdot\boldsymbol{p} = q^{(1)}$ or $\boldsymbol{n}^{(2)}\cdot\boldsymbol{p} = q^{(2)}$, we have*

$$\omega_{[k],[k+2]}P_{[k+2]}(\boldsymbol{p}) + \omega_{[k],[k+1]}P_{[k+1]}(\boldsymbol{p}) + \omega_{[k],[k]}P_{[k]}(\boldsymbol{p}) = 0. \tag{3.5.5}$$

Proof. It follows from (3.5.3). □

We now employ the sample matrix trick used for the elementary Darboux transformation to characterize ω in terms of MVOPR evaluated at some particular points. For the sets $\{\boldsymbol{p}_j^{(1)}\}_{j=1}^{|[k]|}, \{\boldsymbol{p}_j^{(2)}\}_{j=1}^{|[k+1]|}$ we use the notation introduced in Definition 3.4.1, i.e. we use the matrices $\Sigma_{[\ell]}^{(1),k}$ for the first set of points and $\Sigma_{[\ell]}^{(2),k}$ for the second set.

Proposition 3.5.5. *Suppose that* $\{\boldsymbol{p}_j^{(1)}\}_{j=1}^{|[k]|} \cup \{\boldsymbol{p}_j^{(2)}\}_{j=1}^{|[k+1]|} \subset \pi^{1,+} \cup \pi^{2,+}$ *is a poised set for* $\begin{pmatrix} P_{[k]} \\ P_{[k+1]} \end{pmatrix}$, *i.e.* $\begin{vmatrix} \Sigma_{[k]}^{(1),k} & \Sigma_{[k+1]}^{(2),k} \\ \Sigma_{[k+1]}^{(1),k} & \Sigma_{[k+1]}^{(2),k+1} \end{vmatrix} \neq 0$. *Then*

$$\begin{pmatrix} \omega_{[k],[k]} & \omega_{[k],[k+1]} \end{pmatrix} = -\omega_{[k],[k+2]} \begin{pmatrix} \Sigma_{[k+2]}^{(1),k} & \Sigma_{[k+2]}^{(2),k+1} \end{pmatrix} \begin{pmatrix} \Sigma_{[k]}^{(1),k} & \Sigma_{[k]}^{(2),k+1} \\ \Sigma_{[k+1]}^{(1),k} & \Sigma_{[k+1]}^{(2),k+1} \end{pmatrix}^{-1}.$$

Proof. As said we proceed as in the proof of Proposition 3.4.2 and evaluate (3.5.5) in the set $\{\boldsymbol{p}_j^{(1)}\}_{j=1}^{|[k]|} \cup \{\boldsymbol{p}_j^{(2)}\}_{j=1}^{|[k+1]|}$ to get

$$\begin{pmatrix} \omega_{[k],[k]} & \omega_{[k],[k+1]} \end{pmatrix} \begin{pmatrix} \Sigma_{[k]}^{(1),k} & \Sigma_{[k]}^{(2),k+1} \\ \Sigma_{[k+1]}^{(1),k} & \Sigma_{[k+1]}^{(2),k+1} \end{pmatrix} = -\omega_{[k],[k+2]} \begin{pmatrix} \Sigma_{[k+2]}^{(1),k} & \Sigma_{[k+2]}^{(2),k+1} \end{pmatrix},$$

from where the result follows. $\square$

Theorem 3.5.1. *For the composition of two elementary Darboux transformations, when the conditions required in Proposition 3.5.5 hold, we have the following multivariate quasi-determinantal Christoffel formula for the kernel polynomials*

$$\begin{aligned}
&(T^{(2)}T^{(1)}P)_{[k]}(\boldsymbol{x}) \\
&\quad = \frac{\big((\boldsymbol{n}^{(1)} \cdot \boldsymbol{\Lambda})(\boldsymbol{n}^{(2)} \cdot \boldsymbol{\Lambda})\big)_{[k],[k+2]}}{(\boldsymbol{n}^{(2)} \cdot \boldsymbol{x} - q^{(2)})(\boldsymbol{n}^{(1)} \cdot \boldsymbol{x} - q^{(1)})} \Theta_* \begin{pmatrix} \Sigma_{[k]}^{(1),k} & \Sigma_{[k]}^{(2),k+1} & P_{[k]}(\boldsymbol{x}) \\ \Sigma_{[k+1]}^{(1),k} & \Sigma_{[k+1]}^{(2),k+1} & P_{[k+1]}(\boldsymbol{x}) \\ \Sigma_{[k+2]}^{(1),k} & \Sigma_{[k+2]}^{(2),k+1} & P_{[k+2]}(\boldsymbol{x}) \end{pmatrix}.
\end{aligned}$$

Proof. From (3.5.3) we get

$$\begin{aligned}
&(\boldsymbol{n}^{(2)} \cdot \boldsymbol{x} - q^{(2)})(\boldsymbol{n}^{(1)} \cdot \boldsymbol{x} - q^{(1)})(T^{(2)}T^{(1)}P)_{[k]}(\boldsymbol{x}) \\
&\quad = \omega_{[k],[k+2]} P_{[k+2]}(\boldsymbol{x}) + \omega_{[k],[k+1]]} P_{[k+1]}(\boldsymbol{x}) + \omega_{[k],[k]} P_{[k]}(\boldsymbol{x}) \\
&\quad = \omega_{[k],[k+2]} \left(P_{[k+2]}(\boldsymbol{x}) - \begin{pmatrix} \Sigma_{[k+2]}^{(1),k} & \Sigma_{[k+2]}^{(2),k+1} \end{pmatrix} \begin{pmatrix} \Sigma_{[k]}^{(1),k} & \Sigma_{[k]}^{(2),k} \\ \Sigma_{[k+1]}^{(1),k} & \Sigma_{[k+1]}^{(2),k+1} \end{pmatrix}^{-1} \begin{pmatrix} P_{[k]}(\boldsymbol{x}) \\ P_{[k+1]}(\boldsymbol{x}) \end{pmatrix} \right)
\end{aligned}$$

from where the result follows. $\square$

Proposition 3.5.6. *The quasi-tau matrices $H_{[k]}$ and the $\beta_{[k]}$ matrices transform for a 2-step elementary Darboux transformation according to the following quasi-determinantal formulae*

$$(T^{(1)}T^{(2)}H)_{[k]} = \big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]}\begin{vmatrix} \Sigma^{(1),k}_{[k]} & \Sigma^{(2),k+1}_{[k]} & H_{[k]} \\ \Sigma^{(1),k}_{[k+1]} & \Sigma^{(2),k+1}_{[k+1]} & 0_{[k+1],[k]} \\ \Sigma^{(1),k}_{[k+2]} & \Sigma^{(2),k+1}_{[k+2]} & \boxed{0_{[k+2],[k]}} \end{vmatrix},$$

$$\begin{aligned}&(T^{(1)}T^{(2)}\beta_{[k]})\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k-1],[k+1]}\\ &\quad= (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}\\ &\qquad+ \big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]}\begin{vmatrix} \Sigma^{(1),k}_{[k]} & \Sigma^{(2),k+1}_{[k]} & 0_{[k],[k+1]} \\ \Sigma^{(1),k}_{[k+1]} & \Sigma^{(2),k+1}_{[k+1]} & \mathbb{I}_{[k+1]} \\ \Sigma^{(1),k}_{[k+2]} & \Sigma^{(2),k+1}_{[k+2]} & \boxed{\beta_{[k+2]}} \end{vmatrix}.\end{aligned}$$

Proof. See Appendix D.9. □

Observe that in this case we have used Gel'fand style instead of the Olver's notation for quasi-determinant.

3.5.2. The general case: m steps Darboux transformations

We are now ready to consider the general case of m iterated elementary Darboux transformations

$$\mathrm{d}\mu(\boldsymbol{x}) \to \mathcal{Q}(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x}), \qquad \mathcal{Q} := \prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{x} - q^{(i)}),$$

i.e., $\mathrm{d}\mu \to T\mathrm{d}\mu$, where $T := T^{(1)}\cdots T^{(m)}$ is the iteration of m elementary Darboux transformations.

In terms of the lattice resolvents

$$\omega^{(i)} = (T^{(i)}S)(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda} - q^{(i)})S^{-1}, \qquad i\in\{1,\dots,m\},$$

we introduce

Definition 3.5.2. The m-th iterated resolvent is

$$\omega := \big(T^{(m)}\cdots T^{(2)}\omega^{(1)}\big)\big(T^{(m)}\cdots T^{(3)}\omega^{(2)}\big)\cdots\omega^{(m)}.$$

From the definition we see that

Proposition 3.5.7. *The m-th iterated resolvent satisfies*

$$\omega=(TS)\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda}-q^{(i)})\Big)S^{-1}. \tag{3.5.6}$$

Proposition 3.5.8. *The m-th iterated resolvent ω can be expressed in diagonals as follows*

$$\begin{aligned}\omega=&\underbrace{\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)}_{m\text{-th superdiagonal}}\\&+\underbrace{(T\beta)\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)-\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)\beta-\sum_{i=1}^{m}q^{(i)}\prod_{j\neq i}(\boldsymbol{n}^{(j)}\cdot\boldsymbol{\Lambda})}_{(m-1)\text{-th superdiagonal}}\\&\vdots\\&+\underbrace{(TH)H^{-1}}_{\text{diagonal}}\end{aligned} \tag{3.5.7}$$

Proof. Observe that $\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda}-q^{(i)})$ splits into m block superdiagonals. The m-th superdiagonal is $\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})$ while the $(m-1)$-th superdiagonal is given by $-\sum_{i=1}^{m}q^{(i)}\prod_{j\neq i}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})$. Then, applying (3.5.6) we get the two higher superdiagonals of the m-iterated resolvent. Now, from $\omega^{(i)}=\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda}+(T^{(i)}H)H^{-1}$ we get the diagonal part. □

The components of the m-th iterated resolvent ω are

$$\begin{aligned}\omega_{[k],[k+m]}&=\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]},\\ \omega_{[k],[k+m-1]}&=(T\beta)_{[k]}\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k-1],[k+m-1]}-\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\beta_{[k+m]}\\&\quad-\sum_{i=1}^{m}q^{(i)}\Big(\prod_{j\neq i}(\boldsymbol{n}^{(j)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m-1]},\\ \omega_{[k],[k]}&=(TH)_{[k]}H^{-1}_{[k]}.\end{aligned} \tag{3.5.8}$$

Proposition 3.5.9. *The MVOPR and the second kind functions satisfy*

$$\mathcal{Q}(\boldsymbol{x})TP(\boldsymbol{x})=\omega P(\boldsymbol{x}), \tag{3.5.9}$$

$$TC(\boldsymbol{x})=\omega C(\boldsymbol{x}). \tag{3.5.10}$$

Proof. It follows from $(\boldsymbol{n}^{(i)}\cdot\boldsymbol{x}-q^{(i)})T^{(i)}P(\boldsymbol{x})=\omega^{(i)}P(\boldsymbol{x})$ with $i\in\{1,\dots,m\}$ and $T^{(i)}C(\boldsymbol{x})=\omega^{(i)}C(\boldsymbol{x})$. □

We consider again the hyperplanes $\pi^{(i,+)} = \{\boldsymbol{x} \in \mathbb{R}^D : \boldsymbol{n}^{(i)} \cdot \boldsymbol{x} = q^{(i)}\}$ for $i \in \{1, \dots, m\}$ to get

Proposition 3.5.10. *For any $\boldsymbol{p} \in \cup_{i=1}^m \pi^{i,+}$ we have*

$$\omega_{[k],[k+m]} P_{[k+m]}(\boldsymbol{p}) + \omega_{[k],[k+1]} P_{[k+m-1]}(\boldsymbol{p}) + \cdots + \omega_{[k],[k]} P_{[k]}(\boldsymbol{p}) = 0. \quad (3.5.11)$$

Proof. It follows from (3.5.9). $\square$

The sample matrix trick is used again to characterize ω in terms of MVOPR evaluated at some particular points. For the sets $\{\boldsymbol{p}_j^{(i)}\}_{j=1}^{|[k+i-1]|}$, $i \in \{1, \dots, m\}$, we use the notation $\Sigma_{[\ell]}^{(i),k}$ introduced in Definition 3.4.1.

Proposition 3.5.11. *Suppose that $\cup_{i=1}^m \{\boldsymbol{p}_j^{(i)}\}_{j=1}^{|[k-1+i]|} \subset \cup_{i=1}^m \pi^{(i),+}$ is a poised set for $\begin{pmatrix} P_{[k]}(\boldsymbol{x}) \\ \vdots \\ P_{[k+m-1]}(\boldsymbol{x}) \end{pmatrix}$, i.e.*

$$\begin{vmatrix} \Sigma_{[k]}^{(1),k} & \dots & \Sigma_{[k]}^{(m),k+m-1} \\ \vdots & & \vdots \\ \Sigma_{[k+m-1]}^{(1),k} & \dots & \Sigma_{[k+m-1]}^{(m),k+m-1} \end{vmatrix} \neq 0.$$

Then

$$\begin{aligned} &\begin{pmatrix} \omega_{[k],[k]} & \dots & \omega_{[k],[k+m-1]} \end{pmatrix} \\ &= -\omega_{[k],[k+m]} \begin{pmatrix} \Sigma_{[k+m]}^{(1),k} & \dots & \Sigma_{[k+m]}^{(m),k+m-1} \end{pmatrix} \begin{pmatrix} \Sigma_{[k]}^{(1),k} & \dots & \Sigma_{[k]}^{(m),k+m-1} \\ \vdots & & \vdots \\ \Sigma_{[k+m-1]}^{(1),k} & \dots & \Sigma_{[k+m-1]}^{(m),k+m-1} \end{pmatrix}^{-1}. \end{aligned}$$

Proof. Evaluate (3.5.11) in the set $\cup_{i=1}^m \{\boldsymbol{p}_j^{(i)}\}_{j=1}^{|[k-1+i]|}$ to get

$$\begin{aligned} \begin{pmatrix} \omega_{[k],[k]} & \dots & \omega_{[k],[k+m-1]} \end{pmatrix} &\begin{pmatrix} \Sigma_{[k]}^{(1),k} & \dots & \Sigma_{[k]}^{(m),k+m-1} \\ \vdots & & \vdots \\ \Sigma_{[k+m-1]}^{(1),k} & \dots & \Sigma_{[k+m-1]}^{(m),k+m-1} \end{pmatrix} \\ &= -\omega_{[k],[k+m]} \begin{pmatrix} \Sigma_{[k+m]}^{(1),k} & \dots & \Sigma_{[k+m]}^{(m),k+m-1} \end{pmatrix} \end{aligned}$$

and the desired result follows immediately. $\square$

Theorem 3.5.2. *If the conditions specified in Proposition 3.5.11 are fulfilled the following multivariate quasi-determinantal Christoffel formulae hold true*

$$TP_{[k]}(\boldsymbol{x})=\frac{\big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+m]}}{\mathcal{Q}(\boldsymbol{x})}\Theta_*\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & P_{[k]}(\boldsymbol{x})\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & P_{[k+m]}(\boldsymbol{x})\end{pmatrix},$$

$$TC_{[k]}(\boldsymbol{x})=\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\Theta_*\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & C_{[k]}(\boldsymbol{x})\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & C_{[k+m]}(\boldsymbol{x})\end{pmatrix}.$$

Proof. See Appendix D.10. □

In the scalar case, $D=1$, this formula in the OPRL context is known as Christoffel formula, see for example [109].

3.5.3. Quasi-determinantal expressions for the resolvent, quasi-tau matrices and β

An interesting consequence of Proposition 3.5.11 is the following

Proposition 3.5.12. *The resolvent coefficients can be expressed as quasi-determinants as follows*

$$\omega_{[k],[k+i]}=\Big(\prod_{j=1}^{m}(\boldsymbol{n}^{(j)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\Theta_*\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & 0_{[k],[k+i]}\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+i-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+i-1]} & 0_{[k+i-1],[k+i]}\\ \Sigma^{(1),k}_{[k+i]} & \dots & \Sigma^{(m),k+m-1}_{[k+i]} & \mathbb{I}_{[k+i]}\\ \Sigma^{(1),k}_{[k+i+1]} & \dots & \Sigma^{(m),k+m-1}_{[k+i+1]} & 0_{[k+i+1],[k+i]}\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & 0_{[k+m],[k+i]}\end{pmatrix}.$$

Proof. Assuming that $\cup_{i=1}^{m}\{\boldsymbol{p}_j^{(i)}\}_{j=1}^{|[k-1+i]|}\subset\cup_{i=1}^{m}\pi^{(i),+}$ is a poised set for $\begin{pmatrix}P_{[k]}(\boldsymbol{x})\\ \vdots\\ P_{[k+m-1]}(\boldsymbol{x})\end{pmatrix}$, then for $i\in\{0,\dots,m-1\}$

$$\omega_{[k],[k+i]}=-\omega_{[k],[k+m]}\begin{pmatrix}\Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]}\end{pmatrix}$$
$$\times\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]}\\ \vdots & & \vdots\\ \Sigma^{(1),k}_{[k+m-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-1]}\end{pmatrix}^{-1}\begin{pmatrix}0_{[k],[k+i]}\\ \vdots\\ 0_{[k+i-1],[k+i]}\\ \mathbb{I}_{[k+i]}\\ 0_{[k+i+1],[k+i]}\\ \vdots\\ 0_{[k+m-1],[k+i]}\end{pmatrix}. \qquad □$$

The previous proposition together with (3.5.8) gives

Proposition 3.5.13. *The m-th iteration of elementary Darboux transformations has the following effects on the quasi-tau matrices $H_{[k]}$*

$$TH_{[k]}=\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\Theta_*\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & H_{[k]}\\ \Sigma^{(1),k}_{[k+1]} & \dots & \Sigma^{(m),k+m-1}_{[k+1]} & 0_{[k+1],[k]}\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & 0_{[k+m],[k]},\end{pmatrix},$$

and on the matrices $\beta_{[k]}$

$$\begin{aligned}(T\beta)_{[k]}\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k-1],[k+m-1]}&=\sum_{i=1}^{m}q^{(i)}\Big(\prod_{j\neq i}(\boldsymbol{n}^{(j)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m-1]}\\&+\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\Theta_*\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & 0_{[k],[k+m-1]}\\ \vdots & & \vdots & \vdots\\ \Sigma^{(1),k}_{[k+m-2]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-2]} & 0_{[k+m-2],[k+m+1]}\\ \Sigma^{(1),k}_{[k+m-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-1]} & \mathbb{I}_{[k+m+1]}\\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & \beta_{[k+m]}\end{pmatrix}.\end{aligned}$$

We remark that in Proposition 3.5.6 and Proposition 3.5.13 we can get explicitly the transformed β by multiplying on the right by right inverse matrices of the product of factors of type $\boldsymbol{n}\cdot\boldsymbol{\Lambda}$, for that aim see Proposition 6.2.3. We obtain that a right inverse of $\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k-1],[k+m-1]}$ can be expressed in terms of pseudo-inverses as

$$\prod_{i=0}^{m-1}\Big(\mathcal{M}^{-1/2}_{[k+i]}\big((\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})_{[k-1+i],[k+i]}\mathcal{M}^{-1/2}_{[k+i]}\big)^{+}\Big),$$

where we use the multinomial matrix (A.2.3). Therefore

Proposition 3.5.14. *After the iteration of m elementary Darboux transformations the transformed coefficient β can be expressed as a quasi-determinant as follows*

$$\begin{aligned}T\beta_{[k]}&=\sum_{i=1}^{m}q^{(i)}\Big(\prod_{j\neq i}(\boldsymbol{n}^{(j)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m-1]}\prod_{i=0}^{m-1}\Big(\mathcal{M}^{-1/2}_{[k+i]}\big((\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})_{[k-1+i],[k+i]}\mathcal{M}^{-1/2}_{[k+i]}\big)^{+}\Big)\\&+\Big(\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[k],[k+m]}\end{aligned}$$

$$\times\begin{vmatrix} \Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]} & 0_{[k],[k+m-1]} \\ \vdots & & \vdots & \vdots \\ \Sigma^{(1),k}_{[k+m-2]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-2]} & 0_{[k+m-2],[k+m+1]} \\ \Sigma^{(1),k}_{[k+m-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-1]} & \prod_{i=0}^{m-1}\Big(\mathcal{M}^{-1/2}_{[k+i]}((\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})_{[k-1+i],[k+i]}\mathcal{M}^{-1/2}_{[k+i]})^+\Big) \\ \Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]} & \boxed{\beta_{[k+m]}\prod_{i=0}^{m-1}\Big(\mathcal{M}^{-1/2}_{[k+i]}((\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})_{[k-1+i],[k+i]}\mathcal{M}^{-1/2}_{[k+i]})^+\Big)} \end{vmatrix}.$$

3.5.4. Christoffel–Darboux kernel and the kernel polynomials

The transformed polynomials TP_k are known in the 1D case as kernel polynomials because of the nice formula (3.4.1). We will show now that a similar relation holds in the multivariate situation, and thus the transformed MVOPR should receive the name of multivariate kernel polynomials in the same footing as it happens in the 1D scenario.

Now we consider a result similar to Theorem 3.3.1 but for m elementary Darboux transformations. In doing so we need

Definition 3.5.3. We introduce the following truncation matrix of the resolvent matrix

$$\omega^{[\ell,m]} := \begin{pmatrix} \omega_{[\ell],[\ell]} & \omega_{[\ell],[\ell+1]} & \dots & \omega_{[\ell],[\ell+m-2]} & \omega_{[\ell],[\ell+m-1]} \\ 0_{[\ell+1],[\ell]} & \omega_{[\ell+1],[\ell+1]} & \dots & \omega_{[\ell+1],[\ell+m-2]} & \omega_{[\ell+1],[\ell+m-1]} \\ \vdots & & \ddots & \vdots & \vdots \\ 0_{[\ell+m-2],[\ell]} & 0_{[\ell+m-2],[\ell+1]} & \dots & \omega_{[\ell+m-2],[\ell+m-2]} & \omega_{[\ell+m-1],[\ell+m-1]} \\ 0_{[\ell+m-1],[\ell]} & 0_{[\ell+m-1],[\ell+1]} & \dots & 0_{[\ell+m-1],[\ell+m-2]} & \omega_{[\ell+m-1],[\ell+m-1]} \end{pmatrix}$$

and the upper unitriangular matrix $\zeta^{[\ell,m]} := H^{[\ell]}\big(TH^{[\ell]}\big)^{-1}\omega^{[\ell,m]}$.

Notice that $\zeta^{[\ell,m]}_{[\ell+i],[\ell+i]} = \mathbb{I}_{[\ell+i]}$, for $i \in \{0,\dots,m-1\}$ and that $\big(H^{[\ell]}\big)^{-1}\zeta^{[\ell,m]} = \big(TH^{[\ell]}\big)^{-1}\omega^{[\ell,m]}$.

Theorem 3.5.3. *The following formula relating Christoffel–Darboux kernels after and before the iteration of m elementary Darboux transformations holds true*

$$\begin{aligned} K^{(\ell+m)}(\boldsymbol{x},\boldsymbol{y}) = {}& \mathcal{Q}(\boldsymbol{x})TK^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) \\ & + \begin{pmatrix} TP_{[\ell]}(\boldsymbol{y}) \\ \vdots \\ TP_{[\ell+m-1]}(\boldsymbol{y}) \end{pmatrix}^{\top} \big(H^{[\ell]}\big)^{-1}\zeta^{[\ell,m]} \begin{pmatrix} P_{[\ell]}(\boldsymbol{x}) \\ \vdots \\ P_{[\ell+m-1]}(\boldsymbol{x}) \end{pmatrix}. \end{aligned} \tag{3.5.12}$$

Proof. See Appendix D.11. □

For only one elementary Darboux transformation, $m = 1$, the above result reduces to

$$K^{(\ell+1)}(\boldsymbol{x},\boldsymbol{y}) = (\boldsymbol{n}\cdot\boldsymbol{x} - q)TK^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) + TP_{[\ell]}(\boldsymbol{y})^{\top}H^{-1}_{[\ell]}P_{[\ell]}(\boldsymbol{x}),$$

and we recover Theorem 3.3.1. For $m = 2$, i.e., the two step Darboux transformation, we get

$$K^{(\ell+2)}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{Q}(\boldsymbol{x})TK^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) + \begin{pmatrix} H_{[\ell]}^{-1}TP_{[\ell]}(\boldsymbol{y}) \\ H_{[\ell+1]}^{-1}TP_{[\ell+1]}(\boldsymbol{y}) \end{pmatrix}^{\top} \zeta^{([\ell,2]} \begin{pmatrix} P_{[\ell]}(\boldsymbol{x}) \\ P_{[\ell+1]}(\boldsymbol{x}) \end{pmatrix}$$

where

$$\mathcal{Q} = (\boldsymbol{n}^{(1)}\cdot\boldsymbol{x} - q^{(1)})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{x} - q^{(2)}), \qquad \zeta^{[\ell,2]} = \begin{pmatrix} \mathbb{I}_{[\ell]} & \zeta_{[\ell],[\ell+1]} \\ 0_{[\ell+1],[\ell]} & \mathbb{I}_{[\ell+1]} \end{pmatrix}$$

with

$$\begin{aligned}\zeta_{[\ell],[\ell+1]} = H_{[\ell]}(TH_{[\ell]})^{-1}\Big((T\beta)_{[\ell]}\Big(\prod_{i=1}^{2}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[\ell-1],[\ell+1]} \\ - \Big(\prod_{i=1}^{2}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{\Lambda})\Big)_{[\ell],[\ell+2]}\beta_{[\ell+2]} - (\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell],[\ell+1]}\Big)\end{aligned}$$

and $\boldsymbol{n} = q^{(1)}\boldsymbol{n}^{(2)} + q^{(2)}\boldsymbol{n}^{(1)}$. Then,

$$\begin{aligned}K^{(\ell+2)}(\boldsymbol{x},\boldsymbol{y}) = \mathcal{Q}(\boldsymbol{x})TK^{(\ell)}(\boldsymbol{x},\boldsymbol{y}) + TP_{[\ell]}(\boldsymbol{y})^{\top}H_{[\ell]}^{-1}P_{[\ell]}(\boldsymbol{x}) + TP_{[\ell+1]}(\boldsymbol{y})^{\top}H_{[\ell+1]}^{-1}P_{[\ell+1]}(\boldsymbol{x}) \\ + TP_{[\ell]}(\boldsymbol{y})^{\top}H_{[\ell]}^{-1}\zeta_{[\ell],[\ell+1]}P_{[\ell+1]}(\boldsymbol{x}).\end{aligned}$$

Definition 3.5.4. Given the set of points $\mathscr{P} := \cup_{i=1}^{m}\{\boldsymbol{p}_j^{(i)}\}_{j=1}^{|[\ell-1+i]|}$ we define the following Christoffel–Darboux vectors

$$\kappa^{(\ell,m)}(\boldsymbol{y}) := \begin{pmatrix} \kappa_{[\ell]}^{(\ell,m)}(\boldsymbol{y}) \\ \vdots \\ \kappa_{[\ell+m-1]}^{(\ell,m)}(\boldsymbol{y}) \end{pmatrix}^{\top}, \qquad \kappa_{[\ell+i-1]}^{(\ell,m)}(\boldsymbol{y}) := \begin{pmatrix} K^{(\ell+m)}(\boldsymbol{p}_1^{(i)},\boldsymbol{y}), \\ \vdots \\ K^{(\ell+m)}(\boldsymbol{p}_{|[\ell+i-1]|}^{(i)},\boldsymbol{y}) \end{pmatrix}^{\top},$$
$$i = 1,\dots,m.$$

We also introduce

$$P^{(\ell,m)}(\boldsymbol{x}) = \begin{pmatrix} P_{[\ell]}(\boldsymbol{x}) \\ \vdots \\ P_{[\ell+m-1]}(\boldsymbol{x}) \end{pmatrix}^{\top}, \qquad \Sigma^{[\ell,m]} := \begin{pmatrix} \Sigma_{[\ell]}^{(1),\ell} & \dots & \Sigma_{[\ell]}^{(m),\ell+m-1} \\ \vdots & & \vdots \\ \Sigma_{[\ell+m-1]}^{(1),\ell} & \dots & \Sigma_{[\ell+m-1]}^{(m),\ell+m-1} \end{pmatrix}.$$

Now, the sample matrix trick leads to the following finding that relates the Christoffel–Darboux kernel evaluated in a poised set and the transformed polynomials, justifying the denomination of kernel polynomials.

Proposition 3.5.15. *Assume that* $\mathscr{P} \subset \cup_{i=1}^{m} \pi^{i,+}$ *is a poised set, i.e.,* $\det \Sigma^{(\ell,m)} \neq 0$*; then,*

$$\kappa^{(\ell,m)}(\boldsymbol{x}) = (TP)^{(\ell,m)}(\boldsymbol{x})\big(H^{[\ell]}\big)^{-1}\zeta^{[\ell,m]}\Sigma^{[\ell,m]}.$$

The following quasi-determinantal expressions hold

$$TP^{(\ell,m)}(\boldsymbol{x}) = -\Theta_* \begin{pmatrix} \zeta^{[\ell,m]}\Sigma^{[\ell,m]} & H^{[\ell]} \\ \kappa^{(\ell,m)}(\boldsymbol{x}) & 0 \end{pmatrix}, \qquad \Theta_* \begin{pmatrix} \zeta^{[\ell,m]}\Sigma^{[\ell,m]} & H^{[\ell]} \\ \kappa^{(\ell,m)}(\boldsymbol{x}) & TP^{(\ell,m)}(\boldsymbol{x}) \end{pmatrix} = 0,$$

One recognizes the first formula as an extension to arbitrary dimensions and iterations of the 1D formula (3.4.1).

Given another set of points $\tilde{\mathscr{P}} := \cup_{i=1}^{m}\{\tilde{\boldsymbol{p}}_j^{(i)}\}_{j=1}^{|[\ell-1+i]|}$ which do not need to be neither in the union of hyper-planes nor poised, we form the corresponding matrix $\tilde{\Sigma}^{[\ell,m]}$ so that

$$(K(\boldsymbol{p}_j^{(i)}, \tilde{\boldsymbol{p}}_{\tilde{j}}^{(\tilde{i})})) = (T\tilde{\Sigma}^{[\ell,m]})\big(H^{[\ell]}\big)^{-1}\zeta^{[\ell,m]}\Sigma^{[\ell,m]}.$$

We are expressing the evaluation of the Christoffel–Darboux kernel of the LHS as the product of matrices that at the end are expressed as elementary Darboux transforms, or Miwa shifts—see §4.3—, of the quasi-tau matrices, see §3.2.1, as it happens in Theorem 8.1 in [4]. It is not so clear if it is helpful for the computation of the Fredholm determinant $\det(1 - \lambda S_\ell)$ of the projection S_ℓ with kernel $K^{(\ell+1)}(\boldsymbol{x}, \boldsymbol{y})$ [4].

4. On continuous Toda and MVOPR

Now, once we have consider the Toda type discrete flows and the corresponding moments matrices $G(\boldsymbol{m})$ we are ready to add continuous deformations to the moment matrix. We will see that for given appropriate deformations or flows of a given measure we get an integrable hierarchy that extends the 2D Toda lattice hierarchy. In our extension the dependent variables are size varying matrices which satisfy Toda type nonlinear PDE.

4.1. The continuous flows

We first introduce of time deformations

Definition 4.1.1. Let us define the following covector of time variables

$$t = (t_{[0]}, t_{[1]}, \dots), \qquad t_{[k]} = (t_{\alpha_1^{(k)}}, \dots, t_{\alpha_{|[k]|}^{(k)}}), \qquad t_{\alpha_j^{(k)}} \in \mathbb{R}.$$

Observe that the just introduced times can be considered as elements in the symmetric algebra $t^\top \in \mathrm{S}(\mathbb{R}^D)$.

Definition 4.1.2. The deformation matrix is

$$W_0(t,\boldsymbol{m})=\exp\Big(\sum_{k=0}^{\infty}\sum_{j=1}^{|[k]|}t_{\alpha_j^{(k)}}\Lambda_{\alpha_j^{(k)}}\Big)\prod_{a=1}^{D}(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)^{m_a},$$

and the deformed moment matrix is

$$G(t,\boldsymbol{m}):=W_0(t,\boldsymbol{m})G.$$

Notice that $G(t,\boldsymbol{m})=G(\boldsymbol{m})W_0(t,\boldsymbol{m})^\top$, and

$$\Lambda_{\boldsymbol{k}}G(t,\boldsymbol{m})=G(t,\boldsymbol{m})\big(\Lambda_{\boldsymbol{k}}\big)^\top,\qquad\forall\boldsymbol{k}\in\mathbb{Z}_+^D,\qquad G(t,\boldsymbol{m})=(G(t,\boldsymbol{m}))^\top.$$

Definition 4.1.3. We introduce the notation

$$t(\boldsymbol{x}):=t\chi(\boldsymbol{x})=\sum_{k=0}^{\infty}\sum_{j=1}^{|[k]|}t_{\alpha_j^{(k)}}\boldsymbol{x}^{\alpha_j^{(k)}}.$$

It is not difficult to see that

Proposition 4.1.1. *The deformed moment matrix is the moment matrix of the following deformed measure*

$$\mathrm{d}\mu_{t,\boldsymbol{m}}(\boldsymbol{x})=\mathrm{e}^{t(\boldsymbol{x})}\mathrm{d}\mu_{\boldsymbol{m}}(\boldsymbol{x})=\mathrm{e}^{t(\boldsymbol{x})}\Big[\prod_{a=1}^{D}\big(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a\big)^{m_a}\Big]\mathrm{d}\mu(\boldsymbol{x}).$$

The Cholesky factorization

$$G(t,\boldsymbol{m}):=(S(t,\boldsymbol{m}))^{-1}H(t,\boldsymbol{m})\big((S(t,\boldsymbol{m}))^{-1}\big)^\top \tag{4.1.1}$$

leads to new MVOPR depending on both continuous and discrete time parameters. We introduce

Definition 4.1.4. The wave semi-infinite matrices are

$$W_1(t,\boldsymbol{m}):=S(t,\boldsymbol{m})W_0(t,\boldsymbol{m}),\qquad W_2(t,\boldsymbol{m}):=H(t,\boldsymbol{m})\big((S(t,\boldsymbol{m}))^{-1}\big)^\top. \tag{4.1.2}$$

An important fact regarding wave matrices and Gaussian decomposition of the evolved moment matrix is

Proposition 4.1.2. *The wave matrices factorize the nondeformed moment matrix as follows*

$$G=\big(W_1(t,\boldsymbol{m})\big)^{-1}W_2(t,\boldsymbol{m}). \tag{4.1.3}$$

Proof. From (4.1.1) we deduce that[12]

$$\begin{aligned} G &= (W_0(t,\boldsymbol{m})^{-1}G(t,\boldsymbol{m}) && (4.1.4)\\ &= (W_0(t,\boldsymbol{m})^{-1}(S(t,\boldsymbol{m}))^{-1}\,H(t,\boldsymbol{m})\big((S(t,\boldsymbol{m}))^{-1}\big)^{\top} && (4.1.5)\\ &= \big(W_1(t,\boldsymbol{m})\big)^{-1}W_2(t,\boldsymbol{m}). \qquad \square && (4.1.6)\end{aligned}$$

In what follows we will use the splitting as a direct sum of the linear semi-infinite matrices in strictly block lower triangular matrices and upper block triangular matrices. Then, M_+ will denote the projection of M in the upper triangular matrices while M_- the projection in the strictly lower triangular matrices.

Proposition 4.1.3. *The wave matrices W_1 and W_2 solve the following system of linear differential equations*

$$\frac{\partial W}{\partial t_{\alpha_j^{(k)}}} = \big(J_{\alpha_j^{(k)}}\big)_+ W, \qquad j=1,\dots,|[k]|, k=0,1,\dots.$$

Proof. Differentiate (3.1.11) to obtain

$$\frac{\partial S(t)}{\partial t_{\alpha_j^{(k)}}}S(t)^{-1} = -\big(J_{\alpha_j^{(k)}}(t)\big)_-, \qquad \frac{\partial W_2(t)}{\partial t_{\alpha_j^{(k)}}}W_2(t)^{-1} = \big(J_{\alpha_j^{(k)}}(t)\big)_+,$$

and the result follows. $\square$

Let us observe that

$$\frac{\partial S(t)}{\partial t_{\alpha_j^{(k)}}}S(t)^{-1} + \big(J_{\alpha_j^{(k)}}(t)\big)_- = 0$$

is of particular relevance. Put for example $k=1$ and consider the equations for the times $t_{[1]}=(t_1,\dots,t_D)$, the *first level times*,

$$\frac{\partial S}{\partial t_a}S^{-1} + (J_a)_- = 0. \tag{4.1.7}$$

[12] The product of two semi-infinite matrices is a delicate issue. There is no problem if we multiply lower triangular with lower triangular, upper triangular with upper triangular and even lower triangular with upper triangular, as all the coefficients of the resulting matrix are finite sums. But the multiplication of upper triangular with lower triangular could lead to problems as sums are now infinite series that need not to converge. This is why $W_1(t)$ is well defined $S(t)$ being lower triangular and $W_0(t)$ upper triangular. However, for $(W_1)^{-1}$ we need to be more careful as the *naïve* answer $(W_1)^{-1} = W_0^{-1}S_1^{-1}$ involves the product of an upper with a lower triangular. A possible answer is to say that the inverse from the right is $S^{-1}H(S^{-1})^{\top}(W_2(t))^{-1}$, which in fact is a consequence of this proposition. Despite of being a *formal* proposition, as we are assuming the existence of the inverse W_1^{-1} one could compute this inverse in appropriate domains. That is the case for the adjoint Baker functions.

Definition 4.1.5. Let us decompose the matrices by diagonals, we write

$$S = \mathbb{I} + \beta^{(1)} + \beta^{(2)} + \cdots \tag{4.1.8}$$

where $\beta^{(1)}$ is the first subdiagonal, i.e. $\beta = \beta^{(1)}$, and in general $\beta^{(k)}$ is the k-th subdiagonal of S.

Then

Proposition 4.1.4. *The coefficients $S_{[k],[k-j]} = \beta^{(j)}_{[k]}$ of the MVOPR are subject to differential relations and the first three are*

$$\begin{aligned}
\frac{\partial \beta_{[k]}}{\partial t_a} &= J_{[k],[k-1]},\\
\frac{\partial \beta^{(2)}_{[k]}}{\partial t_a} &= \frac{\partial \beta^{(1)}_{[k]}}{\partial t_a}\beta^{(1)}_{[k-1]},\\
\frac{\partial \beta^{(3)}_{[k]}}{\partial t_a} &= \frac{\partial \beta^{(2)}_{[k]}}{\partial t_a}\beta^{(1)}_{[k-1]} + \frac{\partial \beta^{(1)}_{[k]}}{\partial t_a}\beta^{(2)}_{[k-1]} - \frac{\partial \beta^{(1)}_{[k]}}{\partial t_a}\beta^{(1)}_{[k-1]}\beta^{(1)}_{[k-2]}.
\end{aligned}$$

Proof. See Appendix D.12. □

4.2. Baker functions. Lax and Zakharov–Shabat equations

Definition 4.2.1. Baker functions are defined by

$$\Psi_1 := W_1\chi, \qquad \Psi_2 := W_2\chi^*,$$

while adjoint Baker functions are given by

$$\Psi_1^* := (W_1^{-1})^\top\chi^*, \qquad \Psi_2^* := (W_2^{-1})^\top\chi.$$

We notice that Ψ_1 and Ψ_2^* lead to the computation of finite sums, but Ψ_1^* and Ψ_2 involve Laurent series; however $(\Psi_2)_{\alpha_i} = C_{\alpha_i}(t,\boldsymbol{m})$ and its domain of convergence is $\mathscr{D}_{\alpha_i}(t,\boldsymbol{m})$. We will denote by $\mathscr{D}^*_{\alpha_i}(t,\boldsymbol{m})$ the domain of convergence of $(\Psi_1^*)_{\alpha_i}(t,\boldsymbol{m})$.

Proposition 4.2.1. *The following expressions for the Baker functions in terms of MVOPR and its multivariate Cauchy transforms hold true*

$$\begin{aligned}
(\Psi_1)_{\alpha_i}(\boldsymbol{z}) &= \mathrm{e}^{t(\boldsymbol{z})}\Big[\prod_{a=1}^{D}\big(\boldsymbol{n}_a\cdot\boldsymbol{z} - q_a\big)^{m_a}\Big]P_{\alpha_i}(\boldsymbol{z},t,\boldsymbol{m}),\\
(\Psi_2)_{\alpha_i}(\boldsymbol{z}) &= \int_\Omega \frac{P_{\alpha_i}(\boldsymbol{y},t)}{(z_1-y_1)\cdots(z_D-y_D)}\mathrm{d}\mu_{t,\boldsymbol{m}}(\boldsymbol{y}), \qquad \boldsymbol{z}\in\mathscr{D}_{\alpha_i}(t,\boldsymbol{m})\setminus\operatorname{supp}(\mathrm{d}\mu),
\end{aligned}$$

$$(\Psi_1^*)_{\alpha_i}(\boldsymbol{z}) = \sum_{j=1}^{|[k]|} (H(t,\boldsymbol{m})^{-1})_{\alpha_i,\alpha_j} \int_\Omega \frac{P_{\alpha_j}(\boldsymbol{y},t,\boldsymbol{m})}{(z_1-y_1)\cdots(z_D-y_D)} \mathrm{d}\mu(\boldsymbol{y}),$$

$$\boldsymbol{z} \in \cap_{j=1}^{|[k]|} \mathscr{D}^*_{\alpha_j}(t,\boldsymbol{m}) \setminus \operatorname{supp}(\mathrm{d}\mu),$$

$$(\Psi_2^*)_{\alpha_i}(\boldsymbol{z}) = \sum_{j=1}^{|[k]|} (H(t,\boldsymbol{m})^{-1})_{\alpha_i,\alpha_j} P_{\alpha_j}(\boldsymbol{z},t,\boldsymbol{m}).$$

Proof. See Appendix D.13. □

Proposition 4.2.2. *The Baker functions and the adjoint Baker functions satisfy*

$$J_a\Psi_1 = x_a\Psi_1 \qquad J_a\Psi_2, = x_a\Psi_2 - \lim_{x_a\to\infty}[x_a\Psi_2],$$
$$J_a^\top\Psi_1^* = x_a\Psi_1^* - \lim_{x_a\to\infty}[x_a\Psi_1^*], \qquad J_a^\top\Psi_2^* = x_a\Psi_2^*.$$

Proposition 4.2.3.

(1) *The Baker functions are subject to the following linear system of differential equations*

$$\frac{\partial\Psi_i}{\partial t_{\alpha_j}} = \big(J_{\alpha_j^{(\ell)}}\big)_+\Psi_i, \qquad \frac{\partial\Psi_i^*}{\partial t_{\alpha_j}} = -\big(J_{\alpha_j^{(\ell)}}\big)_+^\top\Psi_i^*, \qquad i=1,2.$$

(2) *The MVOPR and its second kind functions satisfy*

$$\frac{\partial P}{\partial t_{\alpha_j}} = -\boldsymbol{x}^{\alpha_j}P + \big(J_{\alpha_j^{(\ell)}}\big)_+P, \qquad \frac{\partial C}{\partial t_{\alpha_j}} = \big(J_{\alpha_j^{(\ell)}}\big)_+C.$$

(3) *The following Lax equations hold*

$$\frac{\partial J_{\alpha_i^{(k)}}}{\partial t_{\alpha_j^{(\ell)}}} = \big[\big(J_{\alpha_j^{(\ell)}}\big)_+, J_{\alpha_i^{(k)}}\big].$$

(4) *The Zakharov–Shabat type equations*

$$\frac{\partial\big(J_{\alpha_i^{(k)}}\big)_+}{\partial t_{\alpha_j^{(\ell)}}} - \frac{\partial\big(J_{\alpha_j^{(\ell)}}\big)_+}{\partial t_{\alpha_i^{(k)}}} + \big[\big(J_{\alpha_i^{(k)}}\big)_+, \big(J_{\alpha_j^{(\ell)}}\big)_+\big] = 0,$$
$$\frac{\partial\omega_a}{\partial t_\alpha} - (T_aJ_\alpha)_+\omega_a + \omega_a(J_\alpha)_+ = 0$$

are fulfilled.

4.3. Miwa shifts and discrete flows

We will reproduce a characteristic fact in integrable systems, the Miwa's coherent shifts in the time variables lead to discrete flows and Darboux transformations. We now will indicate how these Miwa shifts are for this multivariate context. The simplest case is perhaps the most interesting one as it reproduces the discrete flows we have considered previously. The coherent shifts in the times

$$t \to t' = t \pm [q]_a, \qquad t'_\alpha = \begin{cases} t_\alpha, & \alpha \notin \mathbb{Z}_+ \boldsymbol{e}_a, \\ t_{m\boldsymbol{e}_a} \pm \dfrac{1}{mq^m}, & \alpha = n\boldsymbol{e}_a \text{ with } m \in \mathbb{Z}_+, \end{cases}$$

lead to the following deformation of the measure $\mathrm{d}\mu_t$

$$\mathrm{d}\mu_t(\boldsymbol{x}) \longrightarrow \mathrm{d}\mu_{t'} = \Big(1 - \frac{x_a}{q}\Big)^{\mp 1} \mathrm{d}\mu_t(\boldsymbol{x}) = -q^{\pm 1}\mathrm{d}\left(T_a^{\mp}\mu_t(\boldsymbol{x})\right),$$

which follows from

$$\log\left(\Big(1 - \frac{x_a}{q}\Big)^{\mp 1}\right) = \pm \sum_{m=1}^{\infty} \frac{(x_a)^m}{mq^m}.$$

Given $q \in \mathbb{C}$ and $\boldsymbol{n} \in \mathbb{R}^D$, in order to recover T, we need the coherent shift given by

$$[q]_{\boldsymbol{n}} = \Big(\frac{\boldsymbol{n}}{q}, \frac{\boldsymbol{n}^{\odot 2}}{2q^2}, \frac{\boldsymbol{n}^{\odot 3}}{3q^3}, \cdots\Big).$$

In fact, considering $[q]_{\boldsymbol{n}}$ as a semi-infinite vector of time perturbations of the time variables t, we get

$$[q]_{\boldsymbol{n}}(\boldsymbol{x}) = \sum_{m=1}^{\infty} \frac{1}{mq^m}(\boldsymbol{n}\cdot\boldsymbol{x})^m = -\log\Big(1 - \frac{\boldsymbol{n}\cdot\boldsymbol{x}}{q}\Big).$$

Consequently, for the shifted times $t' = t \pm [q]_{\boldsymbol{n}}$ we find that

$$\begin{aligned} \exp(t'(\boldsymbol{x}))) &= \exp(t(\boldsymbol{x}) \pm [q]_{\boldsymbol{n}}(\boldsymbol{x})) = \exp(t(\boldsymbol{x}))\exp\Big(\mp\log\Big(1 - \frac{\boldsymbol{n}\cdot\boldsymbol{x}}{q}\Big)\Big) \\ &= \exp(t(\boldsymbol{x}))\exp\Big(\log\Big(\Big(1 - \frac{\boldsymbol{n}\cdot\boldsymbol{x}}{q}\Big)^{\mp 1}\Big)\Big) \\ &= \Big(1 - \frac{\boldsymbol{n}\cdot\boldsymbol{x}}{q}\Big)^{\mp 1}\exp(t(\boldsymbol{x})), \end{aligned}$$

which immediately leads to the identification

$$\begin{aligned}\mathrm{d}\mu_{t\pm[q]_{\boldsymbol{n}}}(\boldsymbol{x}) &= \Big(1-\frac{\boldsymbol{n}\cdot\boldsymbol{x}}{q}\Big)^{\mp 1}\mathrm{d}\mu_t(\boldsymbol{x})\\ &= -q^{\pm 1}\mathrm{d}(T_{\boldsymbol{n}}^{\mp 1}\mu_t(\boldsymbol{x})).\end{aligned}$$

Proposition 4.3.1. *The Miwa shifts can be constructed as follows*

$$\big([q]_{\boldsymbol{n}}\big)_{[k]} = \frac{\boldsymbol{n}^{\odot k}}{kq^k} = \frac{1}{k}\mathcal{M}_{[k]}\chi_{[k]}\Big(\frac{\boldsymbol{n}}{q}\Big).$$

Proof. Use Proposition 2.2.1. □

For each level of times we consider the corresponding nabla or gradient operators

$$\nabla_{[k]} = \begin{pmatrix}\dfrac{\partial}{\partial t_{\alpha_1^{(k)}}}\\ \vdots \\ \dfrac{\partial}{\partial t_{\alpha_{|[k]|}^{(k)}}}\end{pmatrix}$$

and also the normal derivatives

$$\frac{\partial}{\partial \boldsymbol{n}^{\odot k}} := \langle \boldsymbol{n}^{\odot k}, \nabla_{[k]}\rangle^{(k)}.$$

Then, the Miwa shifts are modeled by the left factor of the vertex type operator

$$\exp(t(\boldsymbol{x}))\exp\Big(\sum_{k=1}^{\infty}\frac{1}{kq^n}\frac{\partial}{\partial \boldsymbol{n}^{\odot k}}\Big).$$

4.4. Bilinear equations

We begin with the following observation

Proposition 4.4.1. *Wave matrices evaluated at different times* $(t,\boldsymbol{m})$ *and* $(t',\boldsymbol{m}')$ *fulfill*

$$W_1(t,\boldsymbol{m})\big(W_1(t',\boldsymbol{m}')\big)^{-1} = W_2(t,\boldsymbol{m})\big(W_2(t',\boldsymbol{m}')\big)^{-1}.$$

Proof. From Proposition 4.1.2 we have

$$W_1(t,\boldsymbol{m})G = W_2(t,\boldsymbol{m}), \qquad W_1(t',\boldsymbol{m}')G = W_2(t',\boldsymbol{m}'),$$

for the same initial moment matrix G, from where the result follows immediately. □

Lemma 4.4.1. *We have*

$$\int_{\mathbb{T}^D(\boldsymbol{r})}\chi(\boldsymbol{z})\chi^*(\boldsymbol{z})^\top\mathrm{d}z_1\cdots\mathrm{d}z_D = \int_{\mathbb{T}^D(\boldsymbol{r})}\chi^*(\boldsymbol{z})\chi(\boldsymbol{z})^\top\mathrm{d}z_1\cdots\mathrm{d}z_D = (2\pi\,\mathrm{i})^D\mathbb{I}.$$

Proof. Observe that

$$\chi(\chi^*)^\top = \begin{pmatrix} Z_{[0],[0]} & Z_{[0],[1]} & \dots \\ Z_{[1],[0]} & Z_{[1],[1]} & \dots \\ \vdots & \vdots & \end{pmatrix},$$

$$Z_{[k],[\ell]} := \frac{1}{z_1 \cdots z_D} \begin{pmatrix} \boldsymbol{z}^{\boldsymbol{k}_1-\boldsymbol{\ell}_1} & \boldsymbol{z}^{\boldsymbol{k}_1-\boldsymbol{\ell}_2} & \dots & \boldsymbol{z}^{\boldsymbol{k}_1-\boldsymbol{\ell}_{|[\ell]|}} \\ \boldsymbol{z}^{\boldsymbol{k}_2-\boldsymbol{\ell}_1} & \boldsymbol{z}^{\boldsymbol{k}_2-\boldsymbol{\ell}_2} & \dots & \boldsymbol{z}^{\boldsymbol{k}_2-\boldsymbol{\ell}_{|[\ell]|}} \\ \vdots & \vdots & & \vdots \\ \boldsymbol{z}^{\boldsymbol{k}_{|[k]|}-\boldsymbol{\ell}_1} & \boldsymbol{z}^{\boldsymbol{k}_{|[k]|}-\boldsymbol{\ell}_2} & \dots & \boldsymbol{z}^{\boldsymbol{k}_{|[k]|}-\boldsymbol{\ell}_{|[\ell]|}} \end{pmatrix}.$$

If we now integrate in the polydisk distinguished border $\mathbb{T}^D(\boldsymbol{r})$ using the Fubini theorem we factor each integral in a product of D factors, where the i-th factor is an integral over z_i on the circle centered at origin of radius r_i. This is zero unless the integrand is z_i^{-1} which occurs only in the principal diagonal. □

Lemma 4.4.2. *Given two semi-infinite matrices U and V we have*

$$\begin{aligned} UV &= \frac{1}{(2\pi\,\mathrm{i})^D} \int_{\mathbb{T}^D(\boldsymbol{r})} U\chi(\boldsymbol{z})\big(V^T\chi^*(\boldsymbol{z})\big)^\top \mathrm{d}z_1\cdots \mathrm{d}z_D \\ &= \frac{1}{(2\pi\,\mathrm{i})^D} \int_{\mathbb{T}^D(\boldsymbol{r})} U\chi^*(\boldsymbol{z})\big(V^T\chi(\boldsymbol{z})\big)^\top \mathrm{d}z_1\cdots \mathrm{d}z_D. \end{aligned}$$

Proof. Use Lemma 4.4.1. □

Theorem 4.4.1. *For any pair of times $(t,\boldsymbol{m})$ and $(t',\boldsymbol{m}')$, points $\boldsymbol{r}_1 \in \mathscr{D}^*_{\alpha_j^{(\ell)}}(t',\boldsymbol{m}')$ and $\boldsymbol{r}_2 \in \mathscr{D}_{\alpha_i^{(k)}}(t,\boldsymbol{m})$ in the respective domains of convergence and D-dimensional tori $\mathbb{T}^D(\boldsymbol{r}_1)$ and $\mathbb{T}^D(\boldsymbol{r}_2)$ (Shilov borders of polydisks) we can ensure that Baker and adjoint Baker functions satisfy the following bilinear identity*

$$\begin{aligned} &\int_{\mathbb{T}^D(\boldsymbol{r}_1)} (\Psi_1)_{\alpha_i^{(k)}}(\boldsymbol{z},t,\boldsymbol{m})(\Psi_1^*)_{\alpha_j^{(\ell)}}(\boldsymbol{z},t',\boldsymbol{m}')\mathrm{d}z_1\cdots \mathrm{d}z_D \\ &= \int_{\mathbb{T}^D(\boldsymbol{r}_2)} (\Psi_2)_{\alpha_i^{(k)}}(\boldsymbol{z},t,\boldsymbol{m})(\Psi_2^*)_{\alpha_j^{(\ell)}}(\boldsymbol{z},t',\boldsymbol{m}')\mathrm{d}z_1\cdots \mathrm{d}z_D. \end{aligned}$$

Proof. We give two different proofs:

- First proof: Use Proposition 4.4.1, Lemma 4.4.2 and Definition 4.2.1 to get the result.
- Second proof: For any couple of set of times $(t,\boldsymbol{m})$ and $(t',\boldsymbol{m}')$ and

$$\boldsymbol{z} \in \big(\mathscr{D}_{\alpha_i^{(k)}}(t,\boldsymbol{m}) \cap \mathscr{D}^*_{\alpha_j^{(\ell)}}(t',\boldsymbol{m}')\big) \setminus \operatorname{supp}(\mathrm{d}\mu)$$

to study

$$\int_{\Omega} \mathrm{d}\mu_t(\boldsymbol{y}) P_{\alpha_i^{(k)}}(\boldsymbol{y},t,\boldsymbol{m}) P_{\alpha_j^{(\ell)}}(\boldsymbol{y},t',\boldsymbol{m}'),$$

we can use the Fubini and the integral Cauchy formula—recalling that we are dealing with domains of holomorphy—in each of the two factors to get

$$\begin{aligned}
&\sum_{j'=1}^{|[\ell]|} (H^{-1}(t',\boldsymbol{m}'))_{\alpha_j^{(\ell)},\alpha_{j'}^{(\ell)}} \int_{\mathbb{T}^D(\boldsymbol{r}_1)} \mathrm{d}z_1 \cdots \mathrm{d}z_D \int_{\Omega} \mathrm{d}\mu(\boldsymbol{y}) \frac{P_{\alpha_i^{(k)}}(\boldsymbol{z},t,\boldsymbol{m}) P_{\alpha_{j'}^{(\ell)}}(\boldsymbol{y},t',\boldsymbol{m}')}{(z_1-y_1)\cdots(z_D-y_D)} \\
&\quad \times \mathrm{e}^{t(\boldsymbol{z})} \Big[\prod_{a=1}^{D} \big((\boldsymbol{n}_a\cdot \boldsymbol{z}) - q_a\big)^{m_a} \Big] \\
&\quad = \sum_{j'=1}^{|[\ell]|} (H^{-1}(t',\boldsymbol{m}'))_{\alpha_j^{(\ell)},\alpha_{j'}^{(\ell)}} \int_{\mathbb{T}^D(\boldsymbol{r}_2)} \mathrm{d}z_1 \cdots \mathrm{d}z_D \\
&\qquad \times \int_{\Omega} \mathrm{d}\mu(\boldsymbol{y}) \frac{P_{\alpha_i^{(k)}}(\boldsymbol{y},t,\boldsymbol{m}) P_{\alpha_{j'}^{(\ell)}}(\boldsymbol{z},t',\boldsymbol{m}')}{(z_1-y_1)\cdots(z_D-y_D)} \mathrm{e}^{t(\boldsymbol{y})} \Big[\prod_{a=1}^{D} \big((\boldsymbol{n}_a\cdot \boldsymbol{y}) - q_a\big)^{m_a} \Big],
\end{aligned}$$

from where the bilinear identify follows. □

4.5. Toda type integrable equations

We explore now the nonlinear partial differential equations satisfied by the quasi-tau matrices and the β matrices.

Proposition 4.5.1. *The following relations hold true*

$$\begin{aligned}
\frac{\partial H_{[k]}}{\partial t_a} H_{[k]}^{-1} &= \beta_{[k]}(\Lambda_a)_{[k-1],[k]} - (\Lambda_a)_{[k],[k+1]}\beta_{[k+1]}, \qquad (4.5.1)\\
H_{[k+1]}\big[(\Lambda_a)_{[k],[k+1]}\big]^{\top} H_{[k]}^{-1} &= -\frac{\partial \beta_{[k+1]}}{\partial t_a}.
\end{aligned}$$

Proof. See Appendix D.14. □

For $k=0$ we have

$$-(\Lambda_a)_{[0],[1]}\beta_{[1]} = \frac{\partial H_{[0]}}{\partial t_a} H_{[0]}^{-1}. \qquad (4.5.2)$$

Observe also that

$$(J_a)_{[k+1],[k]} = -\frac{\partial \beta_{[k+1]}}{\partial t_a},$$

which agrees with

$$\Big(H^{-1}_{[k+1]}(J_a)_{[k+1],[k]}H_{[k]}\Big)^\top = (J_a)_{[k],[k+1]}.$$

Theorem 4.5.1. *The quasi-tau matrices $H_{[k]}$ are subject to the following 2D Toda lattice type equations*

$$\begin{aligned}\frac{\partial}{\partial t_b}\Big(\frac{\partial H_{[k]}}{\partial t_a}H^{-1}_{[k]}\Big) &= (\Lambda_a)_{[k],[k+1]}H_{[k+1]}\big[(\Lambda_b)_{[k],[k+1]}\big]^\top H^{-1}_{[k]}\\ &\quad - H_{[k]}\big[(\Lambda_b)_{[k-1],[k]}\big]^\top H^{-1}_{[k-1]}(\Lambda_a)_{[k-1],[k]},\end{aligned} \tag{4.5.3}$$

which can be rewritten for the rectangular matrices $\beta_{[k]}$ as

$$\begin{aligned}\frac{\partial^2 \beta_{[k]}}{\partial t_a \partial t_b} &= -\frac{\partial}{\partial t_a}\Big(\beta_{[k]}(\Lambda_b)_{[k-1],[k]}\beta_{[k]}\Big) + \frac{\partial \beta_{[k]}}{\partial t_a}\beta_{[k-1]}(\Lambda_b)_{[k-2],[k-1]}\\ &\quad + (\Lambda_b)_{[k],[k+1]}\beta_{[k+1]}\frac{\partial \beta_{[k]}}{\partial t_a}.\end{aligned}$$

We also have the following partial differential–difference equations

$$\begin{aligned}\Delta_b\Big(\frac{\partial H_{[k]}}{\partial t_a}H^{-1}_{[k]}\Big) &= (\Lambda_a)_{[k],[k+1]}H_{[k+1]}[(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})_{[k],[k+1]}]^\top\big(T_bH_{[k]}\big)^{-1}\\ &\quad - H_{[k]}[(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})_{[k-1],[k]}]^\top\big(T_bH_{[k-1]}\big)^{-1}(\Lambda_a)_{[k-1],[k]},\end{aligned}$$

and

$$\begin{aligned}\frac{\partial}{\partial t_a}\big((\Delta_bH_{[k]})H^{-1}_{[k]}\big) &= (\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})_{[k],[k+1]}H_{[k+1]}[(\Lambda_a)_{[k],[k+1]}]^\top(H_{[k]})^{-1}\\ &\quad - (T_bH_{[k]})[(\Lambda_a)_{[k-1],[k]}]^\top\big(T_bH_{[k-1]}\big)^{-1}(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})_{[k-1],[k]}.\end{aligned}$$

Notice that (4.5.3) resembles the non-Abelian Toda lattice discussed in [89, Chapter 5, §3]. However, in this case we have two main differences: in the first place the varying size of the matrices and in the second place we also have the *connectors* $(\Lambda_a)_{[k],[k+1]}$, $(\Lambda_a)_{[k-1],[k]}$ and their transpositions connecting different sized matrices.

Following the ideas of Theorem 3.2.1

Definition 4.5.1. We introduce

$$[\Lambda]_k := \begin{pmatrix}(\Lambda_1)_{[k],[k+1]}\\ \vdots\\ (\Lambda_D)_{[k],[k+1]}\end{pmatrix} \in \mathbb{R}^{D|[k]|\times|[k+1]|},$$

the gradient operator

$$[\nabla H]_k := \begin{pmatrix} \frac{\partial H_{[k]}}{\partial t_1} \\ \vdots \\ \frac{\partial H_{[k]}}{\partial t_D} \end{pmatrix} \in \mathbb{R}^{D|[k]|\times|[k]|},$$

and

$$\beta_{[k]} \otimes \mathbb{I}_D = \operatorname{diag}(\underbrace{\beta_{[k]},\dots,\beta_{[k]}}_{D \text{ times}}) \in \mathbb{R}^{D|[k]|\times D|[k-1]|}.$$

The matrix $[\Lambda]_k$ has full column rank and therefore, see Appendix B.1, the correlation matrix $[\Lambda]_k^\top[\Lambda]_k \in \mathbb{R}^{|[k+1]|\times|[k+1]|}$ is invertible and the Moore–Penrose pseudo-inverse $[\Lambda]_k^+ = \big([\Lambda]_k^\top[\Lambda]_k\big)^{-1}[\Lambda]_k^\top \in \mathbb{R}^{|[k+1]|\times D|[k]|}$ is the left inverse $[\Lambda]_k^+[\Lambda]_k = \mathbb{I}_{[k+1]}$.

Proposition 4.5.2. *The β matrices are subject to the following recurrence*

$$\beta_{[k+1]} = -[\Lambda]_k^+[\nabla H]_k H_{[k]}^{-1} + [\Lambda]_k^+(\beta_{[k]}\otimes\mathbb{I}_D)[\Lambda]_{k-1}, \qquad \beta_{[1]} = -(\nabla H_{[0]})H_{[0]}^{-1}.$$

Proof. The proof follows immediately from (4.5.1), (4.5.2) and the fact that $[\Lambda]_0 = \mathbb{I}_D$. $\square$

Iterating once and twice the above result we get

$$\begin{aligned}
\beta_{[k+1]} &= -[\Lambda]_k^+[\nabla H]_k H_{[k]}^{-1} - [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+[\nabla H]_{k-1}H_{[k-1]}^{-1}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1} \\
&\quad + [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+(\beta_{[k-1]}\otimes\mathbb{I}_D)[\Lambda]_{k-2}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1} \\
&= -[\Lambda]_k^+[\nabla H]_k H_{[k]}^{-1} - [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+[\nabla H]_{k-1}H_{[k-1]}^{-1}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1} \\
&\quad - [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+\Big(\Big[[\Lambda]_{k-2}^+[\nabla H]_{k-2}H_{[k-2]}^{-1}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-2}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1} \\
&\quad + [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+\Big(\Big[[\Lambda]_{k-2}^+(\beta_{[k-2]}\otimes\mathbb{I}_D)[\Lambda]_{k-3}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-2}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1},
\end{aligned}$$

respectively. By induction we deduce the following

Proposition 4.5.3. *In terms of logarithmic right derivatives of the quasi-tau matrices the β matrices are expressed by*

$$\begin{aligned}
\beta_{[k+1]} &= -[\Lambda]_k^+[\nabla H]_k H_{[k]}^{-1} - [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+[\nabla H]_{k-1}H_{[k-1]}^{-1}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1} + \cdots \\
&- [\Lambda]_k^+\Big(\Big[[\Lambda]_{k-1}^+\Big(\Big[[\Lambda]_{k-2}^+\cdots\Big(\Big[[\Lambda]_1^+[\nabla H]_0)H_{[0]}^{-1}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_1^+\cdots\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-2}\Big]\otimes\mathbb{I}_D\Big)[\Lambda]_{k-1}.
\end{aligned}$$

In the 1D scenario the above formula simplifies and gives the classical τ-expressions for β_k. Indeed, now

$$\beta_{k+1} = -\frac{\partial \log H_k}{\partial t_1} - \frac{\partial \log H_{k-1}}{\partial t_1} - \cdots - \frac{\partial \log H_0}{\partial t_1}$$

and from $H_k = \frac{\tau_{k+1}}{\tau_k}$ we get a telescopical series giving

$$\beta_{k+1} = -\frac{\partial \log \tau_{k+1}}{\partial t_1}.$$

5. KP type equations via congruences

We study how the previous construction leads to families of nonlinear partial differential–difference equations involving a fixed site, say the k-th position, in the lattice and therefore not mixing several sites in the lattice—notice that in the Toda type equations derived before, see Theorems 3.1.1 and 4.5.1, we are faced with relations involving three contiguous sites, $k-1$, k and $k+1$. We refer the reader to [72,87,85,86,110].

5.1. The congruence technique

Let us first introduce some notation

Definition 5.1.1. Given two semi-infinite matrices $R_1(t, \boldsymbol{m})$ and $R_2(t, \boldsymbol{m})$ we say that

- $R_1(t) \in \mathfrak{l}W_0$ if $R_1(t)\big(W_0(t, \boldsymbol{m})\big)^{-1}$ is a block strictly lower triangular matrix.
- $R_2(t) \in \mathfrak{u}$ if it is a block upper triangular matrix.

Then, we can state the following *congruences* [85] or *asymptotic module* [71] style result

Proposition 5.1.1. *Given two semi-infinite matrices $R_1(t, \boldsymbol{m})$ and $R_2(t, \boldsymbol{m})$ such that*

- $R_1(t, \boldsymbol{m}) \in \mathfrak{l}W_0(t, \boldsymbol{m})$,
- $R_2(t, \boldsymbol{m}) \in \mathfrak{u}$,
- $R_1(t, \boldsymbol{m})G = R_2(t, \boldsymbol{m})$.

Then

$$R_1(t, \boldsymbol{m}) = 0, \qquad R_2(t, \boldsymbol{m}) = 0.$$

Proof. Observe that

$$
\begin{aligned}
R_2(t,\boldsymbol{m}) &= R_1(t,\boldsymbol{m})G = R_1(t,\boldsymbol{m})\big(W_1(t,\boldsymbol{m})\big)^{-1}W_1(t,\boldsymbol{m})G \\
&= R_1(t,\boldsymbol{m})\big(W_1(t,\boldsymbol{m})\big)^{-1}W_2(t,\boldsymbol{m}),
\end{aligned}
$$

where we have used (4.1.3). From here we get

$$
R_1(t,\boldsymbol{m})\big(W_0(t,\boldsymbol{m})\big)^{-1}\big(S(t,\boldsymbol{m})\big)^{-1} = R_2(t,\boldsymbol{m})\Big(H(t,\boldsymbol{m})\big((S(t,\boldsymbol{m}))^{-1}\big)^{\top}\Big)^{-1},
$$

and, as in the LHS we have a strictly lower triangular matrix while in the RHS we have an upper triangular matrix, both sides must vanish and the result follows. □

We use the congruence notation

Definition 5.1.2. When $A - B \in \mathfrak{l}W_0$ we write $A = B + \mathfrak{l}W_0$ and if $A - B \in \mathfrak{u}$ we write $A = B + \mathfrak{u}$.

We introduce the following notation

$$
\partial_a = \frac{\partial}{\partial t_a}, \quad \partial_{(a,b)} = \frac{\partial}{\partial t_{\boldsymbol{e}_a+\boldsymbol{e}_b}}, \quad \partial_{(a,b,c)} = \frac{\partial}{\partial t_{\boldsymbol{e}_a+\boldsymbol{e}_b+\boldsymbol{e}_c}} \qquad a,b,c = 1,\dots,D, \tag{5.1.1}
$$

and the normal derivative

$$
\frac{\partial}{\partial \boldsymbol{n}_a} = \sum_{b=1}^{D} n_{a,b}\partial_a.
$$

Notice that we have employed the round bracket notation for the subindexes of the higher times. In fact, this is convenient as it reflects the invariance under the action of the symmetric group on the letters in the labels, for example $t_{(a,b)} = t_{(b,a)}$.

5.2. Connecting discrete a continuous flows

To begin with let us show the following "asymptotic" behaviors

Proposition 5.2.1. *We have*

$$
\partial_b W_1 = (\Lambda_b + \beta\Lambda_b)W_0 + \mathfrak{l}W_0, \tag{5.2.1}
$$

$$
T_a W_1 = \big((\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}) + (T_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}) - q_a\big)W_0 + \mathfrak{l}W_0. \tag{5.2.2}
$$

Proof. From (4.1.2) we get

$$
\begin{aligned}
\partial_b W_1 &= (\partial_b S + S\Lambda_b)W_0 \\
&= (\Lambda_b + \beta\Lambda_b)W_0 + \mathfrak{l}W_0,
\end{aligned}
$$

$$\begin{aligned}T_aW_1 &= (T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)W_0\\ &= \big((\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})+(T_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})-q_a\big)W_0+\mathfrak{l}W_0. \qquad \square\end{aligned}$$

This immediately leads to the following finding, translations and derivations are almost the same thing when acting on the Baker functions

Proposition 5.2.2. *The Baker functions Ψ_1, Ψ_2 are both solutions of the following difference–differential linear system*

$$\frac{\partial\Psi}{\partial\boldsymbol{n}_a}=T_a\Psi+\Big(q_a-(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\Big)\Psi.$$

Proof. From Proposition 5.2.1 we get that

$$\Big(\frac{\partial}{\partial\boldsymbol{n}_a}-T_a\Big)W_1=\big(q_a-(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\big)W_0+\mathfrak{l}W_0,$$

and we easily conclude that

$$\Big(\frac{\partial}{\partial\boldsymbol{n}_a}-T_a-q_a+(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\Big)W_1=\mathfrak{l}W_0.$$

As $\big(\frac{\partial}{\partial\boldsymbol{n}_a}-T_a-q_a-(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}\big)W_2\in\mathfrak{u}$ from Proposition 5.1.1, with $R_i=\big(\frac{\partial}{\partial\boldsymbol{n}_a}-T_a-q_a+(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\big)W_i$, $i=1,2$, we deduce

$$\Big(\frac{\partial}{\partial\boldsymbol{n}_a}-T_a-q_a+(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\Big)W_1=0$$

and the result follows. $\square$

For MVOPR we have

Proposition 5.2.3. *The MVOPR satisfy*

$$\frac{\partial P_{[k]}}{\partial\boldsymbol{n}_a}=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)\Delta_aP_{[k]}-(\Delta_a\beta)_{[k]}(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[k+1],[k]}P_{[k]}.$$

Proof. Introduce the form of the Baker function Ψ_1 given in Definition 4.2.1 into Proposition 5.2.2. $\square$

The compatibility of the linear systems satisfied by the Baker functions implies

Theorem 5.2.1. *The following equation for β holds*

$$\begin{aligned}\Delta_b\left[\frac{\partial\beta}{\partial\boldsymbol{n}_a}+(\Delta_a\beta)\big(q_a+(\boldsymbol{n_a}\cdot\boldsymbol{\Lambda})\beta\big)\right]\boldsymbol{n_b}\cdot\boldsymbol{\Lambda}\\ =\Delta_a\left[\frac{\partial\beta}{\partial\boldsymbol{n}_b}+(\Delta_b\beta)\big(q_b+(\boldsymbol{n_b}\cdot\boldsymbol{\Lambda})\beta\big)\right]\boldsymbol{n_a}\cdot\boldsymbol{\Lambda}.\end{aligned}\tag{5.2.3}$$

Proof. See Appendix D.15. □

A remarkable fact regarding (5.2.3) is that, when written componentwise, it only involves the rectangular matrix $\beta_{[k]} \in \mathbb{R}^{|[k]|\times|[k-1]|}$ and no other, for example near neighbors $\beta_{[k\pm 1]}$, as happens for Toda type equations.

5.3. Second order flows

We introduce the diagonal matrices $V_{a,b} = \operatorname{diag}((V_{a,b})_{[0]}, (V_{a,b})_{[1]}, (V_{a,b})_{[2]}, \dots)$

$$V_{a,b} := \frac{\partial \beta}{\partial t_a}\Lambda_b, \quad (V_{a,b})_{[k]} = \frac{\partial \beta_{[k]}}{\partial t_a}(\Lambda_b)_{[k-1],[k]}, \quad U_{a,b} := -V_{a,b} - V_{b,a}. \tag{5.3.1}$$

Proposition 5.3.1. *Both Baker functions Ψ_1 and Ψ_2 are solutions of*

$$\frac{\partial \Psi}{\partial t_{(a,b)}} = \frac{\partial^2 \Psi}{\partial t_a \partial t_b} + U_{a,b}\Psi. \tag{5.3.2}$$

Proof. On the one hand, from (4.1.2) we find

$$\begin{aligned}\partial_{(a,b)} W_1 &= (\partial_{a,b} S + S\Lambda_a\Lambda_b)W_0,\\ \partial_a\partial_b W_1 &= (\partial_a\partial_b S + \partial_a S\Lambda_b + \partial_b S\Lambda_a + S\Lambda_a\Lambda_b)W_0\end{aligned}$$

and therefore $(\partial_{(a,b)} - \partial_a\partial_b)W_1 = -(\partial_a S\Lambda_b + \partial_b S\Lambda_a)W_0 + \mathfrak{l}W_0$ so that

$$\Big(\partial_{a,b} - \partial_a\partial_b + V_{a,b} + V_{b,a}\Big)W_1 \in \mathfrak{l}W_0.$$

On the other hand, it is obvious that

$$\Big(\partial_{(a,b)} - \partial_a\partial_b + V_{a,b} + V_{b,a}\Big)W_2 \in \mathfrak{u}.$$

Now, we apply Proposition 5.1.1 with

$$R_i = \Big(\partial_{(a,b)} - \partial_a\partial_b + V_{a,b} + V_{b,a}\Big)W_i, \qquad i = 1, 2,$$

to get the result. □

Observe that for $a = b$ (5.3.2) reads

$$\frac{\partial \Psi_{[k]}}{\partial t_a^{(2)}} = \frac{\partial^2 \Psi_{[k]}}{\partial t_a^2} + (U_a)_{[k]}\Psi_{[k]}, \qquad t^{(2)} := t_{(a,a)}, \qquad U_a = -2V_{a,a},$$

which is a time dependent one-dimensional Schrödinger type equation for the square matrices $\Psi_{[k]}$, the wave functions, and potential of the square matrix $(U_a)_{[k]}$. Moreover, multidimensional matrix Schrödinger equations appear if we look to other directions, thus given $(a_1,\dots,a_d)\subset\{1,\dots,D\}$, $a_1<\dots<a_d$ we can look at the second order time flow generated by $\frac{\partial}{\partial t}:=\frac{\partial}{\partial t_{a_1,a_1}}+\dots+\frac{\partial}{\partial t_{a_d,a_d}}$ to get in terms of the d-dimensional nabla operator $\nabla:=(\frac{\partial}{\partial t_{a_1}},\dots,\frac{\partial}{\partial t_{a_d}})^\top$, Laplacian $\Delta:=\nabla^2=\frac{\partial^2}{\partial t_{a_1}^2}+\dots+\frac{\partial^2}{\partial t_{a_d}^2}$ and matrix potential $U:=U_{a_1,a_1}+\dots+U_{a_d,a_d}=2\nabla(\beta)\cdot\mathbf{\Lambda}$

$$\frac{\partial\Psi_{[k]}}{\partial t}=\Delta\Psi_{[k]}+U_{[k]}\Psi_{[k]}.$$

Corollary 5.3.1. *The MVOPR satisfy*

$$\begin{aligned}\frac{\partial P_{[k]}}{\partial t_{(a,b)}}(\boldsymbol{x})&=\frac{\partial^2 P_{[k]}}{\partial t_a\partial t_b}(\boldsymbol{x})+x_a\frac{\partial P_{[k]}}{\partial t_b}(\boldsymbol{x})+x_b\frac{\partial P_{[k]}}{\partial t_a}(\boldsymbol{x})\\&\quad-\Big(\frac{\partial\beta_{[k]}}{\partial t_a}(\Lambda_b)_{[k-1],[k]}+\frac{\partial\beta_{[k]}}{\partial t_b}(\Lambda_a)_{[k-1],[k]}\Big)P_{[k]}(\boldsymbol{x}).\end{aligned}$$

Proof. Just introduce expressions for the Baker functions in Proposition 4.2.1 in the previous proposition. □

We see that again only k-th site of the lattice is involved in these linear equations and, consequently, its compatibility will lead to equations for the coefficients evaluated at that site. These nonlinear equation for which $\beta_{[k]}$ is a solution are

Theorem 5.3.1. *The following nonlinear partial differential equation*

$$\begin{aligned}&\partial_{(c,d)}(\partial_a\beta\Lambda_b+\partial_b\beta\Lambda_a)-\partial_{(a,b)}(\partial_c\beta\Lambda_d+\partial_d\beta\Lambda_c)\\&\quad=\partial_a\partial_b(\partial_c\beta\Lambda_d+\partial_d\beta\Lambda_c)-\partial_c\partial_d(\partial_a\beta\Lambda_b+\partial_b\beta\Lambda_a)\\&\qquad+(\partial_b\partial_c\beta)(\Lambda_d\beta\Lambda_a-\Lambda_a\beta\Lambda_d)+(\partial_b\partial_d\beta)(\Lambda_c\beta\Lambda_a-\Lambda_a\beta\Lambda_c)\\&\qquad+(\partial_a\partial_c\beta)(\Lambda_d\beta\Lambda_b-\Lambda_b\beta\Lambda_d)+(\partial_a\partial_d\beta)(\Lambda_c\beta\Lambda_b-\Lambda_b\beta\Lambda_c)\\&\qquad+\big[\partial_a\beta\Lambda_b+\partial_b\beta\Lambda_a,\partial_c\beta\Lambda_d+\partial_d\beta\Lambda_c\big]\end{aligned}\tag{5.3.3}$$

is satisfied for $a,b,c,d\in\{1,\dots,D\}$.

Proof. See Appendix D.16. □

Observe that this equation decouples giving for each k the same equation (5.3.3) up to the replacements $\beta\to\beta_{[k]}$ and $\Lambda_A\to(\Lambda_A)_{[k-1],[k]}$, $k=1,2,\dots$ and $A=a,b,c,d$. For the particular case $a=b=A$ and $c=d=B$ and using the notation $t_A=x$, $t_B=y$, $t_{(A,A)}=$ and $t_{B,B}=t$ we get

$$\begin{aligned}&\frac{\partial^2\beta}{\partial t\partial x}\Lambda_A-\frac{\partial\beta}{\partial s\partial y}\Lambda_B\\&\quad=\frac{\partial^3\beta}{\partial x^2\partial y}\Lambda_B-\frac{\partial^3\beta}{\partial x\partial y^2}\Lambda_A+2\Big[\frac{\partial\beta}{\partial x}\Lambda_A,\frac{\partial\beta}{\partial y}\Lambda_B\Big]+2\frac{\partial^2\beta}{\partial x\partial y}(\Lambda_B\beta\Lambda_A-\Lambda_A\beta\Lambda_B).\end{aligned}$$

5.4. Exploring third order flows

Associated with the third order times $t_{(a,b,c)}$ we introduce the following block diagonal matrices

$$V_{a,b,c} = \operatorname{diag}((V_{a,b,c})_{[0]}, (V_{a,b,c})_{[1]}, (V_{a,b,c})_{[2]}, \dots)$$

with

$$\begin{aligned} V_{a,b,c} &:= \frac{\partial \beta}{\partial t_a}[\beta, \Lambda_b]\Lambda_c, \\ &= \frac{\partial \beta^{(2)}}{\partial t_a}\Lambda_b\Lambda_c - \frac{\partial \beta}{\partial t_a}\Lambda_b\beta\Lambda_c, \\ (V_{a,b,c})_{[k]} &= \frac{\partial \beta_{[k]}}{\partial t_a}\Big(\beta_{[k-1]}\big(\Lambda_b\big)_{[k-2],[k-1]} - \big(\Lambda_b\big)_{[k-1],[k]}\beta_{[k]}\Big)\big(\Lambda_c\big)_{[k-1],[k]}, \\ &= \frac{\partial \beta^{(2)}_{[k]}}{\partial t_a}\big(\Lambda_b\big)_{[k-2],[k-1]} - \frac{\partial \beta_{[k]}}{\partial t_a}\big(\Lambda_b\big)_{[k-1],[k]}\beta_{[k]}\big(\Lambda_c\big)_{[k-1],[k]}. \end{aligned}$$

Observe the use of Proposition 4.1.4; we remark that $(V_{a,b,c})_{[k]}$ depends on $\beta_{[k]}$ and $\beta^{(2)}_{[k]}$ only, coefficients of the MVOPR for the second and third higher degree monomials, $P_{[k]}(\boldsymbol{x}) = \chi_{[k]}(\boldsymbol{x}) + \beta_{[k]}\chi_{[k-1]}(\boldsymbol{x}) + \beta^{(2)}_{[k]}\chi_{[k-2]}(\boldsymbol{x}) + \dots + \beta^{(k)}_{[k]}$. If we insist in using only the second higher total degree coefficient and not the third higher total degree coefficient there is price we must pay, now we involve two polynomials $P_{[k]}$ and $P_{[k-1]}$—as we require of $\beta_{[k]}$ and $\beta_{[k-1]}$. Then

Proposition 5.4.1. *The Baker functions Ψ_1 and Ψ_2 are both solutions of the third order linear differential equations*

$$\begin{aligned} \frac{\partial \Psi}{\partial t_{(a,b,c)}} = &\frac{\partial^3 \Psi}{\partial t_a \partial t_b \partial t_c} - V_{a,b}\frac{\partial \Psi}{\partial t_c} - V_{c,a}\frac{\partial \Psi}{\partial t_b} - V_{b,c}\frac{\partial \Psi}{\partial t_a} \\ &- \Big(\frac{\partial V_{a,b}}{\partial t_c} + \frac{\partial V_{b,c}}{\partial t_a} + \frac{\partial V_{c,a}}{\partial t_b} + V_{a,b,c} + V_{b,c,a} + V_{c,b,a}\Big)\Psi. \end{aligned}$$

Proof. See Appendix D.19. □

For $a = b = c$ in terms of $t^{(3)}_a := t_{(a,a,a)}$ and

$$\tilde{U}_a := -3\Big(\frac{\partial^2 \beta}{\partial t_a^2}\Lambda_a + \frac{\partial \beta^{(2)}}{\partial t_a}\Lambda_a^2 - \frac{\partial \beta}{\partial t_a}\Lambda_a\beta\Lambda_a\Big),$$

the Baker functions satisfy

$$\frac{\partial \Psi}{\partial t^{(3)}_a} = \frac{\partial^3 \Psi}{\partial t_a^3} + \frac{3}{2}U_a\frac{\partial \Psi}{\partial t_a} + \tilde{U}_a\Psi.$$

Therefore, in terms of the differential operators

$$
\begin{aligned}
\mathcal{L}_a &:= \frac{\partial^2}{\partial t_a^2} + U_a, & U_a &= -2\frac{\partial \beta}{\partial t_a}\Lambda_a,\\
\mathcal{P}_b &:= \frac{\partial^3}{\partial t_b^3} + \frac{3}{2}U_b\frac{\partial}{\partial t_a} + \tilde{U}_b, & \tilde{U}_b &= -3\frac{\partial^2\beta}{\partial t_b^2}\Lambda_b - 3\frac{\partial \beta^{(2)}}{\partial t_b}\Lambda_b^2 + 3\frac{\partial\beta}{\partial t_b}\Lambda_b\beta\Lambda_b,
\end{aligned}
\tag{5.4.1}
$$

the wave matrix fulfills

$$
\frac{\partial W_1}{\partial t_a^{(2)}} = \mathcal{L}_a(W_1), \qquad \frac{\partial W_1}{\partial t_b^{(3)}} = \mathcal{P}_b(W_1).
$$

Hence, the following compatibility equations hold

$$
R_{ab}(W_1) = 0, \qquad R_{ab} := \frac{\partial \mathcal{L}_a}{\partial t_b^{(3)}} - \frac{\partial \mathcal{P}_b}{\partial t_a^{(2)}} + \big[\mathcal{L}_a, \mathcal{P}_b\big].
$$

Observe that

$$
\begin{aligned}
R_{ab} = {} & -3\frac{\partial U_a}{\partial t_b}\frac{\partial^2}{\partial t_b^2} + 3\frac{\partial U_b}{\partial t_a}\frac{\partial^2}{\partial t_a \partial t_b} + 2\frac{\partial \tilde{U}_b}{\partial t_a}\frac{\partial}{\partial t_a}\\
& + \Big(-\frac{3}{4}\frac{\partial U_b}{\partial t_a^{(2)}} + \frac{3}{2}\frac{\partial^2 U_b}{\partial t_a^2} - 3\frac{\partial^2 U_a}{\partial t_b^2} + \frac{3}{2}[U_a, U_b]\Big)\frac{\partial}{\partial t_b}\\
& + \frac{1}{2}\frac{\partial U_a}{\partial t_b^{(3)}} - \frac{1}{2}\frac{\partial \tilde{U}_b}{\partial t_a^{(2)}} + \frac{\partial^2 \tilde{U}_b}{\partial t_a^2} - \frac{\partial^3 U_a}{\partial t_b^3} - \frac{3}{2}U_b\frac{\partial U_a}{\partial t_b} + [U_a, \tilde{U}_b]
\end{aligned}
$$

is a matrix second order partial differential operator in t_a, t_b. Its action on the wave matrix can be split into diagonals. $\big(R_{ab}(W)\big)W_0^{-1}$ has all its superdiagonals above the second superdiagonal identically zero.[13] Consequently, we request to the main diagonal and to the first and second superdiagonals to cancel.

Proposition 5.4.2. *The matrices β and $\beta^{(2)}$ are subject to*

$$
\begin{aligned}
\frac{\partial^2\beta^{(2)}}{\partial t_a\partial t_b}\Lambda_a &= \frac{\partial}{\partial t_b}\left[\frac{\partial\beta}{\partial t_a}\Lambda_a\beta - \frac{1}{2}\frac{\partial^2\beta}{\partial t_a^2} + \frac{1}{4}\frac{\partial\beta}{\partial t_a^{(2)}}\right],\\
0 &= 3\frac{\partial^2}{\partial t_b^2}\left[\frac{1}{2}\frac{\partial\beta}{\partial t_a^{(2)}} - \frac{\partial^2\beta}{\partial t_a^2} + 2\frac{\partial\beta}{\partial t_a}\Lambda_a\beta\right]\Lambda_b\\
&\quad + \frac{\partial}{\partial t_a}\left[2\frac{\partial^3\beta}{\partial t_b^3} - \frac{\partial\beta}{\partial t_b^{(3)}} + \left(\frac{\partial\beta}{\partial t_b}\Lambda_b\beta - \frac{\partial\beta^{(2)}\Lambda_b}{\partial t_b}\right)\Lambda_b\beta\right]\Lambda_a
\end{aligned}
$$

[13] It is a consequence of $\partial_a W_1 = (\Lambda_a + \beta\Lambda_a)W_0 + \mathfrak{u}W_0$ and $\partial_a\partial_b W_1 = \big(\Lambda_a\Lambda_b + \beta\Lambda_a\Lambda_b + \frac{\partial\beta}{\partial t_a}\Lambda_b + \frac{\partial\beta}{\partial t_b}\Lambda_a + \beta^{(2)}\Lambda_a\Lambda_b\big)W_0 + \mathfrak{u}W_0$.

$$+3\frac{\partial}{\partial t_b}\left[\left(2\frac{\partial\beta}{\partial t_a}\Lambda_a\beta^{(2)}+\frac{1}{2}\frac{\partial\beta^{(2)}}{\partial t_a^{(2)}}-\frac{\partial^2\beta^{(2)}}{\partial t_a^2}\right)\Lambda_b^2-2\frac{\partial\beta}{\partial t_b}\Lambda_b\frac{\partial\beta}{\partial t_a}\Lambda_a\right]$$
$$+3\frac{\partial\beta}{\partial t_b}\Lambda_b\left[\frac{\partial^2\beta}{\partial t_a\partial t_b}-2\frac{\partial\beta}{\partial t_a}\Lambda_a\beta-\frac{1}{2}\frac{\partial\beta}{\partial t_a}\right]\Lambda_b-6\frac{\partial^2\beta}{\partial t_b\partial t_a}\Lambda_b\beta^{(2)}\Lambda_a\Lambda_b.$$

6. Linear isometry invariant measures and MVOPR

In this section we consider orthogonal transformations $R\in\mathrm{O}(\mathbb{R}^D)$; i.e., linear isometries $R:\mathbb{R}^D\to\mathbb{R}^D$ preserving the dot or scalar product: $R\boldsymbol{u}\cdot R\boldsymbol{v}=\boldsymbol{u}\cdot\boldsymbol{v}$, $\forall\boldsymbol{u},\boldsymbol{v}\in\mathbb{R}^D$. For the matrix $[R]_B$ in the canonical basis $B=\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_D\}$ of $\mathbb{R}^D$ of the orthogonal endomorphism means $[R]_B^\top=[R]_B^{-1}$. Given such an orthogonal transformation $\boldsymbol{x}\to R\boldsymbol{x}$, we assume the linear *isometry* invariance condition $\mathrm{d}\mu(\boldsymbol{x})=\mathrm{d}\mu(R\boldsymbol{x})$.[14]

6.1. Symmetric powers of a linear isometry. Orthonormal basis and biorthogonal systems

What is the action of this linear isometry in the set of MVOPR? or putting it in other equivalent terms, how do it act of the corresponding symmetric tensor powers? Given any set of linear transformations $\{f_i\}_{i=1}^m\in\operatorname{End}(\mathbb{R}^D)$ one can construct a map $f_1\odot\cdots\odot f_m\in\operatorname{End}\left(\left(\mathbb{R}^D\right)^{\odot k}\right)$ such that

$$(f_1\odot\cdots\odot f_m)(\boldsymbol{u}_1\odot\cdots\odot\boldsymbol{u}_m)=f_1(\boldsymbol{u}_1)\odot\cdots\odot f_m(\boldsymbol{u}_m),\qquad\forall\boldsymbol{u}_i\in\mathbb{R}^D.$$

In this manner we introduce the k-th symmetric power of the endomorphism R acting on symmetric tensor powers, $R^{\odot k}\in\operatorname{End}\left(\left(\mathbb{R}^D\right)^{\odot k}\right)$ and, moreover, a diagonal block endomorphism in the symmetric algebra, $\mathscr{R}:=\operatorname{diag}(1,\mathscr{R}_{[1]},\mathscr{R}_{[2]},\dots)=1\oplus\mathscr{R}_{[1]}\oplus\mathscr{R}_{[2]}\oplus\cdots\in\operatorname{End}\left(\mathrm{S}(\mathbb{R}^D)\right)$, with its diagonal blocks given by $\mathscr{R}_{[k]}:=R^{\odot k}$. For a given invertible endomorphism R, with inverse R^{-1}, the corresponding endomorphism $\mathscr{R}$ in the symmetric algebra is invertible with inverse $(\mathscr{R}^{-1})_{[k]}=(R^{-1})^{\odot k}$. The $\mathrm{S}^k(\mathbb{R}^D)$ is equipped with a natural scalar product $\langle\cdot,\cdot\rangle^{(k)}$ given in terms of permanents, see Appendix A.2.2. In particular, for decomposable symmetric tensors (A.2.2)

$$\begin{aligned}\langle R^{\odot k}(\boldsymbol{u}_1\odot\cdots\odot\boldsymbol{u}_k),R^{\odot k}(\boldsymbol{v}_1\odot\cdots\odot\boldsymbol{v}_k)\rangle^{(k)}&=\frac{1}{k!}\operatorname{perm}\begin{pmatrix}R\boldsymbol{u}_1\cdot R\boldsymbol{v}_1&\cdots&R\boldsymbol{u}_1\cdot R\boldsymbol{v}_k\\ \vdots&&\vdots\\ R\boldsymbol{u}_k\cdot R\boldsymbol{v}_1&\cdots&R\boldsymbol{u}_k\cdot R\boldsymbol{v}_k\end{pmatrix}\\&=\frac{1}{k!}\operatorname{perm}\begin{pmatrix}\boldsymbol{u}_1\cdot\boldsymbol{v}_1&\cdots&\boldsymbol{u}_1\cdot\boldsymbol{v}_k\\ \vdots&&\vdots\\ \boldsymbol{u}_k\cdot\boldsymbol{v}_1&\cdots&R\boldsymbol{u}_k\cdot\boldsymbol{v}_k\end{pmatrix}\\&=\langle\boldsymbol{u}_1\odot\cdots\odot\boldsymbol{u}_k),\boldsymbol{v}_1\odot\cdots\odot\boldsymbol{v}_k\rangle^{(k)}.\end{aligned}$$

[14] The measure μ is said to be invariant under R if for every measurable set $A\subset\mathbb{R}^D$ we have $\mu\left(R^{-1}(A)\right)=\mu(A)$.

Thus, the k-th tensor power $R^{\odot k}$ is an orthogonal transformation in $\{\mathrm{S}^k(\mathbb{R}^D), \langle\cdot,\cdot\rangle^{(k)}\}$. At this point we stress that care must be taken when we express this fact in terms of matrices. Observe that the *canonical* basis $\{\boldsymbol{e}^{\alpha_i}\}_{i=1}^{|[k]|}$, with $\boldsymbol{e}^{\alpha_i} = \boldsymbol{e}_1^{\odot\alpha_{i,1}} \odot \cdots \odot \boldsymbol{e}_D^{\odot\alpha_{i,D}}$, is an orthogonal set for $\langle\cdot,\cdot\rangle^{(k)}$, but is not orthonormal. In fact, we know that, see Appendix A.2.2, $\|\boldsymbol{e}^{\alpha_i}\|^2 = \binom{k}{\alpha_i}^{-1}$, the metric matrix in this basis being the inverse of the multinomial matrix. Let us find the matrix representing $R^{\odot k}$ in the canonical basis $B_c = \{\boldsymbol{e}^{\alpha_i}\}_{i=1}^{|[k]|}$; we proceed to compute

$$
\begin{aligned}
R^{\odot k}\boldsymbol{e}^{\alpha_i} &= (R\boldsymbol{e}_1)^{\odot\alpha_{i,1}} \odot \cdots \odot (R\boldsymbol{e}_D)^{\odot\alpha_{i,D}} \\
&= \Big(\sum_{j=1}^{D} R_{j_1,1}\boldsymbol{e}_j\Big)^{\odot\alpha_{i,1}} \odot \cdots \odot \Big(\sum_{j=1}^{D} R_{j_1,1}\boldsymbol{e}_j\Big)^{\odot\alpha_{i,D}} \\
&= \Big[\sum_{|\alpha'|=\alpha_{i,1}} \binom{\alpha_{i,1}}{\alpha'} \prod_{j=1}^{D} R_{j,1}^{\alpha'_j}\boldsymbol{e}^{\alpha'}\Big] \odot \cdots \odot \Big[\sum_{|\alpha'|=\alpha_{i,D}} \binom{\alpha_{i,D}}{\alpha'} \prod_{j=1}^{D} R_{j,D}^{\alpha'_j}\boldsymbol{e}^{\alpha'}\Big] \\
&= \sum_{j=1}^{|[k]|} (\mathscr{R}_{[k]})_{\alpha_j,\alpha_i}\boldsymbol{e}^{\alpha_j},
\end{aligned}
$$

which gives the matrix $[\mathscr{R}_{[k]}]_{B_c} = \big[(\mathscr{R}_{[k]})_{\alpha_i,\alpha_j}\big]$. Then, as the transformation preserves the scalar product with metric matrix given by $\mathcal{M}_{[k]}^{-1}$, the matrix in the canonical basis of the k-th symmetric tensor power satisfies

$$
[\mathscr{R}_{[k]}]_{B_c}^{\top}\mathcal{M}_{[k]}^{-1}[\mathscr{R}_{[k]}]_{B_c} = \mathcal{M}_{[k]}^{-1}. \tag{6.1.1}
$$

Instead of the canonical basis B_c we could consider the orthonormal linear basis $B = \{\boldsymbol{u}_i\}_{i=1}^{|[k]|}$ with the normalized vectors $\boldsymbol{u}_i = \Big(\sqrt{\binom{k}{\alpha_i}}\Big)^{-1}\boldsymbol{e}^{\alpha_i}$. In this basis the matrix for $R^{\odot k}$ is $[\mathscr{R}_{[k]}]_B = \mathcal{M}_{[k]}^{-1/2}[\mathscr{R}_{[k]}]_{B_c}\mathcal{M}_{[k]}^{1/2}$, which happens to be an orthogonal matrix; i.e., $[\mathscr{R}]_B^{\top} = [\mathscr{R}]_B^{-1} = [\mathscr{R}^{-1}]_B$. Another orthogonal basis is $\tilde{B}_c = \{\tilde{\boldsymbol{e}}^{\alpha_i}\}_{i=1}^{|[k]|}$, with $\tilde{\boldsymbol{e}}^{\alpha_i} = \binom{k}{\alpha_i}^{-1}\boldsymbol{e}^{\alpha_i}$, that despite not being orthonormal, forms with the canonical basis a biorthogonal system, i.e. $\langle\tilde{\boldsymbol{e}}^{\alpha_i}, \boldsymbol{e}^{\alpha_j}\rangle^{(k)} = \delta_{i,j}$. The matrix representing $\mathscr{R}_{[k]}$ in this orthogonal basis is

$$
\begin{aligned}
\eta_{R,[k]} := [\mathscr{R}_{[k]}]_{\tilde{B}_c} &= \mathcal{M}_{[k]}^{-1}[\mathscr{R}_{[k]}]_{B_c}\mathcal{M}_{[k]} \\
&= \mathcal{M}_{[k]}^{-1/2}[\mathscr{R}_{[k]}]_{B}\mathcal{M}_{[k]}^{1/2},
\end{aligned}
$$

and in the symmetric algebra we have the corresponding block diagonal matrix

$$
\eta_R = \mathcal{M}^{-1}[\mathscr{R}]_{B_c}\mathcal{M}, \tag{6.1.2}
$$

which satisfies $\eta_R^{\top}\mathcal{M}\eta_R = \mathcal{M}$. Here $[\mathscr{R}]_{B_c} := \operatorname{diag}([\mathscr{R}_{[0]}]_{B_c}, [\mathscr{R}_{[1]}]_{B_c}, \dots)$ and $[\mathscr{R}]_B := \operatorname{diag}([\mathscr{R}_{[0]}]_B, [\mathscr{R}_{[1]}]_B, \dots)$. Observe that

$$
\begin{aligned}
\eta_{[R]^\top}=\eta_{R^{-1}}=\eta_R^{-1}&=\mathcal{M}^{-1/2}[\mathscr{R}^{-1}]_B\mathcal{M}^{1/2}=\mathcal{M}^{-1/2}[\mathscr{R}]_B^\top\mathcal{M}^{1/2}=(\mathcal{M}^{1/2}[\mathscr{R}]_B\mathcal{M}^{-1/2})^\top\\
&=\mathcal{M}^{-1}(\mathcal{M}^{-1/2}[\mathscr{R}]_B\mathcal{M}^{1/2})^\top\mathcal{M}\\
&=\mathcal{M}^{-1}\eta_R^\top\mathcal{M},
\end{aligned}
$$

and also that, as B_c and $\tilde{B}_c$ form a biorthogonal system, we have the following relations

$$
\eta_R^\top=[\mathscr{R}]_{B_c}^{-1},\quad \eta_R^{-1}=[\mathscr{R}]_{B_c}^\top. \tag{6.1.3}
$$

6.2. Applications to monomials and shift matrices

We now see how the above developments apply to the monomials χ and the shift matrices Λ introduced previously.

Proposition 6.2.1. *We have*

$$
\begin{aligned}
&\chi(R\boldsymbol{x})=\eta_R\chi(\boldsymbol{x}),\\
&R\boldsymbol{n}\cdot\boldsymbol{\Lambda}=\eta_R(\boldsymbol{n}\cdot\boldsymbol{\Lambda})\eta_R^{-1}.
\end{aligned}\tag{6.2.1}
$$

Proof. To prove the first relation notice that form (2.2.1) we get

$$
\begin{aligned}
\chi_{[k]}(R\boldsymbol{x})=&\mathcal{M}_{[k]}^{-1}[(R\boldsymbol{x})^{\odot k}]_{B_c}=\mathcal{M}_{[k]}^{-1}[\mathscr{R}_{[k]}]_{B_c}[\boldsymbol{x}^{\odot k}]_{B_c}=\mathcal{M}_{[k]}^{-1}[\mathscr{R}_{[k]}]_{B_c}\mathcal{M}_{[k]}\chi_{[k]}(\boldsymbol{x})\\
=&\eta_{R,[k]}\chi_{[k]}(\boldsymbol{x}).
\end{aligned}
$$

For the second formula we observe that

$$
\begin{aligned}
\eta_R(\boldsymbol{n}\cdot\boldsymbol{\Lambda})\eta_R^{-1}\chi(\boldsymbol{x})=&\eta_R(\boldsymbol{n}\cdot\boldsymbol{\Lambda})\chi(R^{-1}\boldsymbol{x})=(\boldsymbol{n}\cdot R^{-1}\boldsymbol{x})\eta_R\chi(R^{-1}x)=(R\boldsymbol{n}\cdot\boldsymbol{x})\chi(\boldsymbol{x})\\
=&(R\boldsymbol{n}\cdot\boldsymbol{\Lambda})\chi(\boldsymbol{x}),
\end{aligned}
$$

which holds $\forall\boldsymbol{x}\in\mathbb{R}^D$ so that the result follows. □

When $R\in\mathfrak{S}_D\subset\mathrm{O}(\mathbb{R}^D)$ we also have $\chi^*(R\boldsymbol{x})=\eta_R\chi^*(\boldsymbol{x})$. We know that Λ_a has no left inverse but it does have a right inverse, its transpose, $\Lambda_a\Lambda_a^\top=\mathbb{I}$. In this paper we have derived a number of results with $\boldsymbol{n}\cdot\boldsymbol{\Lambda}$ and sometimes, for example in Proposition 3.5.6 and Proposition 3.5.13, it is useful to find the right inverse of $\boldsymbol{n}\cdot\boldsymbol{\Lambda}$.

Proposition 6.2.2. *Given any vector $\boldsymbol{n}\in\mathbb{R}^D$ find $R\in\mathrm{O}(\mathbb{R}^D)$ such that $R\boldsymbol{e}_a=\boldsymbol{n}$. Then, a right inverse of $(\boldsymbol{n}\cdot\boldsymbol{\Lambda})$ is $\eta_R\Lambda_a^\top\eta_R^{-1}$; i.e.,*

$$
(\boldsymbol{n}\cdot\boldsymbol{\Lambda})\eta_R\Lambda_a^\top\eta_R^{-1}=\mathbb{I}.
$$

Proof. From (6.2.1) we know that $\boldsymbol{n}\cdot\boldsymbol{\Lambda}=(R\boldsymbol{e}_a)\cdot\boldsymbol{\Lambda}=\eta_R\Lambda_a\eta_R^{-1}$ but the right inverse of this matrix is $\eta_R\Lambda_a^\top\eta_R^{-1}$ and the result is proven. □

In our opinion a nicer result, as it does not depend on any alien isometry R, is

Proposition 6.2.3. *A right inverse for $(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}$, $k>0$, is*

$$\begin{aligned}&\mathcal{M}_{[k]}^{-1/2}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}\big)^{+}\\&\quad=\mathcal{M}_{[k]}^{-1}((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]})^{\top}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1}((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]})^{\top})\big)^{-1}.\end{aligned}$$

Proof. See Appendix D.17. □

6.3. Consequences of the measure symmetry

Proposition 6.3.1. *Whenever, for a given orthogonal endomorphism $R\in\mathrm{O}(\mathbb{R}^D)$, there is a symmetry in the measure of the form $\mathrm{d}\mu(\boldsymbol{x})=\mathrm{d}\mu(R\boldsymbol{x})$ we have*

(1) *The moment matrix satisfies*

$$\eta_R G\eta_R^\top=G.$$

(2) *The factors of the Cholesky factorization (2.3.2) are such that*

$$\eta_R S\eta_R^{-1}=S,\qquad \eta_R H\eta_R^\top=H. \tag{6.3.1}$$

(3) *Moreover,*

$$\eta_{R,[k]}\beta_{[k]}=\beta_{[k]}\eta_{R,[k-1]}.$$

Proof. See Appendix D.18. □

Observe that while we can write for $[\eta_R,S]=0$ for the quasi-tau matrices we can write $[\eta_R,H\mathcal{M}]=0$ $(\big(\eta_R^\top\big)^{-1}=\mathcal{M}\eta_R\mathcal{M}^{-1}=[\mathscr{R}]_{B_c}^{-1})$.

Now, we are ready to deduce how the MVOPR of a symmetric measure behaves under the symmetry of the measure

Proposition 6.3.2. *Let us assume that for an orthogonal transformation $R\in\mathrm{O}(\mathbb{R}^D)$ the measure satisfies $\mathrm{d}\mu(\boldsymbol{x})=\mathrm{d}\mu(R\boldsymbol{x})$, then:*

(1) *The MVOPR fulfill*

$$P\big(Rx\big)=\eta_R P(\boldsymbol{x}). \tag{6.3.2}$$

(2) *The Jacobi matrices are such that*

$$R\boldsymbol{n}\cdot\boldsymbol{J}=\eta_R(\boldsymbol{n}\cdot\boldsymbol{J})\eta_R^{-1}. \tag{6.3.3}$$

(3) *The Christoffel–Darboux kernel remains invariant*

$$K^{(\ell)}(R\boldsymbol{x},R\boldsymbol{y})=K^{(\ell)}(\boldsymbol{x},\boldsymbol{y}). \tag{6.3.4}$$

Proof. See Appendix D.20. □

When $R\in\mathfrak{S}_D$, as $x_1\cdots x_D$ is invariant under permutation of coordinates, we also have $C\big(R\boldsymbol{x}\big)=\eta_R C(\boldsymbol{x})$ and $Q^{(\ell)}(R\boldsymbol{x},R\boldsymbol{y})=Q^{(\ell)}(\boldsymbol{x},\boldsymbol{y})$.

6.4. Compatible Toda flows

We now request that the symmetry $\mathrm{d}\mu(\boldsymbol{x})=\mathrm{d}\mu(R\boldsymbol{x})$ is preserved under the integrable deformations discussed previously. As before we distinguish two cases, the discrete case and the continuous case. For the first situation we have

Proposition 6.4.1. *If $\boldsymbol{n}$ is invariant under the transformation $R\in\mathrm{O}(\mathbb{R}^D)$, $\boldsymbol{n}=R\boldsymbol{n}$, then the corresponding discrete transformation preserves the isometry invariance of the measure*

$$T\mathrm{d}\mu(\boldsymbol{x})=T\mathrm{d}\mu(R\boldsymbol{x}).$$

Proof. The new measure is $T\mathrm{d}\mu(\boldsymbol{x})=(\boldsymbol{n}\cdot\boldsymbol{x}-q)\mathrm{d}\mu(\boldsymbol{x})$ so that

$$\begin{aligned}T\mathrm{d}\mu(R\boldsymbol{x})&=(\boldsymbol{n}\cdot R\boldsymbol{x}-q)\mathrm{d}\mu(R\boldsymbol{x})\\&=(\boldsymbol{n}\cdot R\boldsymbol{x}-q)\mathrm{d}\mu(\boldsymbol{x})\\&=(R^\top\boldsymbol{n}\cdot\boldsymbol{x}-q)\mathrm{d}\mu(\boldsymbol{x}).\end{aligned}$$

Therefore when $\boldsymbol{n}=R^\top\boldsymbol{n}$, as $R^\top=R^{-1}$ we get that the new measure is invariant. □

For the continuous flows recall Definition 4.1.3 and Proposition 4.1.1 and notice that if we order the times $t=(t_{[0]},t_{[1]},\dots)$, $t_{[k]}=(t_{\alpha_1},\dots,t_{\alpha_{|[k]|}})$

Proposition 6.4.2. *The continuous Toda flows preserve the symmetry $\mathrm{d}\mu(\boldsymbol{x})=\mathrm{d}\mu(R\boldsymbol{x})$ whenever the times are such that*

$$t_{[k]}\eta_{R,[k]}=t_{[k]}.$$

Proof. As we know from Proposition 4.1.1

$$
\begin{aligned}
\mathrm{d}\mu_t(R\boldsymbol{x}) &= \exp\big(t\chi(R\boldsymbol{x})\big)\mathrm{d}\mu(R\boldsymbol{x})\\
&= \exp\big(t\eta_R\chi(\boldsymbol{x})\big)\mathrm{d}\mu(\boldsymbol{x})
\end{aligned}
$$

and when $t\eta_R = t$ we get $\mathrm{d}\mu_t(R\boldsymbol{x}) = \mathrm{d}\mu_t(\boldsymbol{x})$. $\square$

6.4.1. Linear subspaces of fixed points of a linear isometry

We have seen that to analyze compatible flows with the linear isometry invariance it is crucial to find fixed points for the linear isometries.

Definition 6.4.1. The linear subspace of fixed points of the linear isometry R is

$$V_R = \{\boldsymbol{v} \in \mathbb{R}^D : R\boldsymbol{v} = \boldsymbol{v}\}.$$

Interesting examples of linear isometries are provided by reflections. Given a nonzero vector $\boldsymbol{n}$ the corresponding Householder reflection is

$$r_{\boldsymbol{n}} = \mathbb{I}_D - 2\frac{1}{\boldsymbol{n}\cdot\boldsymbol{n}}\boldsymbol{n}\boldsymbol{n}^\top.$$

This is a reflection in the hyperplane $\boldsymbol{n}^\perp$, it is idempotent $r_{\boldsymbol{n}}^2 = \mathbb{I}_D$, $r_{\boldsymbol{n}}\big|_{\mathbb{R}\boldsymbol{n}} = -\operatorname{id}$, (negating any vector component parallel to $\boldsymbol{n}$), and $r_{\boldsymbol{n}}\big|_{\boldsymbol{n}^\perp} = \operatorname{id}$. Therefore, for the Householder case $V_R = \boldsymbol{n}^\perp$. Any orthogonal matrix $R \in \mathrm{O}(\mathbb{R}^D)$ is as a product of at most D Householder reflections. Given an orthogonal set $\{\boldsymbol{n}_1, \dots, \boldsymbol{n}_m\} \subset \mathbb{R}^D$ the product R of the corresponding Householder reflections (which happens to be commutative as the order of the factors does not affect the result) is $R = \mathbb{I}_D - 2\sum_{i=1}^m \frac{1}{\boldsymbol{n}_i\cdot\boldsymbol{n}_i}\boldsymbol{n}_i\boldsymbol{n}_i^\top$. This is a reflection, with reflection hyperplane $\{\boldsymbol{n}_1, \dots, \boldsymbol{n}_m\}^\perp$, negating the components parallel to $\mathbb{R}\{\boldsymbol{n}_1, \dots, \boldsymbol{n}_m\}$. Now, the fixed point subspace $V_R = \{\boldsymbol{n}_1, \dots, \boldsymbol{n}_m\}^\perp$ is the reflection plane.

For $D = 2$ the orthogonal transformations could be of two types, a rotation of angle θ, $R_\theta = \left(\begin{smallmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{smallmatrix}\right)$ and a reflection according to the vector $\boldsymbol{n} = \begin{pmatrix}-\sin(\theta/2)\\ \cos(\theta/2)\end{pmatrix}$, $R = \mathbb{I}_2 - 2\boldsymbol{n}\boldsymbol{n}^\top = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$ with reflection line given by $\mathbb{R}\begin{pmatrix}\cos(\theta/2)\\ \sin(\theta/2)\end{pmatrix}$. For $\theta = \pi/2$ the reflection is about the line $y = x$ and therefore exchanges x and y: it is a permutation matrix $\left(\begin{smallmatrix}0 & 1\\ 1 & 0\end{smallmatrix}\right)$.

In general, given a linear isometry $R \in \mathrm{O}(\mathbb{R}^D)$ there exists an orthonormal basis $\{\boldsymbol{u}_a\}_{a=1}^D$ such that its matrix reads

$$R = \operatorname{diag}(\underbrace{1, \dots, 1}_{p \text{ of them}}, \underbrace{-1, \dots, -1}_{q \text{ of them}}, R_{\theta_1}, \dots, R_{\theta_m},)$$

$p-q$ being even or odd depending whether D is even or odd, here R_θ is a two-dimensional nontrivial rotation of angle θ. Therefore $V_R = \mathbb{R}\{\boldsymbol{u}_1, \dots, \boldsymbol{u}_p\}$; notice that for D even it could happen that $V_R = \{0\}$, but for D odd we always have a nontrivial fixed point subspace, $\dim V_R \geq 1$.

6.4.2. Secant varieties of Veronese varieties and linear isometry invariance preserving flows

Proposition 6.4.3. *If the times $t^\top \in \operatorname{Sym}(\mathbb{R}^D)$ are restricted to belong to the symmetric algebra of the fixed point subspace of the linear isometry R, i.e. $t^\top \in \operatorname{Sym}(V_R) \subset \operatorname{Sym}(\mathbb{R}^D)$, the linear isometry invariance condition of the measure is preserved as the time passes by,* $\mathrm{d}\mu_t(\boldsymbol{x}) = \mathrm{d}\mu_t(R\boldsymbol{x})\ \forall t^\top \in \operatorname{Sym}(V_R)$.

Proof. Observing that $\eta_R^\top = [\mathscr{R}]_{B_c}$, the linear isometry invariance preserving condition can be written as

$$[R^{\odot k}]_{B_c} t_{[k]}^\top = t_{[k]}^\top.$$

The first nontrivial condition is that

$$R[t]_1^\top = t_{[1]} \qquad [t]_{[1]}^\top = \begin{pmatrix} t_1 \\ \vdots \\ t_D \end{pmatrix}$$

and therefore $[t]_{[1]}^\top \in V_R$.

In order to explore what kind of higher flows will preserve the linear isometry invariance condition of the measure we observe that if $V_{[1]} = \{\boldsymbol{t} \in \mathbb{R}^D : R\boldsymbol{t} = \boldsymbol{t}\}$ any symmetric power tensor in $V_{[k]} = (V_{[1]})^{\odot k}$ will be subspace of fixed point for $R^{\odot k}$. Indeed, $V_{[k]}$ is linearly generated by decomposable symmetric tensors $\boldsymbol{v}_1 \odot \cdots \odot \boldsymbol{v}_k$ with $\boldsymbol{v}_i \in V_{[1]}$ and

$$\begin{aligned} R^{\odot k}(\boldsymbol{v}_1 \odot \cdots \odot \boldsymbol{v}_k) &= (R\boldsymbol{v}_1) \odot \cdots \odot (R\boldsymbol{v}_k) \\ &= \boldsymbol{v}_1 \odot \cdots \odot \boldsymbol{v}_k. \qquad \square \end{aligned}$$

The map $\mathbb{R}^m \to (\mathbb{R}^m)^{\odot k}$ taking $\boldsymbol{v} \to \boldsymbol{v}^{\odot k}$ has as its image the Veronese variety $\mathscr{V}_{m,k} := \{\boldsymbol{x}^{\odot k} \in (\mathbb{R}^m)^{\odot k} : \boldsymbol{x} \in \mathbb{R}^m\}$. It happens [35] that every symmetric power tensor can be written for some $r \geq 0$ as $\sum_{i=1}^r \boldsymbol{v}_i^{\odot k}$ and the symmetric tensor rank is the minimum when this holds. Hence, the symmetric power can be described by r points of the Veronese variety; the closure of the union of all linear spaces spanned by r points of the Veronese variety $\mathscr{V}_{m,k}$ is called the $(r-1)$-th secant variety of $\mathscr{V}_{m,k}$. Consequently, the times are constrained to belong to one the secant varieties of the Veronese variety $\mathscr{V}_{\dim(V_R),k}$.

Appendix A. Compositions, multisets and symmetric algebras

A.1. Compositions and multisets

From combinatorics [108] we know that a weak D-composition of an integer k is a way of writing k as the sum of D non-negative integers. Notice that while for a composition we require the parts to be positive integers (excluding therefore the zero) for weak compositions the zero is allowed. The problem of counting the number $N(k,D)$ of weak compositions, i.e. the cardinality of the set $[k]$, is related to the problem of counting the number of compositions, which is $\binom{k-1}{D-1}$. In fact, given a weak D-composition $k = k_{i,1}+\cdots+k_{i,D}$ if we put $q_{i,j}=k_{i,j}+1$, $j\in\{1,\dots,D\}$ we have $q_{i,1}+\cdots+q_{i,D}=k+D$ and we are dealing with a $(k+D)$-composition. Thus $N(k,D)=|[k]|=\binom{k+D-1}{D-1}=\binom{k+D-1}{k}$—consider all the possible permutations of $k+(D-1)$ elements out of which k and $(D-1)$ are repeated. Two sequences that differ in the order of their terms define different weak compositions of their sum, while they are considered to define the same partition of that number. Every integer has finitely many distinct compositions.

A multiset [27] is 2-tuple (I,M) where I is some set, the underlying set of elements, and the multiplicity $M: I\to\mathbb{N}$ is a function from I to the set of positive integers; for each $a\in I$ the multiplicity or number of occurrences is $M(a)$. For an indexed family, (a_i), where i is some index-set, we define a multiset $\{a_i\}$, where the multiplicity of any element a is the number of indices i such that $a_i=a$. A form of describing a multiset that is used in this article is considering non-negative integers $(a_i)_{i=1}^k$ such that $1\leq a_1\leq\cdots\leq a_k\leq D$, where repetitions are allowed, e.g. for $k=5$ we could have $a_1=a_2=a_3<a_4=a_5$, denoting $a_1=a$ and $a_4=b$ we are dealing with the multiset $\{a,a,a,b,b\}$ being three the multiplicity of a, $M(a)=3$, and two the multiplicity of b, $M(b)=2$.

A.2. *Symmetric tensor powers and symmetric algebras*

We give a brief description of notions and results regarding symmetric algebras, for further information we refer the reader to [35,48,77,117].

A.2.1. *Symmetric tensors*

A symmetric tensor of order k is a tensor of order k that is invariant under a permutation of its vector arguments:

$$T(u_1,\dots,u_k)=\tau_\sigma T(u_1,\dots,u_k)=T(u_{\sigma 1},\dots,u_{\sigma k})$$

for every $\sigma\in\mathfrak{S}_k$, $\mathfrak{S}_k$ being the symmetric group of k letters. The coefficients of a symmetric tensor of order k satisfy $T_{i_1,\dots,i_k}=T_{i_{\sigma 1},\dots,i_{\sigma k}}$. The space of symmetric tensors of order k on $\mathbb{R}^D$ is naturally isomorphic to the dual of the space of homogeneous multivariate polynomials of total degree k and the graded vector space of all symmetric tensors can be naturally identified with the symmetric algebra $\operatorname{Sym}(\mathbb{R}^D)$.

The symmetric part of a tensor $T\in\left(\mathbb{R}^D\right)^{\otimes k}$ of order k is defined by

$$\operatorname{Sym} T=\frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k}\tau_\sigma T, \tag{A.2.1}$$

where the summation extends over the symmetric group on k symbols. If the tensor coefficients of the tensor are $T_{i_1,i_2,\dots,i_k}$, those of the symmetric part are $T_{(i_1,i_2,\dots,i_k)}=$

$\frac{1}{k!}\sum_{\sigma\in\mathfrak{S}_k} T_{i_{\sigma 1},i_{\sigma 2},\dots,i_{\sigma k}}$. For two arbitrary pure tensors $T = v_1 \otimes v_2 \otimes \cdots \otimes v_r$ the corresponding symmetrization or symmetric part is given by $v_1 \odot v_2 \odot \cdots \odot v_r \equiv \operatorname{Sym}(v_1 \otimes v_2 \otimes \cdots \otimes v_r) := \frac{1}{r!}\sum_{\sigma\in\mathfrak{S}_r} v_{\sigma 1} \otimes v_{\sigma 2} \otimes \cdots \otimes v_{\sigma r}$. Given two tensors $T_i \in \operatorname{Sym}^{k_i}(\mathbb{R}^D)$, $i \in \{1,2\}$, the symmetrization operator allows us to define $T_1 \odot T_2 = \operatorname{Sym}(T_1 \otimes T_2) \in \operatorname{Sym}^{k_1+k_2}(\mathbb{R}^D)$. As the resulting product is commutative and associative some authors write $T_1T_2 = T_1 \odot T_2$. Given a vector $v \in \mathbb{R}^D$ we will use the exponential notation $v^{\odot k} = \underbrace{v \odot v \odot \cdots \odot v}_{k \text{ times}} = \underbrace{v \otimes v \otimes \cdots \otimes v}_{k \text{ times}} = v^{\otimes k}$.

A.2.2. Symmetric tensor powers and symmetric algebra

Symmetric tensor powers $\mathrm{S}^k(\mathbb{R}^D) = (\mathbb{R}^D)^{\odot k}$ are generated by the so-called decomposable (or simple or pure) symmetric tensors $u_1 \odot \cdots \odot u_k$, where $u_1, \dots, u_k \in \mathbb{R}^D$. Given a basis $\{\boldsymbol{e}_a\}_{a=1}^D$ we can construct an explicit linear basis of $S^k(\mathbb{R}^D)$ using the concept of multiset. The mentioned linear basis for the k-th symmetric power is $\{\boldsymbol{e}_{a_1} \odot \cdots \odot \boldsymbol{e}_{a_k}\}_{\substack{1\leq a_1\leq\cdots\leq a_k\leq D\\ k\in\mathbb{Z}_+}}$, or in terms of multisets $I = \{a_1, \dots, a_p\}$ with multiplicities $M(a_i)$, such that $M(a_1) + \cdots + M(a_p) = k$ we have $\{\boldsymbol{e}_{a_1}^{\odot M(a_1)} \odot \cdots \odot \boldsymbol{e}_{a_p}^{\odot M(a_p)}\}_I$. The dual space of the symmetric powers happens to be isomorphic to the set of symmetric multilinear functionals on $\mathbb{R}^D$, $\big(\mathrm{S}^k(\mathbb{R}^D)\big)^* \cong S((\mathbb{R}^D)^k, \mathbb{R})$.

The number of multisets of cardinality k, with elements taken from a finite set of cardinality D, is known as the multiset coefficient $\left(\!\binom{D}{k}\!\right)$, see [108], which resembles the binomial coefficients and we say "D multichoose k" instead of "D choose k" for $\binom{n}{k}$. We have that $\left(\!\binom{D}{k}\!\right) = \binom{D+k-1}{k} = \frac{(D+k-1)!}{k!\,(D-1)!} = \frac{D(D+1)(D+2)\cdots(D+k-1)}{k!}$, and the number of such multisets is the same as the number of subsets of cardinality k in a set of cardinality $D + k - 1$. Thus, $\dim \mathrm{S}^k(\mathbb{R}^D) \equiv |[k]| = \left(\!\binom{D}{k}\!\right)$.

We can define a surjective map $\pi : \big(\mathbb{R}^D\big)^{\otimes k} \to \mathrm{S}(\mathbb{R}^D)$ by the symmetrization $\pi(u_1 \otimes \cdots \otimes u_k) := u_1 \odot \cdots \odot u_k$. This map has a section, i.e. an injective map $\imath : \mathrm{S}(\mathbb{R}^D) \to \big(\mathbb{R}^D\big)^{\otimes k}$ such that $\pi \circ \imath = \mathrm{id}$. The map gives $\imath(u_1 \odot \cdots \odot u_k) = \operatorname{Sym}(u_1 \otimes \cdots \otimes u_k)$ so that its image is just the space of symmetric tensors just discussed. Moreover, for the symmetrization of tensors of (A.2.1) we have $\operatorname{Sym} := \imath \circ \pi : \big(\mathbb{R}^D\big)^{\otimes k} \to \big(\mathbb{R}^D\big)^{\otimes k}$; notice also that this symmetrization is a projection $\operatorname{Sym}^2 = \operatorname{Sym}$ so that

$$\big(\mathbb{R}^D\big)^{\otimes k} = \operatorname{Sym}\Big(\big(\mathbb{R}^D\big)^{\otimes k}\Big) \oplus \ker(\operatorname{Sym}) = \imath\Big(\big(\mathbb{R}^D\big)^{\odot k}\Big) \oplus \ker(\operatorname{Sym}).$$

The direct sum $\mathrm{S}(\mathbb{R}^D) := \oplus_{k\geq 0}\mathrm{S}^k(\mathbb{R}^D)$ is the symmetric algebra of $\mathbb{R}^D$, which is commutative and associative. The symmetric algebra of $\mathbb{R}^D$ can be constructed as the tensor algebra $\mathrm{T}(\mathbb{R}^D)$ quotient with the ideal generated tensors of the form $t_1(\boldsymbol{x} \otimes \boldsymbol{y} - \boldsymbol{y} \otimes \boldsymbol{x})t_2$ with t_1, t_2 homogeneous tensors of arbitrary degree.

A.2.3. Dot product

There is an interesting inner product in the symmetric tensor power $\mathrm{S}^k(\mathbb{R}^D)$ which is a symmetric positive definite bilinear form, see appendix A in [77], and also [35]. It is given by the linear extension of the following definition for decomposable symmetric tensor powers

$$\begin{aligned}&\langle \boldsymbol{u}_1 \odot \cdots \odot \boldsymbol{u}_k, \boldsymbol{v}_1 \odot \cdots \odot \boldsymbol{v}_k \rangle^{(k)}\\ &= \frac{1}{k!} \sum_{\sigma \in \mathfrak{S}_k} \prod_{a=1}^{k} \boldsymbol{u}_a \cdot \boldsymbol{v}_{\sigma a} = \frac{1}{k!} \operatorname{perm} \begin{pmatrix} \boldsymbol{u}_1 \cdot \boldsymbol{v}_1 & \cdots & \boldsymbol{u}_1 \cdot \boldsymbol{v}_k \\ \vdots & & \vdots \\ \boldsymbol{u}_k \cdot \boldsymbol{v}_1 & \cdots & \boldsymbol{u}_k \cdot \boldsymbol{v}_k \end{pmatrix},\end{aligned} \tag{A.2.2}$$

where $\boldsymbol{u}_a, \boldsymbol{v}_a \in \mathbb{R}^D$ and we have used permanents [91,94].

We introduce the semi-infinite multinomial matrix $\mathcal{M} = \operatorname{diag}(\mathcal{M}_{[0]}, \mathcal{M}_{[1]}, \dots)$ where

$$\mathcal{M}_{[k]} := \operatorname{diag}\left(\binom{k}{\boldsymbol{\alpha}_1}, \dots, \binom{k}{\boldsymbol{\alpha}_{|[k]|}} \right) \in \mathbb{R}^{|[k]| \times |[k]|} \tag{A.2.3}$$

are diagonal matrices with coefficients the multinomial numbers

$$\binom{k}{\boldsymbol{\alpha}_j} = \frac{k!}{\prod_{a=1}^{D} \alpha_{j,a}!}, \qquad j \in \{1, \dots, |[k]|\}.$$

According to the proof of Corollary A.24 of [77] for the canonical basis vectors $\boldsymbol{e}^{\boldsymbol{\alpha}_i} = \boldsymbol{e}_1^{\odot \alpha_{i,1}} \odot \cdots \odot \boldsymbol{e}_D^{\odot \alpha_{i,D}}$, for all $\boldsymbol{\alpha}_i \in [k] := \{\alpha_{i,1}\boldsymbol{e}_1 + \cdots + \alpha_{i,D}\boldsymbol{e}_D \in \mathbb{Z}_+^D \text{ with } \alpha_{i,1} + \cdots + \alpha_{i,D} = k\}$,[15] we have the following metric coefficients

$$\langle \boldsymbol{e}_{\boldsymbol{\alpha}_i}, \boldsymbol{e}_{\boldsymbol{\alpha}_j} \rangle^{(k)} = \delta_{i,j} \binom{k}{\boldsymbol{\alpha}_i}^{-1}, \qquad i, j \in \left\{1, \dots, \left(\!\binom{D}{k}\!\right)\right\}.$$

Hence, the interior product $\langle \cdot, \cdot \rangle^{(k)} : \mathrm{S}^k(\mathbb{R}^D) \times \mathrm{S}^k(\mathbb{R}^D) \to \mathbb{R}$ is given by

$$\left\langle \sum_{i=1}^{|[k]|} a_{\boldsymbol{k}_i} \boldsymbol{e}^{\boldsymbol{k}_i}, \sum_{i=1}^{|[k]|} b_{\boldsymbol{k}_i} \boldsymbol{e}^{\boldsymbol{k}_i} \right\rangle^{(k)} = \begin{pmatrix} a_{\boldsymbol{k}_1} \\ \vdots \\ a_{\boldsymbol{k}_{|[k]|}} \end{pmatrix}^{\top} \mathcal{M}_{[k]}^{-1} \begin{pmatrix} b_{\boldsymbol{k}_1} \\ \vdots \\ b_{\boldsymbol{k}_{|[k]|}} \end{pmatrix}. \tag{A.2.4}$$

It is easy to check that

$$\begin{aligned} \langle \boldsymbol{u}_1 \odot \cdots \odot \boldsymbol{u}_k, \boldsymbol{v}^{\odot k} \rangle^{(k)} &= \prod_{a=1}^{k} (\boldsymbol{u}_a \cdot \boldsymbol{v}), \\ \langle \boldsymbol{u}^{\odot k}, \boldsymbol{v}^{\odot k} \rangle^{(k)} &= (\boldsymbol{u} \cdot \boldsymbol{v})^k. \end{aligned} \tag{A.2.5}$$

We remark that all these developments are connected with Quantum Physics. Indeed, when Quantum Mechanics of large systems describes sets of an arbitrary number of

15 In [48] the set $[k]$ is denoted by $\Xi(D, k)$.

bosons, if $\mathscr{H}$ is the Hilbert space for the states of a single particle, then $\mathrm{S}^k(\mathscr{H})$ will describe the pure states of k identical bosons, and in general the symmetric algebra $\mathrm{S}(\mathscr{H})$ is the Hilbert space of pure states of an arbitrary number of bosons, see [111]. Thus, multivariate polynomials are connected, naively if you want, with bosons such that its single particle Hilbert space of pure states is $\mathbb{R}^D$ (which is a real Hilbert space, that still has a physical meaning for even dimensions D).

A.2.4. The shift matrices and symmetric tensors

The shift matrices have a natural description in the symmetric algebra, as well.

Proposition A.2.1. *In the symmetric tensor power setting these blocks can be thought of as*

$$(\Lambda_a)_{[k-1],[k]}:\mathrm{S}^k(\mathbb{R}^D)\to\mathrm{S}^{k-1}(\mathbb{R}^D)$$

with

$$(\Lambda_a)_{[k-1],[k]}\boldsymbol{e}^{\alpha_i}=\begin{cases}\boldsymbol{e}^{\alpha_i-\boldsymbol{e}_a}, & \text{if } \alpha_i\cdot\boldsymbol{e}_a\neq 0,\\ 0, & \text{if } \alpha_i\cdot\boldsymbol{e}_a=0.\end{cases}$$

Following §1.10.3 of [48] we introduce interior multiplications. First we consider dual linear space $(\mathbb{R}^D)^*$ of linear functionals $\mathbb{R}^D\to\mathbb{R}$, and the dual basis $\{\boldsymbol{\omega}_a\}_{a=1}^D\subset(\mathbb{R}^D)^*$, i.e. $\boldsymbol{\omega}_a(\boldsymbol{e}_b)=\delta_{a,b}$. The symmetric tensors powers $\boldsymbol{\omega}^\alpha=\boldsymbol{\omega}_1^{\odot\alpha_1}\odot\cdots\odot\boldsymbol{\omega}_D^{\odot\alpha_D}$ give rise to a linear basis $\{\boldsymbol{\omega}^\alpha\}_{\alpha\in[k]}$ of $\mathrm{S}^k\big((\mathbb{R}^D)^*\big)$ dual to $\big\{\frac{1}{\alpha!}\boldsymbol{e}^\alpha\big\}_{\alpha\in[k]}$

$$\boldsymbol{\omega}^\alpha(\boldsymbol{e}^{\boldsymbol{\beta}})=\alpha!\delta_{\boldsymbol{\alpha},\boldsymbol{\beta}}.$$

Interior multiplications or right contractions are maps

$$\llcorner:\mathrm{S}^k(\mathbb{R}^D)\times\mathrm{S}^{k'}\big((\mathbb{R}^D)^*\big)\to\mathrm{S}^{k-k'}(\mathbb{R}^D),$$

with

$$\boldsymbol{e}^{\boldsymbol{\alpha}}\llcorner\boldsymbol{\omega}^{\alpha'}:=\begin{cases}0, & \alpha_a<\alpha'_a,\ \text{for some } a\in\{1,\dots,D\},\\ \frac{\boldsymbol{\alpha}!}{(\boldsymbol{\alpha}-\alpha')!}\boldsymbol{e}^{\boldsymbol{\alpha}-\alpha'}, & \alpha_a\geq\alpha'_a,\forall a\in\{1,\dots,D\}.\end{cases}$$

When we take $k'=1$ and consider the linear maps $\llcorner\boldsymbol{\omega}_a:\mathrm{S}^k(\mathbb{R}^D)\to\mathrm{S}^{k-1}(\mathbb{R}^D)$ we find

$$\boldsymbol{e}^{\boldsymbol{\alpha}}\llcorner\boldsymbol{\omega}_a:=(\boldsymbol{\alpha}\cdot\boldsymbol{e}_a)\boldsymbol{e}^{\boldsymbol{\alpha}-\boldsymbol{e}_a}.$$

In terms of the number operators[16] $N_a:\mathrm{S}^k(\mathbb{R}^D)\to\mathrm{S}^k(\mathbb{R}^D)$ given by $N_a(\sum_{\alpha\in[k]}c_\alpha\boldsymbol{e}^\alpha)=\sum_{\alpha\in[k]}(\alpha\cdot\boldsymbol{e}_a)c_\alpha\boldsymbol{e}^\alpha$ we can write

[16] If we follow the quantum interpretation of our symmetric algebra as system of bosons with the single particle described by $\mathbb{R}^D$, we could understand N_a as the number operator of particles in state $\boldsymbol{e}_a$.

$$(\Lambda_a)_{[k-1],[k]}(T) = (N_a^{-1}T) \llcorner \boldsymbol{\omega}_a, \qquad \forall T \in (\mathbb{R}^D)^{\odot k},$$

as the composition of a number operator with an interior multiplication.[17]

Appendix B. Complements of linear algebra: Pseudo-inverses, Schur complements and quasi-determinants

B.1. The Moore–Penrose pseudo-inverse

As this paper requires pseudo-inverses a number of times we decided to include a short resume on the subject, for more information see [22]. Given a rectangular matrix $M \in \mathbb{R}^{m\times n}$, its Moore–Penrose pseudo-inverse $M^+ \in \mathbb{R}^{n\times m}$ [93,26,98] is a generalized inverse, i.e.,

$$MM^+M = M, \qquad M^+MM^+ = M^+,$$

that in addition satisfies

$$(MM^+)^\top = MM^+, \qquad (M^+M)^\top = M^+M.$$

Let us say that any matrix has a generalized inverse and a unique pseudo-inverse. Obviously, if M is invertible then $M^+ = M^{-1}$, any rectangular zero matrix has as it pseudo-inverse its transpose. The pseudo-inverse operation is idempotent $(M^+)^+ = M$ and $(M^\top)^+ = (M^+)^\top$.

The square matrices $P := MM^+ \in \mathbb{R}^{m\times m}$ and $Q := M^+M \in \mathbb{R}^{n\times n}$ are orthogonal projection operators, i.e. $P = P^\top$ and $Q = Q^\top$, $P^2 = P$, $Q^2 = Q$. Moreover, we have

(1) $PM = M = MQ$ and $M^+P = M^+ = QM^+$.
(2) $\operatorname{Im}(M) = \ker(M)^\perp = \operatorname{Im}(P) = \ker(\mathbb{I} - P)^\perp$.
(3) $\operatorname{Im}(M^\top) = \ker(M^\top)^\perp = \operatorname{Im}(Q) = \ker(\mathbb{I} - Q)^\perp$.

When $P = M^\top M$ is invertible, e.g. when we have full column rank, there is a unique matrix M^+ which satisfies these properties and is given by $M^+ = (M^\top M)^{-1}M^\top$, which in addition is a left inverse. When $Q = MM^\top$ is invertible, e.g. when we have full row rank, then $M^+ = M^\top(MM^\top)^{-1}$; that moreover is a right inverse. In these cases $P = M^\top M$ or $Q = MM^\top$ are denominated *correlation matrices*, respectively.

B.2. Schur complements

Given $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ in block form, the Schur complement with respect to A (if $\det A \neq 0$) is

[17] Following with the boson analogy we have the destruction operators $a_b = N_b^{-1/2}\Lambda_b$ [111].

$$\mathrm{SC}\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \equiv M/A := D - CA^{-1}B.$$

The Schur complement with respect to D (if $\det D \neq 0$) is

$$\mathrm{SC}_D\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \equiv M/D := A - BD^{-1}C.$$

Observe that we have the block Gauss factorization

$$\begin{aligned}\begin{pmatrix} A & B \\ \mathbb{C} & D \end{pmatrix} &= \begin{pmatrix} \mathbb{I} & 0 \\ CA^{-1} & \mathbb{I} \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix}\begin{pmatrix} \mathbb{I} & A^{-1}B \\ 0 & \mathbb{I} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{I} & BD^{-1} \\ 0 & \mathbb{I} \end{pmatrix}\begin{pmatrix} M/D & 0 \\ 0 & D \end{pmatrix}\begin{pmatrix} \mathbb{I} & 0 \\ D^{-1}C & \mathbb{I} \end{pmatrix}\end{aligned}$$

implies the Schur determinant formula $\det M = \det(A)\det(M/A)$. This is in fact the Schur lemma in a disguise form, in fact Schur lemma in [107] assumes that $[A, C] = 0$ so that $\det M = \det(AD - BC)$. In terms of the Schur complements we have the following well known expressions for the inverse matrices

$$\begin{aligned} M^{-1} &= \begin{pmatrix} \mathbb{I} & -A^{-1}B \\ 0 & \mathbb{I} \end{pmatrix}\begin{pmatrix} A^{-1} & 0 \\ 0 & (M/A)^{-1} \end{pmatrix}\begin{pmatrix} \mathbb{I} & 0 \\ -CA^{-1} & \mathbb{I} \end{pmatrix} \\ &= \begin{pmatrix} A^{-1} + A^{-1}B(M/A)^{-1}CA^{-1} & -A^{-1}B(M/A) \\ -(M/A)^{-1}CA^{-1} & (M/A)^{-1} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{I} & 0 \\ -D^{-1}C & \mathbb{I} \end{pmatrix}\begin{pmatrix} (M/D)^{-1} & 0 \\ 0 & D^{-1} \end{pmatrix}\begin{pmatrix} \mathbb{I} & -BD^{-1} \\ 0 & \mathbb{I} \end{pmatrix} \\ &= \begin{pmatrix} (M/D)^{-1} & -(M/D)^{-1}BD^{-1} \\ -D^{-1}B(M/D)^{-1} & D^{-1} + D^{-1}(M/D)^{-1}BD^{-1} \end{pmatrix}. \end{aligned} \tag{B.2.1}$$

The two expressions found for the inverse of M are known as the Matrix Inversion Lemma in Linear Estimation Theory [75] and as Sherman–Morrison–Woodbury formula in Linear Algebra [33]. If both A and D are invertible we deduce that M/A is invertible if and only if M/D is invertible.

B.3. Quasi-determinants and the heredity principle

Given any partitioned matrix

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,k} \\ A_{2,1} & A_{2,2} & \dots & A_{2,k} \\ \vdots & \vdots & & \vdots \\ A_{k,1} & A_{k,2} & \dots & A_{k,k} \end{pmatrix} \tag{B.3.1}$$

where $A_{i,j} \in \mathbb{R}^{m_i\times m_j}$ for $i,j \in \{1,\dots,k-1\}$, and $A_{k,k} \in \mathbb{R}^{\kappa_1\times\kappa_2}$,$A_{i,k} \in \mathbb{R}^{m_i\times\kappa_2}$ and $A_{k,j} \in \mathbb{R}^{\kappa_1\times m_j}$, we are going to define its quasi-determinant *à la Olver* recursively. We start with $k=2$, so that $A = \left(\begin{smallmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2}\end{smallmatrix}\right)$, in this case the first quasi-determinant $\Theta_1(A) := A/A_{1,1}$; i.e., a Schur complement which requires $\det A_{1,1} \neq 0$. The notation of Olver is different to that of the Gel'fand school where $\Theta_1(A) = |A|_{2,2} = \begin{vmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & \boxed{A_{2,2}}\end{vmatrix}$. There is another quasi-determinant $\Theta_2(A) = A/A_{22} = |A|_{1,1} = \begin{vmatrix} \boxed{A_{1,1}} & A_{1,2} \\ A_{2,1} & A_{2,2}\end{vmatrix}$, the other Schur complement, and we need $A_{2,2}$ to be an invertible square matrix. Other quasi-determinants that can be considered for regular square blocks are $\begin{vmatrix} A_{1,1} & A_{1,2} \\ \boxed{A_{2,1}} & A_{2,2}\end{vmatrix}$ and $\begin{vmatrix} A_{1,1} & \boxed{A_{1,2}} \\ A_{2,1} & A_{2,2}\end{vmatrix}$. Notice that these last two quasi-determinants are out of the scope of the paper, as the partitioned matrices considered here have rectangular off diagonal blocks and therefore are not invertible.

Following [96] we remark that quasi-determinantal reduction is a commutative operation. This is the heredity principle formulated by Gel'fand and Retakh [54,58]: quasi-determinants of quasi-determinants are quasi-determinants. Let us illustrate this by reproducing a nice example discussed in [96]. We consider the matrix

$$A = \left(\begin{array}{c|c:c} A_{1,1} & A_{1,2} & A_{1,3} \\ \hline A_{2,1} & A_{2,2} & A_{2,3} \\ \hdashline A_{3,1} & A_{3,2} & A_{3,3} \end{array}\right)$$

and take the quasi-determinant with respect to the first diagonal block, which we define as the Schur complement indicated by the non-dashed lines, to get

$$\Theta_1(A) = \begin{vmatrix} A_{11,1} & A_{1,2} \quad A_{1,3} \\ \begin{matrix} A_{2,1} \\ A_{3,1}\end{matrix} & \boxed{\begin{matrix} A_{2,2} & A_{2,3} \\ A_{3,2} & A_{3,3}\end{matrix}} \end{vmatrix} = \begin{pmatrix} A_{2,2} & A_{2,3} \\ A_{3,2} & A_{3,3}\end{pmatrix} - \begin{pmatrix} A_{2,1} \\ A_{3,1}\end{pmatrix} A_{1,1}^{-1} \begin{pmatrix} A_{1,2} & A_{1,3}\end{pmatrix}$$

$$= \left(\begin{array}{c:c} A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2} & A_{2,3} - A_{2,1}A_{1,1}^{-1}A_{1,3} \\ \hdashline A_{3,2} - A_{3,1}A_{1,1}^{-1}A_{1,2} & A_{3,3} - A_{3,1}A_{1,1}^{-1}A_{1,3} \end{array}\right),$$

a matrix with blocks with subindexes involving 2 and 3 but not 1. Notice also, that as we are allowed to take blocks of different sizes we have taken the quasi-determinant with respect to a bigger block, composed of two rows and columns of basic blocks. This is the Olver's generalization of Gel'fand's et al. construction. Now, we take the quasi-determinant given by the Schur complement as indicated by the dashed lines, to get

$$\Theta_2(\Theta_1(A)) = \begin{vmatrix} A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2} & A_{2,3} - A_{2,1}A_{1,1}^{-1}A_{1,3} \\ A_{3,2} - A_{3,1}A_{1,1}^{-1}A_{1,2} & \boxed{A_{3,3} - A_{3,1}A_{1,1}^{-1}A_{1,3}} \end{vmatrix} \tag{B.3.2}$$

$$\begin{aligned} &= A_{3,3} - A_{3,1}A_{1,1}^{-1}A_{1,3} \\ &\quad - \big(A_{3,2} - A_{3,1}A_{1,1}^{-1}A_{1,2}\big)\big(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}\big)^{-1}\big(A_{2,3} - A_{2,1}A_{1,1}^{-1}A_{1,3}\big). \end{aligned} \tag{B.3.3}$$

We are ready to compute, for the very same matrix

$$A = \left(\begin{array}{cc|c} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ \hline A_{3,1} & A_{3,2} & A_{3,3} \end{array}\right), \tag{B.3.4}$$

the quasi-determinant associated to the two first diagonal blocks, that we label as $\{1,2\}$; i.e., the Schur complement indicated by the non-dashed lines in (B.3.4), to get

$$\begin{aligned} \Theta_{\{1,2\}}(A) &= \begin{vmatrix} A_{1,1} & A_{1,2} & A_{1,3} \\ A_{2,1} & A_{2,2} & A_{2,3} \\ A_{1,3} & A_{2,3} & \boxed{A_{3,3}} \end{vmatrix} \\ &= A_{3,3} - \begin{pmatrix} A_{3,1} & A_{3,2} \end{pmatrix} \begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}^{-1} \begin{pmatrix} A_{1,3} \\ A_{2,3} \end{pmatrix}. \end{aligned}$$

But recalling (B.2.1)

$$\begin{aligned} &\begin{pmatrix} A_{1,1} & A_{1,2} \\ A_{2,1} & A_{2,2} \end{pmatrix}^{-1} \\ &= \begin{pmatrix} A_{1,1}^{-1} + A_{1,1}^{-1}A_{1,2}(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2})^{-1}A_{2,1}A_{1,1}^{-1} & -A_{1,1}^{-1}A_{1,2}(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}) \\ -(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2})^{-1}A_{2,1}A_{1,1}^{-1} & (A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2})^{-1} \end{pmatrix} \end{aligned}$$

we get

$$\begin{aligned} \Theta_{\{1,2\}}(A) = A_{3,3} &- A_{3,1}A_{1,1}^{-1}A_{1,3} + A_{3,1}A_{1,1}^{-1}A_{1,2}\big(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}\big)^{-1}A_{2,1}A_{1,1}^{-1}A_{1,3} \\ &- A_{3,2}\big(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}\big)^{-1}A_{2,1}A_{1,1}^{-1}A_{1,3} \\ &- A_{3,1}A_{1,1}^{-1}A_{1,2}\big(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}\big)^{-1}A_{2,3} \\ &+ A_{3,2}\big(A_{2,2} - A_{2,1}A_{1,1}^{-1}A_{1,2}\big)^{-1}A_{2,3} \end{aligned}$$

which is identical to (B.3.2), so that

$$\Theta_2(\Theta_1(A)) = \Theta_{\{1,2\}}(A).$$

Given any set $I=\{i_1,\dots,i_m\}\subset\{1,\dots,k\}$ the heredity principle allows us to define the quasi-determinant[18]

$$\Theta_I(A)=\Theta_{i_1}(\Theta_{i_2}(\cdots\Theta_{i_m}(A)\cdots))$$

and the ℓ-th quasi-determinant is

$$\Theta^{(\ell)}(A)=\Theta_{\{1,\dots,\ell-1,\ell+1,\dots,k\}}(A)=|A|_{\ell,\ell}=\begin{vmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,\ell} & \dots & A_{1,k}\\ A_{2,1} & A_{2,2} & \dots & A_{2,\ell} & \dots & A_{2,k}\\ \vdots & \vdots & & \vdots & & \vdots\\ A_{\ell,1} & A_{\ell,2} & \dots & \boxed{A_{\ell,\ell}} & \dots & A_{\ell,k}\\ \vdots & \vdots & & \vdots & & \vdots\\ A_{k,1} & A_{k,2} & \dots & A_{k,\ell} & \dots & A_{k,k}\end{vmatrix}.$$

The last quasi-determinant is denoted by

$$\Theta_*(A)=\Theta^{(k)}(A)=|A|_{k,k}=\begin{vmatrix} A_{1,1} & A_{1,2} & \dots & A_{1,k}\\ A_{2,1} & A_{2,2} & \dots & A_{2,k}\\ \vdots & \vdots & & \vdots\\ A_{k,1} & A_{k,2} & \dots & \boxed{A_{k,k}}\end{vmatrix}.$$

B.4. Quasi-determinants and Gauss–Borel factorization

An important application of quasi-determinants presented in [96] is the characterization of the factors of the block Gauss–Borel factorization of a partitioned matrix A as in (B.3.1) (in the case of interest in this paper a Cholesky factorization) in terms of quasi-determinants of A. To present this result we need to introduce for two sets of indices $\{i_1,\dots,i_m\}$ and $\{j_1,\dots,j_m\}$ subset, with m elements, of $\{1,\dots,k\}$

$$A^{i_1\dots i_m}_{j_1\dots j_m}=\begin{pmatrix} A_{i_1,j_1} & \dots & A_{i_1,j_m}\\ \vdots & & \vdots\\ A_{i_m,j_1} & \dots & A_{i_m,j_m}\end{pmatrix}.$$

Theorem B.4.1 *(Theorem 3 in [96]). A regular block matrix as in (B.3.1) factors as*

$$A=LDV$$

with $L=(L_{i,j})$, $D=\operatorname{diag}(D_1,\dots,D_k)$ *and* $V=(V_{i,j})$, *where*

[18] In [96] it is defined as the Schur complement with respect to a big block built up by the blocks determined by the indices I.

$$
\begin{aligned}
L_{i,j} &= \begin{cases} 0, & i < j, \\ \Theta_*(A^{12\dots j-1,i}_{12\dots j-1,j})\Theta_*(A^{12\dots j}_{12\dots j})^{-1}, & i \geq j, \end{cases} \\
D_j &= \Theta_*(A^{12\dots j}_{12\dots j}), \\
V_{i,j} &= \begin{cases} 0, & i > j, \\ \Theta_*(A^{12\dots i}_{12\dots i})^{-1}\Theta_*(A^{12\dots i-1,i}_{12\dots i-1,j}), & i \leq j. \end{cases}
\end{aligned}
$$

Regularity of A requires invertibility of $\Theta_*(A^{12\dots j}_{12\dots j})$ for $j = 1, \dots, k$.

For a symmetric case, $A = A^\top$, we have

$$
\Theta_*(A^{12\dots i}_{12\dots i}) = \Theta_*(A^{12\dots i}_{12\dots i})^\top, \qquad \Theta_*(A^{12\dots i-1,i}_{12\dots i-1,j}) = \Theta_*(A^{12\dots i-1,j}_{12\dots i-1,i})^\top.
$$

Appendix C. Several complex variables

In this paper we discuss multivariate second kind functions in the realm of the block Cholesky factorization and for that aim some facts regarding complex analysis in several variables are needed. Here we just recall them, see for example [21,69,106,79] for more information.

(1) Given the vector $\boldsymbol{r} = (r_1, \dots, r_D)^\top \in \mathbb{R}^D_+$, we consider the polydisk

$$
\Delta(\boldsymbol{r}) = \{\boldsymbol{z} = (z_1, \dots, z_D)^\top : |z_i| < r_i, i = 1, \dots, D\} \subset \mathbb{C}^D
$$

centered at the origin of polyradius $\boldsymbol{r}$. Its distinguished boundary is the D-dimensional torus

$$
\mathbb{T}^D(\boldsymbol{r}) = \{\boldsymbol{z} \in \mathbb{C}^D : |z_i| = r_i, i = 1, \dots, D\}.
$$

Recall that the border of the polydisk $\Gamma = \partial\Delta$ splits in D sets of dimension $2D-1$, the distinguished border being its skeleton; i.e. the intersection of all them. The distinguished border is also known as Shilov border.

(2) Given two polyradii $\boldsymbol{r}$ and $\boldsymbol{R}$ we construct the associated polyannulus centered at the origin

$$
A^D(\boldsymbol{r}, \boldsymbol{R}) := \{\boldsymbol{z} \in \mathbb{C}^D : r_i < z_i < R_i, i = 1, \dots, D\}.
$$

(3) A set $A \subset \mathbb{C}^n$ is a complete Reinhardt domain if the unit polydisk acts on it by componentwise multiplication.

(4) Any set $A \subset \mathbb{C}^D$ is called Reinhardt (D-circled) if for each $\boldsymbol{\lambda} := (\mathrm{e}^{\mathrm{i}\,\theta_1}, \dots, \mathrm{e}^{\mathrm{i}\,\theta_D}) \in \mathbb{T}^D$ with $\theta_i \in [0, 2\pi)$ for every $\boldsymbol{c} \in A$ we have that $(\mathrm{e}^{\mathrm{i}\,\theta_1} c_1, \dots, \mathrm{e}^{\mathrm{i}\,\theta_D} c_D)^\top \in A$; i.e., $\mathbb{T}^D$ acts on A componentwise.

(5) If $\mathscr{D} \subset \mathbb{C}^D$ is the domain of convergence of a Laurent series $L(\boldsymbol{z})$, then for any $\boldsymbol{c} = (c_1, \dots, c_D)^\top \in \mathscr{D}$ we have that $\mathbb{T}^D(|c_1|, \dots, |c_D|) \subset \mathscr{D}$. Thus, the domain of convergence is a Reinhardt (D-circled) domain.
(6) The domain of convergence $\mathscr{D}_{\boldsymbol{\ell}_a}$ is logarithmically convex; i.e., the set

$$\log \mathscr{D}_{\boldsymbol{\ell}_a} := \{(\log |z_1|, \dots, |z_D|) : (z_1, \dots, z_D^\top \in \mathscr{D}_{\boldsymbol{\ell}_a}\}$$

is convex (given any pair of points $\boldsymbol{c}_1, \boldsymbol{c}_2 \in \mathscr{D}_{\boldsymbol{\ell}_a}$, the segment $[\boldsymbol{c}_1, \boldsymbol{c}_2] := \{(1-t)\boldsymbol{c}_1 + t\boldsymbol{c}_2 : t \in [0,1]\} \subset \mathscr{D}_{\boldsymbol{\ell}_a})$.
(7) For all polyradii $\boldsymbol{r}$ and $\boldsymbol{R}$ the annulus $A^D(\boldsymbol{r}, \boldsymbol{R})$ is a Reinhardt domain. Any Reinhardt domain is the union of polyannuli and so is the domain of convergence $\mathscr{D}$.
(8) The polydisk of convergence of a power series is such that any other polydisk $\Delta(\boldsymbol{r}')$ with $r_j < r'_j$ for some j contains points where the power series diverge.
(9) The Laurent series is locally normally summable in its domain of convergence and therefore locally absolutely uniformly summable. We remind the reader that a Laurent series $\sum_{\boldsymbol{k} \in \mathbb{Z}^D} a_{\boldsymbol{k}} \boldsymbol{z}^{\boldsymbol{k}}$ is locally normally summable if for any compact set $K \subset \mathscr{D}$ there exist $C > 0$ and $\theta \in (0,1)$ such that $|a_{\boldsymbol{k}} \boldsymbol{z}^{\boldsymbol{k}}| \leq C\theta^{|\boldsymbol{k}|}$ for $\boldsymbol{z} \in K$ and $\boldsymbol{k} \in \mathbb{Z}^D$.
(10) The function $L(\boldsymbol{z})$ is holomorphic (holomorphic in each variable z_i, $i = 1, \dots, D$) in $\mathscr{D}$, which is its domain of holomorphy.
(11) Given a holomorphic function $L(\boldsymbol{z})$ in $A^D_{\boldsymbol{c}}(\boldsymbol{r}, \boldsymbol{R})$ (a polyannulus centered at $\boldsymbol{c} \in \mathbb{C}^D$), and a polyradius $\boldsymbol{\rho}$ such that $r_i < \rho_i < R_i$, $i = 1, \dots, D$ then

$$L(\boldsymbol{z}) = \sum_{\boldsymbol{k} \in \mathbb{Z}^D} c_{\boldsymbol{k}} (\boldsymbol{z} - \boldsymbol{c})^{\boldsymbol{k}}, \qquad c_{\boldsymbol{k}} = \frac{1}{(2\pi\,\mathrm{i})^D} \int\limits_{\mathbb{T}^D_{\boldsymbol{c}}(\boldsymbol{\rho})} \frac{L(\boldsymbol{w})}{(\boldsymbol{w} - \boldsymbol{c})^{\boldsymbol{k}}} \mathrm{d}w_1 \dots \mathrm{d}w_D,$$

where $\mathbb{T}^D_{\boldsymbol{c}}(\boldsymbol{\rho})$ is the distinguished border of the polycircle centered at $\boldsymbol{c}$ with polyradius $\boldsymbol{\rho}$.

Appendix D. Proofs

D.1. Proof of Proposition 2.3.1

Proof. Assuming $\det A \neq 0$ for any block matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$ we can write in terms of Schur complements

$$M = \begin{pmatrix} \mathbb{I} & 0 \\ CA^{-1} & \mathbb{I} \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix} \begin{pmatrix} \mathbb{I} & A^{-1}B \\ 0 & \mathbb{I} \end{pmatrix}.$$

Thus, as $\det G^{[k]} \neq 0$ $\forall k = 0, 1, \dots,$ we can write

$$G^{[\ell+1]} = \left(\begin{array}{c|c} \mathbb{I}^{[\ell]} & 0 \\ \hline v^{[\ell],[\ell-1]} & \mathbb{I}_{[\ell]} \end{array}\right) \left(\begin{array}{c|c} G^{[\ell]} & 0 \\ \hline 0 & G^{[\ell+1]}/G^{[\ell]} \end{array}\right) \left(\begin{array}{c|c} \mathbb{I}^{[\ell]} & (v^{[\ell],[\ell-1]})^\top \\ \hline 0 & \mathbb{I}_{[\ell]} \end{array}\right),$$

where

$$v^{[\ell],[\ell-1]} := \begin{pmatrix} v_{[\ell],[0]} & v_{[\ell],[1]} & \dots & v_{[\ell],[\ell-1]} \end{pmatrix}.$$

Applying the same factorization to $G^{[\ell]}$ we get

$$G^{[\ell+1]} = \left(\begin{array}{c|cc} \mathbb{I}^{[\ell-1]} & 0 & 0 \\ \hline r^{[\ell-1][\ell-2]} & \mathbb{I}_{[\ell-1]} & 0 \\ s^{[\ell][\ell-2]} & t_{[\ell][\ell-1]} & \mathbb{I}_{[\ell]} \end{array}\right) \left(\begin{array}{c|cc} G^{[\ell-1]} & 0 & 0 \\ \hline 0 & G^{[\ell]}/G^{[\ell-1]} & 0 \\ 0 & 0 & G^{[\ell+1]}/G^{[\ell]} \end{array}\right)$$
$$\times \left(\begin{array}{c|cc} \mathbb{I}^{[\ell-1]} & (r^{[\ell-1][\ell-2]})^\top & (s^{[\ell][\ell-2]})^\top \\ \hline 0 & \mathbb{I}_{[\ell-1]} & (t_{[\ell],[\ell-1]})^\top \\ 0 & 0 & \mathbb{I}_{[\ell]} \end{array}\right).$$

Here the zeroes indicates zero rectangular matrices of different sizes. Finally, the iteration of these factorizations leads to

$$G^{[\ell+1]} = \begin{pmatrix} \mathbb{I}_{[|0]} & 0 & \dots & 0 \\ * & \mathbb{I}_{[1]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \dots & * & \mathbb{I}_{[\ell]} \end{pmatrix} \operatorname{diag}(G^{[1]}/G^{[0]}, G^{[2]}/G^{[1]}, \dots, G^{[\ell+1]}/G^{[\ell]})$$
$$\times \begin{pmatrix} \mathbb{I}_{[|0]} & 0 & \dots & 0 \\ * & \mathbb{I}_{[1]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \dots & * & \mathbb{I}_{[\ell]} \end{pmatrix}^\top.$$

Since this would have been valid for any ℓ it would also hold for the direct limit $\varinjlim G^{[\ell]}$. $\square$

D.2. Proof of Proposition 2.5.1

Proof.

$$\begin{aligned} C_{[\ell]}(\boldsymbol{z}) &= \sum_{n=0}^{\infty} (SG)_{[\ell],[n]} \chi^*_{[n]}(\boldsymbol{z}) && \text{use the Cholesky factorization (2.3.2)} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^{\ell} S_{[\ell],[k]} G_{[k],[n]} \chi^*_{[n]}(\boldsymbol{z}) \\ &= \sum_{n=0}^{\infty} \int_\Omega \sum_{k=0}^{\ell} S_{[\ell],[k]} \chi_{[k]}(\boldsymbol{y}) \mathrm{d}\mu(\boldsymbol{y}) \big(\chi_{[n]}(\boldsymbol{y})\big)^\top \chi^*_{[n]}(\boldsymbol{z}) && \text{recall (2.3.1)} \\ &= \sum_{n=0}^{\infty} \int_\Omega P_{[\ell]}(\boldsymbol{y}) \mathrm{d}\mu(\boldsymbol{y}) \big(\chi_{[n]}(\boldsymbol{y})\big)^\top \chi^*_{[n]}(\boldsymbol{z}) && \text{because (2.4.1)} \end{aligned}$$

$$
\begin{aligned}
&= \int_\Omega P_{[\ell]}(\boldsymbol{y})\mathrm{d}\mu(\boldsymbol{y}) \sum_{n=0}^{\infty}\big(\chi_{[n]}(\boldsymbol{y})\big)^\top \chi^*_{[n]}(\boldsymbol{z}) && \text{interchange of series and integral}\\
&= \int_\Omega P_{[\ell]}(\boldsymbol{y})\mathrm{d}\mu(\boldsymbol{y}) \frac{1}{(z_1-y_1)\cdots(z_D-y_D)} && \text{recall (2.1.1).} \qquad \square
\end{aligned}
$$

D.3. Proof of Proposition 2.6.2

Proof. From (2.6.3) we deduce that

$$
\Big[\prod_{i=1}^{k}(\Lambda_{a_i}^\top - q_{a_i})\Big]\chi^* = \Big[\prod_{i=1}^{k}(x_{a_i}\Pi_{a_i} - q_{a_i})\Big]\chi^*
$$

but

$$
\begin{aligned}
\prod_{i=1}^{k}(x_{a_i}\Pi_{a_i} - q_{a_i}) &= \prod_{i=1}^{k}(x_{a_i} - q_{a_i} - x_{a_i}\Pi_{a_i}^\perp)\\
&= \Big[\prod_{i=1}^{k}(x_{a_i} - q_{a_i})\Big] + (-1)^k\Big[\prod_{i=1}^{k} x_{a_i}\Pi_{a_i}^\perp\Big]\\
&\quad + \sum_{j=1}^{k-1}\frac{(-1)^j}{(k-j)!j!}\sum_{\sigma\in\mathfrak{S}_k}\Big(\Big[\prod_{i=j+1}^{k}(x_{a_{\sigma i}} - q_{a_{\sigma i}})\Big]\Big[\prod_{i=1}^{j} x_{a_{\sigma i}}\Pi_{a_{\sigma i}}^\perp\Big]\Big)
\end{aligned}
$$

and (2.6.1) implies the result. $\square$

D.4. Proof of Proposition 2.7.2

Proof. Just follow the chain of equalities

$$
\begin{aligned}
\Big[\prod_{i=1}^{k}(J_{a_i} - q_{a_i})\Big]C &= S\prod_{i=1}^{k}(\Lambda_{a_i} - q_{a_i})S^{-1}H(S^{-1})^\top\chi^* && \text{use (2.5.1) and (2.7.1)}\\
&= S\prod_{i=1}^{k}(\Lambda_{a_i} - q_{a_i})G\chi^* && \text{use (2.3.2)}\\
&= SG\prod_{i=1}^{k}(\Lambda_{a_i}^\top - q_{a_i})\chi^* && \text{from (2.6.5)}\\
&= H(S^{-1})^\top\prod_{i=1}^{k}(\Lambda_{a_i}^\top - q_{a_i})\chi^* && \text{follows from (2.3.2).}
\end{aligned}
$$

Finally, (2.6.4) implies the announced result. $\square$

D.5. Proof of Proposition 3.1.8

Proof.

(1) It is proven as follows

$$\begin{aligned}T_aW_1 &= (T_aS)(T_aW_0)\\ &= (T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)W_0\\ &= (T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)S^{-1}W_1\\ &= \omega_aW_1,\\ T_aW_2 &= (T_aH)\big(T_aS^{-1}\big)^\top\\ &= (T_aH)\big(T_aS^{-1}\big)^\top\big(S^\top H^{-1}\big)\big(S^\top H^{-1}\big)^{-1}\\ &= \omega_aW_2.\end{aligned}$$

(2) For the first equation observe that

$$\begin{aligned}(T_a(\boldsymbol{n}_b\cdot\boldsymbol{J}))\omega_a &= (T_aS)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})(T_aS)^{-1}(T_aH)\big((T_aS)^{-1}\big)^\top S^\top H^{-1}\\ &= (T_aS)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})(T_aG)S^\top H^{-1}\\ &= (T_aS)(T_aG)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})^\top S^\top H^{-1}\\ &= (T_aH)\big((T_aS)^{-1}\big)^\top(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})^\top S^\top H^{-1}\\ &= (T_aH)\big((T_aS)^{-1}\big)^\top S^\top H^{-1}H\big(S^{-1}\big)^\top(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})^\top S^\top H^{-1}\\ &= \omega_aH(\boldsymbol{n}_b\cdot\boldsymbol{J})^\top H^{-1}\\ &= \omega_a(\boldsymbol{n}_b\cdot\boldsymbol{J}),\end{aligned}$$

and for the second one

$$M_b(T_bM_a) = S(T_bS)^{-1}(T_bS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})(T_bS)^{-1} = (\boldsymbol{n}_a\cdot\boldsymbol{J})M_b.$$

(3) For the first equation from (3.1.19) we get $T_b(T_aW) = (T_b\omega_a)(T_bW) = \big[(T_b\omega_a)\omega_b\big]W$ and interchanging $a\leftrightarrow b$ we get $\big[(T_a\omega_b)\omega_a - (T_b\omega_a)\omega_b\big]SW_0 = 0$. For the second equation, from the definitions, it is easy to see that

$$M_a(T_aM_b) = S(T_aT_bS)^{-1} = M_b(T_bM_a). \qquad\square$$

D.6. Proof of Proposition 3.2.3

Proof. From (2.6.3) we get

$$
\begin{aligned}
(\boldsymbol{n}_a\cdot\Lambda-q_a)^\top\chi^* &= \Big(\sum_{b=1}^{D} n_{a,b}x_b\Pi_b-q_a\Big)\chi^*\\
&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)\chi^*-\Big(\sum_{b=1}^{D} n_{a,b}x_b\Pi_b^{\perp}\Big)\chi^*\\
&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)\chi^*-\Big(\sum_{b=1}^{D} n_{a,b}\lim_{x_b\to\infty}x_b\chi^*\Big)\\
&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)\chi^*-\boldsymbol{n}_a\cdot\widehat{\boldsymbol{\chi}^*}
\end{aligned}
$$

where

$$
\widehat{\boldsymbol{\chi}^*}=(\lim_{x_1\to\infty}x_1\chi^*,\dots,\lim_{x_D\to\infty}x_D\chi^*).
$$

Consequently

$$
\begin{aligned}
M_a(T_aC)&=H\big((T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)S^{-1}\big)^\top\\
&\quad\times(T_aH)^{-1}(T_aH)\Big(\big(T_aS\big)^{-1}\Big)^\top\chi^* \qquad \text{from (2.5.1) and (3.1.4)}\\
&=H\big(S^{-1}\big)^\top(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)^\top\chi^*\\
&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)C-\boldsymbol{n}_a\cdot\widehat{\boldsymbol{C}},
\end{aligned}
$$

that together with

$$
\omega_aC=T_aC
$$

implies the result. □

D.7. Proof of Theorem 3.2.2

Proof. Previously to the proof we need

Lemma D.7.1. *The following relation is satisfied by the second kind functions*

$$
\begin{aligned}
M(TC)&=\Big[\prod_{a=1}^{D}(x_a-q_a)\Big]C+(-1)^D\widehat{C}_{1,\dots,D}\\
&\quad+\sum_{j=1}^{D}\frac{(-1)^j}{(D-j)!j!}\sum_{\sigma\in\mathfrak{S}_D}\Big(\Big[\prod_{a=j+1}^{D}\big(x_{\sigma a}-q_{\sigma a}\big)\Big]\widehat{C}_{\sigma 1,\dots,\sigma j}\Big). \qquad \text{(D.7.1)}
\end{aligned}
$$

Proof. The proof is a consequence of

$$
\begin{aligned}
M(TC) &= (H)(S^{-1})^\top \Big[\prod_{a=1}^{D}(\Lambda_a-q_a)\Big]^\top (TS)^\top (TH)^{-1}(TH)((TS)^{-1})^\top \chi^* \quad \text{from (3.2.6)}\\
&= H(S^{-1})^\top \Big[\prod_{a=1}^{D}(\Lambda_a-q_a)\Big]^\top \chi^*\\
&= (HS^{-1})^\top \Big(\Big[\prod_{a=1}^{D}(x_a-q_a)\Big]\chi^* \quad \text{from (2.6.4)}\\
&\quad + \sum_{j=1}^{D}\frac{(-1)^j}{(k-j)!j!}\\
&\quad \times \sum_{\sigma\in\mathfrak{S}_D}\Big(\Big[\prod_{a=j+1}^{D}\big(x_{\sigma a}-q_{\sigma a}\big)\Big]\lim_{x_{\sigma 1}\to\infty}\cdots\lim_{x_{\sigma j}\to\infty}\Big(\Big[\prod_{i=1}^{j}x_{\sigma i}\Big]\chi^*\Big)\Big)\Big). \qquad \square
\end{aligned}
$$

Now, in Lemma D.7.1 we put $\boldsymbol{x}=\boldsymbol{q}$ into (D.7.1), observe that $\widehat{C}_{[k],1,\dots,D}=\delta_{k,0}H_{[0]}$, and multiply by the inverse of the lower unitriangular matrix M to get

$$
(C)_{[k]}(\boldsymbol{q}) = (-1)^D\big(T^{-1}M^{-1}\big)_{[k],[0]}T^{-1}H_{[0]}. \tag{D.7.2}
$$

According to (3.2.7) with $\sigma=1$ we have

$$
T^{-1}M^{-1}=\big(T_D^{-1}M_D\big)^{-1}\cdots\big(T_2^{-1}\cdots T_D^{-1}M_2\big)^{-1}(T^{-1}M_1)^{-1}
$$

and given the particular structure of M_a, $a\in\{1,\dots,D\}$, we have the following simple expression

$$
\big(T_a^{-1}\cdots T_D^{-1}M_a\big)^{-1} = \begin{pmatrix}
\mathbb{I}_{[0]} & 0 & 0 & 0 & \cdots\\
-\rho^{(a)}_{[1]} & \mathbb{I}_{[1]} & 0 & 0 & \cdots\\
\rho^{(a)}_{[2]}\rho^{(a)}_{[1]} & -\rho^{(a)}_{[2]} & \mathbb{I}_{[2]} & 0 & \cdots\\
-\rho^{(a)}_{[3]}\rho^{(a)}_{[2]}\rho^{(a)}_{[1]} & \rho^{(a)}_{[3]}\rho^{(a)}_{[2]} & -\rho^{(a)}_{[3]} & \mathbb{I}_{[3]} & \ddots\\
\vdots & \ddots & \ddots & & \ddots
\end{pmatrix}
$$

for $a\in\{1,\dots,D\}$. This allows explicit computation of the elements of the inverse matrix M^{-1} and in particular leads to products over multisets, see Appendix A,

$$
\big(T^{-1}M^{-1}\big)_{[k],[0]} = (-1)^k\sum_{1\leq a_1\leq\cdots\leq a_k\leq D}\rho^{(a_k)}_{[k]}\cdots\rho^{(a_1)}_{[1]}
$$

so that (D.7.2) reads

$$
C_{[k]}(\boldsymbol{q}) = (-1)^{k+D}\sum_{1\leq a_1\leq\cdots\leq a_k\leq D}\rho^{(a_k)}_{[k]}\cdots\rho^{(a_1)}_{[1]}T^{-1}H_{[0]}
$$

and recalling (3.2.9) we get the desired result. $\square$

D.8. Proof of Theorem 3.3.1

Proof. To obtain the result we consider the expressions of $P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} M_a^{[\ell]} \times (T_a P^{[\ell]})(\boldsymbol{y})$ when letting the operator between square brackets act to the right or to the left. Acting on its right gives the Christoffel–Darboux kernel

$$\begin{aligned} &P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} M_a^{[\ell]} (T_a P^{[\ell]})(\boldsymbol{y}) \\ &\quad = P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} S^{[\ell]} (T_a S^{[\ell]})^{-1} (T_a P^{[\ell]})(\boldsymbol{y}), && \text{consequence of (3.1.3)} \\ &\quad = P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} P^{[\ell]}(\boldsymbol{y}) && \text{see (2.4.1)} \\ &\quad = K^{(\ell)}(\boldsymbol{x}, \boldsymbol{y}), && \text{see Definition 2.8.1.} \end{aligned}$$

If we act on the left, recalling (3.1.4) we get

$$\begin{aligned} &P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} M_a^{[\ell]} (T_a P^{[\ell]})(\boldsymbol{y}) \\ &\quad = P^{[\ell]}(\boldsymbol{x})^\top (H^{[\ell]})^{-1} H^{[\ell]} \Big(\big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} \big)^{[\ell]} \Big)^\top (T_a H^{[\ell]})^{-1} (T_a P^{[\ell]})(\boldsymbol{y}) \\ &\quad = \Big(\big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} \big)^{[\ell]} P^{[\ell]}(\boldsymbol{x}) \Big)^\top (T_a H^{[\ell]})^{-1} (T_a P^{[\ell]})(\boldsymbol{y}). \end{aligned}$$

Now, with the help of the block decomposition of any block semi-infinite matrix $M = \begin{pmatrix} M^{[\ell]} & M^{[\ell],[\geq \ell]} \\ M^{[\geq \ell],[\ell]} & M^{[\geq \ell]} \end{pmatrix}$ we write

$$\begin{aligned} &\big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} \big)^{[\ell]} P^{[\ell]}(\boldsymbol{x}) \\ &\quad = \big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} P(\boldsymbol{x}) \big)^{[\ell]} - \big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} \big)^{[\ell],[\geq \ell]} P^{[\geq \ell]}(\boldsymbol{x}). \end{aligned}$$

On the one hand, if we take into account (2.4.1) and (2.6.2) the first term in the LHS reads

$$\big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} P(\boldsymbol{x}) \big)^{[\ell]} = (\boldsymbol{n}_a \cdot \boldsymbol{x} - q_a) T_a P^{[k]}(\boldsymbol{x})$$

and on the other hand, given the lower unitriangular form of $T_a S$ and S and that $\boldsymbol{n}_a \cdot \boldsymbol{\Lambda}$ is zero but for the first superdiagonal

$$\big((T_a S)(\boldsymbol{n}_a \cdot \boldsymbol{\Lambda} - q_a) S^{-1} \big)^{[\ell],[\geq \ell]} = \begin{pmatrix} 0_{[0],[\ell]} & 0_{[0],[\ell+1]} & \cdots \\ 0_{[1],[\ell]} & 0_{[1],[\ell+1]} & \cdots \\ \vdots & \vdots & \\ 0_{[\ell-2],[\ell]} & 0_{[\ell-2],[\ell+1]} & \cdots \\ (\boldsymbol{n} \cdot \boldsymbol{\Lambda})_{[\ell-1],[\ell]} & 0_{[\ell-1],[\ell+1]} & \cdots \end{pmatrix}$$

and therefore

$$\big((T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)S^{-1}\big)^{[\ell],[\geq\ell]}P^{[\geq\ell]}(\boldsymbol{x})=\begin{pmatrix}0_{[0]}\\0_{[1]}\\\vdots\\0_{[\ell-2]}\\(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}P_{[\ell]}(\boldsymbol{x})\end{pmatrix}.$$

Hence,

$$\begin{aligned}\Big(\big((T_aS)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}-q_a)S^{-1}\big)^{[\ell]}P^{[\ell]}(\boldsymbol{x})\Big)^\top&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aP^{[\ell]}(\boldsymbol{x}))^\top\\&\quad-(0_{[0]},0_{[1]},\dots,0_{[\ell-2]},P_{[\ell]}(\boldsymbol{x})^\top((\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]})^\top)\end{aligned}$$

so that

$$\begin{aligned}P^{[\ell]}(\boldsymbol{x})^\top(H^{[\ell]})^{-1}M_a^{[\ell]}(T_aP^{[\ell]})(\boldsymbol{y})&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)T_a\big(P^{[\ell]}(\boldsymbol{x})^\top(H^{[\ell]})^{-1}P^{[\ell]}(\boldsymbol{y})\big)\\&\quad-P_{[\ell]}(\boldsymbol{x})^\top\big((\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}\big)^\top(T_aH_{[\ell-1]})^{-1}(T_aP_{[\ell-1]})(\boldsymbol{y}).\end{aligned}$$

Consequently, equating both results we conclude

$$\begin{aligned}K^{(\ell)}(\boldsymbol{x},\boldsymbol{y})&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)T_aK^{(\ell)}(\boldsymbol{x},\boldsymbol{y})\\&\quad-\big((T_aH_{[\ell-1]})^{-1}(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[\ell-1],[\ell]}P_{[\ell]}(\boldsymbol{x})\big)^\top(T_aP_{[\ell-1]})(\boldsymbol{y}).\end{aligned}\tag{D.8.1}$$

Now we recall (3.2.2) in the following form

$$(T_aH_{[\ell-1]})^{-1}(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})_{[\ell-1],[k]}P_{[\ell]}=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aH_{[\ell-1]})^{-1}(T_aP)_{[\ell-1]}-H_{[\ell-1]}^{-1}P_{[\ell-1]},$$

and introduce it into (D.8.1) to get

$$\begin{aligned}K^{(\ell)}(\boldsymbol{x},\boldsymbol{y})&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)T_aK^{(\ell)}(\boldsymbol{x},\boldsymbol{y})\\&\quad-\big((\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)(T_aH_{[\ell-1]})^{-1}(T_aP_{[\ell-1]}(\boldsymbol{x}))-H_{[\ell-1]}^{-1}P_{[\ell-1]}(\boldsymbol{x}))\big)^\top\\&\quad\times(T_aP_{[\ell-1]})(\boldsymbol{y})\\&=(\boldsymbol{n}_a\cdot\boldsymbol{x}-q_a)T_aK^{(\ell-1)}(\boldsymbol{x},\boldsymbol{y})+P_{[\ell-1]}(\boldsymbol{x})^\top)H_{[\ell-1]}^{-1}(T_aP_{[\ell-1]})(\boldsymbol{y}).\qquad\square\end{aligned}$$

D.9. Proof of Proposition 3.5.6

Proof. Observe that Proposition 3.5.5 implies

$$\omega_{[k],[k]}=-\omega_{[k],[k+2]}\begin{pmatrix}\Sigma^{(1),k}_{[k+2]}&\Sigma^{(2),k+1}_{[k+2]}\end{pmatrix}\begin{pmatrix}\Sigma^{(1),k}_{[k]}&\Sigma^{(2),k+1}_{[k]}\\\Sigma^{(1),k}_{[k+1]}&\Sigma^{(2),k+1}_{[k+1]}\end{pmatrix}^{-1}\begin{pmatrix}\mathbb{I}_{[k]}\\0_{[k+1],[k]}\end{pmatrix},$$

$$\omega_{[k],[k+1]}=-\omega_{[k],[k+2]}\begin{pmatrix}\Sigma^{(1),k}_{[k+2]}&\Sigma^{(2),k+1}_{[k+2]}\end{pmatrix}\begin{pmatrix}\Sigma^{(1),k}_{[k]}&\Sigma^{(2),k+1}_{[k]}\\\Sigma^{(1),k}_{[k+1]}&\Sigma^{(2),k+1}_{[k+1]}\end{pmatrix}^{-1}\begin{pmatrix}0_{[k],[k+1]}\\\mathbb{I}_{[k+1]}\end{pmatrix}.$$

Now, from Proposition 3.5.3 we get

$$(T^{(1)}T^{(2)}H_{[k]})H_{[k]}^{-1} = -\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]}\begin{pmatrix}\Sigma^{(1),k}_{[k+2]} & \Sigma^{(2),k+1}_{[k+2]}\end{pmatrix}$$
$$\times\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \Sigma^{(2),k+1}_{[k]}\\ \Sigma^{(1),k}_{[k+1]} & \Sigma^{(2),k+1}_{[k+1]}\end{pmatrix}^{-1}\begin{pmatrix}\mathbb{I}_{[k]}\\ 0_{[k+1],[k]}\end{pmatrix},$$

and

$$\begin{aligned}&(T^{(1)}T^{(2)}\beta_{[k]})\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k-1],[k+1]}\\ &\quad=(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k],[k+1]}+\big((\boldsymbol{n}^{(1)}\cdot\boldsymbol{\Lambda})(\boldsymbol{n}^{(2)}\cdot\boldsymbol{\Lambda})\big)_{[k],[k+2]}\Big(\beta_{[k+2]}\\ &\qquad-\begin{pmatrix}\Sigma^{(1),k}_{[k+2]} & \Sigma^{(2),k+1}_{[k+2]}\end{pmatrix}\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \Sigma^{(2),k+1}_{[k]}\\ \Sigma^{(1),k}_{[k+1]} & \Sigma^{(2),k+1}_{[k+1]}\end{pmatrix}^{-1}\begin{pmatrix}0_{[k],[k+1]}\\ \mathbb{I}_{[k+1]}\end{pmatrix}\Big).\end{aligned}$$

From here the stated result follows easily by recalling the expressions of the quasi-determinant. □

D.10. Proof of Theorem 3.5.2

Proof. From (3.5.9) we have

$$\begin{aligned}&\prod_{i=1}^{m}(\boldsymbol{n}^{(i)}\cdot\boldsymbol{x}-q^{(i)})\Big(\prod_{i=1}^{m}T_{\boldsymbol{n}^{(i)}}P\Big)_{[k]}(\boldsymbol{x})\\ &\quad=\omega_{[k],[k+m]}P_{[k+m]}(\boldsymbol{x})+\omega_{[k],[k+m-1]}P_{[k+m-1]}(\boldsymbol{x})+\cdots+\omega_{[k],[k]}P(\boldsymbol{x})\\ &\quad=\omega_{[k],[k+m]}\Big(P_{[k+m]}(\boldsymbol{x})-\begin{pmatrix}\Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]}\end{pmatrix}\\ &\qquad\times\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]}\\ \vdots & & \vdots\\ \Sigma^{(1),k}_{[k+m-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-1]}\end{pmatrix}^{-1}\begin{pmatrix}P_{[k]}(\boldsymbol{x})\\ \vdots\\ P_{[k+m-1]}(\boldsymbol{x})\end{pmatrix}\Big)\end{aligned}$$

from where the result follows. For the second kind functions we proceed similarly

$$\begin{aligned}&\Big(\prod_{i=1}^{m}T_{\boldsymbol{n}^{(i)}}C\Big)_{[k]}(\boldsymbol{x})\\ &\quad=\omega_{[k],[k+m]}C_{[k+m]}(\boldsymbol{x})+\omega_{[k],[k+m-1]}C_{[k+m-1]}(\boldsymbol{x})+\cdots+\omega_{[k],[k]}C(\boldsymbol{x})\\ &\quad=\omega_{[k],[k+m]}\Big(C_{[k+m]}(\boldsymbol{x})-\begin{pmatrix}\Sigma^{(1),k}_{[k+m]} & \dots & \Sigma^{(m),k+m-1}_{[k+m]}\end{pmatrix}\end{aligned}$$

$$
\times\begin{pmatrix}\Sigma^{(1),k}_{[k]} & \dots & \Sigma^{(m),k+m-1}_{[k]}\\ \vdots & & \vdots\\ \Sigma^{(1),k}_{[k+m-1]} & \dots & \Sigma^{(m),k+m-1}_{[k+m-1]}\end{pmatrix}^{-1}\begin{pmatrix}C_{[k]}(\boldsymbol{x})\\ \vdots\\ C_{[k+m-1]}(\boldsymbol{x})\end{pmatrix}\Bigg). \qquad \square
$$

D.11. Proof of Theorem 3.5.3

Proof. Let us consider a similar matrix to that discussed in Definition 3.2.2, $M = S(TS)^{-1}$ which factors out as $M = M^{(1)}(T^{(1)}M^{(2)})\cdots(T^{(1)}\cdots T^{(m-1}M^{(m)})$. From the symmetry of the moment matrix, $G\prod_{i=1}^{m}\big(n^{(i)}\cdot\boldsymbol{\Lambda}-q^{(i)}\big)^{\top} = TG$, we conclude that $M = H\omega^{\top}(TH)^{-1}$. Notice that $MTP = P$ and $\omega P = \mathcal{Q}TP$. Now, we proceed as in the proof of Theorem 3.3.1 and evaluate $P^{[\ell+m]}(\boldsymbol{x})^{\top}\big(H^{[\ell+m]}\big)^{-1}M^{[\ell+m]}TP^{[\ell+m]}(\boldsymbol{y})$; a sandwich constructed in terms of $(\ell+m)$-th block truncations of semi-infinite matrices. We do it in two ways; first by acting on the right and, as S is a block lower unitriangular matrix and therefore $M^{[\ell+m]} = S^{[\ell+m]}\big(TS^{[\ell+m]}\big)^{-1}$, we get

$$
\begin{aligned}
P^{[\ell+m]}(\boldsymbol{x})^{\top}\big(H^{[\ell+m]}\big)^{-1}M^{[\ell+m]}TP^{[\ell+m]}(\boldsymbol{y}) &= P^{[\ell+m]}(\boldsymbol{x})^{\top}\big(H^{[\ell+m]}\big)^{-1}P^{[\ell+m]}(\boldsymbol{y})\\
&= K^{(\ell+m)}(\boldsymbol{x},\boldsymbol{y}).
\end{aligned}
$$

To act on the left we first evaluate the truncation $M^{[\ell+m]} = H^{[\ell+m]}\big(\omega^{[\ell+m]}\big)^{\top}\times\big(TH^{[\ell+m]}\big)^{-1}$ in terms of the corresponding truncated resolvent. Then,

$$
\begin{aligned}
&P^{[\ell+m]}(\boldsymbol{x})^{\top}\big(H^{[\ell+m]}\big)^{-1}M^{[\ell+m]}TP^{[\ell+m]}(\boldsymbol{y})\\
&\quad= \big(\omega^{[\ell+m]}P^{[\ell+m]}(\boldsymbol{x})\big)^{\top}\big(TH^{[\ell+m]}\big)^{-1}P^{[\ell+m]}(\boldsymbol{y}).
\end{aligned}
$$

As we know the resolvent ω is a block upper triangular semi-infinite matrix with all its superdiagonals equal to zero but for the first m. Thus, the $(\ell+m)$-th truncation gives a matrix of the following form

$$
\omega^{[\ell+m]} = \left(\begin{array}{c}\omega^{[l],[\ell+m]}\\ \hline 0\,\big|\,\omega^{[\ell,m]}\end{array}\right)
$$

where $\omega^{[\ell],[\ell+m]}$ is a truncation built up with the first ℓ block rows and the first $\ell+m$ block columns of the resolvent ω, that for ℓ big enough looks like

$$
\omega^{[\ell],[\ell+m]} := \begin{pmatrix}
\omega_{[0],[0]} & \omega_{[0],[1]} & \dots & \omega_{[0],[m]} & 0_{[0],[m+1]} & 0_{[0],[m+2]} & \dots & 0_{[0],[\ell-1]} & \dots & 0_{[0],[m+\ell-1]}\\
0_{[1],[0]} & \omega_{[1],[1]} & \dots & \omega_{[1],[m]} & \omega_{[1],[m+1]} & 0_{[1],[m+2]} & \dots & 0_{[1],[\ell-1]} & \dots & 0_{[1],[m+\ell-1]}\\
\vdots & & & & & & \ddots & \vdots & & \\
0_{[\ell-1],[0]} & 0_{[\ell-1],[1]} & \dots & 0_{[\ell-1],[m]} & 0_{[\ell-1],[m+1]} & 0_{[\ell-1],[m+2]} & \dots & \omega_{[\ell-1],[\ell-1]} & \dots & \omega_{[\ell-1],[m+\ell-1]}
\end{pmatrix}.
$$

Then,

$$\omega^{[\ell+m]}P^{[\ell+m]}=\left(\begin{array}{c|c}\multicolumn{2}{c}{\omega^{[l],[\ell+m]}}\\ \hline 0&\omega^{[\ell,m]}\end{array}\right)P^{[\ell+m]}(\boldsymbol{x})=\left(\begin{array}{c}\omega^{[\ell],[\ell+m]}P^{[\ell+m]}(\boldsymbol{x})\\ \hline \omega^{[\ell,m]}\begin{pmatrix}P_{[\ell]}(\boldsymbol{x})\\ \vdots\\ P_{\ell+m}(\boldsymbol{x})\end{pmatrix}\end{array}\right)$$

It is important to notice that each row of the truncation $\omega^{[\ell],[\ell+m]}$ contains the complete nontrivial part of the corresponding row of the resolvent; i.e.

$$\begin{aligned}\omega^{[\ell],[\ell+m]}P^{[\ell+m]}(\boldsymbol{x})=&(\omega P(\boldsymbol{x}))^{[\ell]},\\ =&\mathcal{Q}(\boldsymbol{x})TP^{[\ell]}(\boldsymbol{x}).\end{aligned}$$

Therefore

$$\begin{aligned}&P^{[\ell+m]}(\boldsymbol{x})^\top\left(H^{[\ell+m]}\right)^{-1}M^{[\ell+m]}TP^{[\ell+m]}(\boldsymbol{y})\\ &=\left(\begin{array}{c}\mathcal{Q}(\boldsymbol{x})TP^{[\ell]}(\boldsymbol{x})\\ \hline \omega^{[\ell,m]}\begin{pmatrix}P_{[\ell]}(\boldsymbol{x})\\ \vdots\\ P_{\ell+m}(\boldsymbol{x})\end{pmatrix}\end{array}\right)^\top\begin{pmatrix}(TH^{[\ell]})^{-1}TP^{[\ell]}(\boldsymbol{y})\\ (TH_{[\ell]})^{-1}TP_{[\ell]}(\boldsymbol{y})\\ \vdots\\ (TH_{[\ell+m]})^{-1}TP_{\ell+m}(\boldsymbol{y})\end{pmatrix}\\ &=\mathcal{Q}(\boldsymbol{x})(TP^{[\ell]}(\boldsymbol{x}))^\top(TH^{[\ell]})^{-1}TP^{[\ell]}(\boldsymbol{y})\\ &\quad+\begin{pmatrix}P_{[\ell]}(\boldsymbol{x})\\ \vdots\\ P_{\ell+m}(\boldsymbol{x})\end{pmatrix}^\top\left(\omega^{[\ell,m]}\right)^\top\begin{pmatrix}P_{[\ell]}(\boldsymbol{x})\\ \vdots\\ P_{\ell+m}(\boldsymbol{x})\end{pmatrix}.\qquad\square\end{aligned}$$

D.12. Proof of Proposition 4.1.4

Proof. We first notice that

$$\begin{aligned}\frac{\partial S}{\partial t_a}=&\frac{\partial\beta^{(1)}}{\partial t_a}+\frac{\partial\beta^{(2)}}{\partial t_a}+\cdots\\ S^{-1}=&\mathbb{I}-\beta^{(1)}\underbrace{-\beta^{(2)}+\left(\beta^{(1)}\right)^2}_{\text{second subdiagonal}}+\cdots\end{aligned}$$

so that we can split the right t_a derivative of S into subdiagonals as follows

$$\begin{aligned}\frac{\partial S}{\partial t_a}S^{-1}&=\left(\frac{\partial\beta^{(1)}}{\partial t_a}+\frac{\partial\beta^{(2)}}{\partial t_a}+\cdots\right)\left(\mathbb{I}-\beta^{(1)}-\beta^{(2)}+\left(\beta^{(1)}\right)^2\right)\\ &=\frac{\partial\beta^{(1)}}{\partial t_a}+\underbrace{\frac{\partial\beta^{(2)}}{\partial t_a}-\frac{\partial\beta^{(1)}}{\partial t_a}\beta^{(1)}}_{\text{second subdiagonal}}\end{aligned}$$

$$+\underbrace{\frac{\partial \beta^{(3)}}{\partial t_a}-\frac{\partial \beta^{(2)}}{\partial t_a}\beta^{(1)}-\frac{\partial \beta^{(1)}}{\partial t_a}\beta^{(2)}+\frac{\partial \beta^{(1)}}{\partial t_a}\big(\beta^{(1)}\big)^2}_{\text{third subdiagonal}}+\cdots.$$

Now, recalling that the basic Jacobi operators have only a nonvanishing subdiagonal from (4.1.7) we get

$$\begin{aligned}\frac{\partial \beta_{[k]}}{\partial t_a}&=J_{[k],[k-1]},\\ \frac{\partial \beta^{(2)}}{\partial t_a}&=\frac{\partial \beta^{(1)}}{\partial t_a}\beta^{(1)},\\ \frac{\partial \beta^{(3)}}{\partial t_a}&=\frac{\partial \beta^{(2)}}{\partial t_a}\beta^{(1)}+\frac{\partial \beta^{(1)}}{\partial t_a}\beta^{(2)}-\frac{\partial \beta^{(1)}}{\partial t_a}\big(\beta^{(1)}\big)^2. \qquad \square\end{aligned}$$

D.13. Proof of Proposition 4.2.1

Proof. We first consider the expressions for the Baker functions Ψ_1, Ψ_2 and the adjoint Baker function Ψ_2^*:

$$\begin{aligned}\Psi_1(\boldsymbol{z},t,\boldsymbol{m})&=S(t,\boldsymbol{m})\mathrm{e}^{t(\boldsymbol{z})}\Big[\prod_{a=1}^{D}\big((\boldsymbol{n}_a\cdot\boldsymbol{x})-q_a\big)^{m_a}\Big]\chi(\boldsymbol{z})\\ &=\mathrm{e}^{t(\boldsymbol{z})}\Big[\prod_{a=1}^{D}\big((\boldsymbol{n}_a\cdot\boldsymbol{x})-q_a\big)^{m_a}\Big]P(\boldsymbol{z},t,\boldsymbol{m}),\\ \Psi_2^*(\boldsymbol{z},t,\boldsymbol{m})&=H(t,\boldsymbol{m})^{-1}S(t,\boldsymbol{m})\chi(\boldsymbol{z})=H(t,\boldsymbol{m})^{-1}P(\boldsymbol{z},t,\boldsymbol{m}),\\ \Psi_2(\boldsymbol{z},t)&=H(t,\boldsymbol{m})(S(t,\boldsymbol{m})^{-1})^\top\chi^*(\boldsymbol{z})=C(\boldsymbol{z},t,\boldsymbol{m}).\end{aligned}$$

For the remaining Baker function Ψ_1^* we proceed as follows

$$\begin{aligned}\Psi_1^*(\boldsymbol{z},t,m)&=\big[(W_1(t,\boldsymbol{m}))^{-1}\big]^\top\chi=(W_2(\boldsymbol{z},t,\boldsymbol{m})^{-1})^\top G^\top\chi^*(\boldsymbol{z})\\ &=H(t,\boldsymbol{m})^{-1}S(t,\boldsymbol{m})G\chi^*(\boldsymbol{z}),\end{aligned}$$

and recall the proof of Proposition 2.5.1 where we replace $S\to S(t,\boldsymbol{m})$ but keep G (not replacing it by $G(t,\boldsymbol{m})$) to get

$$\Psi_1^*(\boldsymbol{z},t,\boldsymbol{m})=H(t,\boldsymbol{m})^{-1}\int_\Omega P(\boldsymbol{y},t,\boldsymbol{m})\mathrm{d}\mu(\boldsymbol{y})\frac{1}{(z_1-y_1)\cdots(z_D-y_D)},$$

and we get the desired result. $\square$

D.14. Proof of Proposition 4.5.1

Proof. From (4.1.3) we get

$$S(t)W_0(t)G = H(t)(S(t)^{-1})^\top,$$

that, by differentiation, leads to

$$\frac{\partial S}{\partial t_a}S^{-1} + S\Lambda_a S^{-1} = \frac{\partial H}{\partial t_a}H^{-1} - H\Big(\frac{\partial S}{\partial t_a}S^{-1}\Big)^\top H^{-1}.$$

Then, we have

$$(J_a)_{[k],[k]} = \beta_{[k]}(\Lambda_a)_{[k-1],[k]} - (\Lambda_a)_{[k],[k+1]}\beta_{[k+1]} = \frac{\partial H_{[k]}}{\partial t_a}H_{[k]}^{-1},$$
$$(J_a)_{[k],[k+1]} = (\Lambda_a)_{[k],[k+1]} = -H_{[k]}\frac{\partial(\beta_{[k+1]})^\top}{\partial t_a}H_{[k+1]}^{-1}. \qquad \square$$

D.15. Proof of Theorem 5.2.1

Proof. We obviously have

$$\Big[\frac{\partial}{\partial \boldsymbol{n}_b} - T_b, \frac{\partial}{\partial \boldsymbol{n}_a} - T_a\Big](W_1) = 0, \qquad a,b \in \{1,\dots,D\}.$$

Recalling the proof of Proposition 5.2.2 we can write

$$\begin{aligned}\Big[\frac{\partial}{\partial \boldsymbol{n}_b} - T_b, \frac{\partial}{\partial \boldsymbol{n}_a} - T_a\Big](W_1) &= \Big(\frac{\partial \Delta_a\beta(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})}{\partial \boldsymbol{n}_b} - \frac{\partial \Delta_b\ \beta(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})}{\partial \boldsymbol{n}_a}\Big)(W_1)\\ &\quad + \big(-q_a + (\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\big)\frac{\partial W_1}{\partial \boldsymbol{n}_b} - \big(-q_b + (\Delta_b\beta)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})\big)\frac{\partial W_1}{\partial \boldsymbol{n}_a}\\ &\quad - (-q_a + T_b(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda}))(T_bW_1) + (-q_b + T_a(\Delta_b\beta)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda}))(T_aW_1).\end{aligned}$$

To evaluate this expression we recall Proposition 5.2.1 that splits it by diagonals giving

$$\begin{aligned}&(\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda}) - (\Delta_b\beta)(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\\ &\quad = (T_b\Delta_a\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda}) - (T_a\Delta_b\beta)((\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})\end{aligned}$$

or

$$(\Delta_a\Delta_b\beta)(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_b\cdot\boldsymbol{\Lambda}) = (\Delta_b\Delta_a\beta)((\boldsymbol{n}_b\cdot\boldsymbol{\Lambda})(\boldsymbol{n}_a\cdot\boldsymbol{\Lambda})$$

which happens to be an identity. Next we look at the main diagonal where we have

$$\Delta_b \left[\frac{\partial \beta}{\partial \boldsymbol{n}_a} + (\Delta_a \beta)\big(q_a + (\boldsymbol{n_a} \cdot \boldsymbol{\Lambda})\beta\big)\right] \boldsymbol{n_b} \cdot \boldsymbol{\Lambda}$$
$$= \Delta_a \left[\frac{\partial \beta}{\partial \boldsymbol{n}_b} + (\Delta_b \beta)\big(q_b + (\boldsymbol{n_b} \cdot \boldsymbol{\Lambda})\beta\big)\right] \boldsymbol{n_a} \cdot \boldsymbol{\Lambda}. \qquad \square$$

D.16. Proof of Theorem 5.3.1

Proof. If we denote

$$L_{a,b} := \partial_a \partial_b + U_{a,b}, \quad a, b \in \{1, \dots, D\}, \tag{D.16.1}$$

(5.3.2) reads as $\partial_{a,b}(W_i) = L_{a,b}(W_i)$. The compatibility conditions for this linear system are

$$\Big(\partial_{(a,b)}(L_{c,d}) - \partial_{(c,d)}(L_{a,b}) + [L_{c,d}, L_{a,b}]\Big)(W_i) = 0, \qquad i = 1, 2,$$

and consequently

$$R_{a,b,c,d}(W_i) = 0, \quad a, b, c, d \in \{1, \dots, D\}, \tag{D.16.2}$$

where

$$\begin{aligned} R_{a,b,c,d} := {}& (\partial_b U_{c,d})\partial_a + (\partial_a U_{c,d})\partial_b - (\partial_d U_{a,b})\partial_c - (\partial_c U_{a,b})\partial_d - \partial_{(a,b)}(U_{c,d}) + \partial_{(c,d)}(U_{a,b}) \\ & + [U_{c,d}, U_{a,b}] - \partial_c \partial_d U_{a,b} + \partial_a \partial_b U_{c,d}. \end{aligned}$$

Notice that we have that W_1 satisfies an equation of the form

$$\Big(\sum_{j=1}^{D} B_j \partial_j + A\Big)(W_1) = 0.$$

Recalling (5.2.1), we find that

$$0 = \Big(\sum_{j=1}^{D} B_j \partial_j + A\Big)(W_1) = \Big(\underbrace{\sum_{j=1}^{D} B_j \Lambda_j}_{\text{first superdiagonal}} + \underbrace{\sum_{j=1}^{D} B_j \beta \Lambda_j + A}_{\text{diagonal}} + \mathfrak{l}\Big) W_0$$

and, decoupling by diagonals, we get

$$\sum_{j=1}^{D} B_j \Lambda_j = 0, \qquad \sum_{j=1}^{D} B_j \beta \Lambda_j + A = 0.$$

From (D.16.2) we find, in the first place, that

$$(\partial_b U_{c,d})\Lambda_a + (\partial_a U_{c,d})\Lambda_b - (\partial_d U_{a,b})\Lambda_c - (\partial_c U_{a,b})\Lambda_d = 0,$$

which is identically satisfied because of (5.3.1). In the second place, we get the following nonlinear equation

$$\begin{aligned}&(\partial_b U_{c,d})\beta\Lambda_a + (\partial_a U_{c,d})\beta\Lambda_b - (\partial_d U_{a,b})\beta\Lambda_c - (\partial_c U_{a,b})\beta\Lambda_d + \partial_{(a,b)}U_{c,d} - \partial_{(c,d)}U_{a,b}\\&\quad + [U_{c,d}, U_{a,b}] - \partial_c\partial_d U_{a,b} + \partial_a\partial_b U_{c,d} = 0,\end{aligned}$$

and recalling (5.3.1) we get the desired result. □

D.17. Proof of Proposition 6.2.3

Proof. Observe that for $\boldsymbol{n} = R\boldsymbol{e}_a$ we have (6.2.1) that

$$M = (\boldsymbol{n}\cdot\boldsymbol{\Lambda})\mathcal{M}^{-1/2} = \eta_R\Lambda_a\eta_R^{-1}\mathcal{M}^{-1/2},$$

from where we deduce that

$$\begin{aligned}MM^\top &= \eta_R\Lambda_a\eta_R^{-1}\mathcal{M}^{-1}\big(\eta_R^{-1}\big)^\top\Lambda_a^\top\eta_R^\top\\&= \eta_R\Lambda_a[\mathscr{R}]_{B_c}^\top\mathcal{M}^{-1}[\mathscr{R}]_{B_c}\Lambda_a^\top\eta_R^\top && \text{because (6.1.3)}\\&= \eta_R\Lambda_a\mathcal{M}^{-1}\Lambda_a^\top\eta_R^\top && \text{from (6.1.1)}\end{aligned}$$

We now introduce the matrix built up of multinomial coefficients involving $\boldsymbol{e}_a$,

$$\mathcal{M}_{[k+1]_a} = \operatorname{diag}\left(\binom{k+1}{\alpha_{j_1}^{(k+1)}},\dots,\binom{k+1}{\alpha_{j_{|[k]|}}^{(k+1)}}\right) \in \mathbb{R}^{|[k]|\times|[k]|},$$

where $[k+1]_a := \{\alpha_{j_m}^{(k+1)}\}_{m=1}^{|[k]|} \subset [k+1]$ is the set containing only the multi-indices such that $\boldsymbol{e}_a\cdot\alpha_{j_m}^{(k+1)} \neq 0$, assuming the reverse lexicographical order; notice that $|[k+1]_a| = |[k]|$. Then, we can write

$$(\Lambda_a)_{[k],[k+1]}\mathcal{M}_{[k+1]}^{-1}((\Lambda_a)_{[k],[k+1]})^\top = \mathcal{M}_{[k+1]_a}^{-1},$$

which is clearly invertible and, consequently

$$M_{[k],[k+1]}(M_{[k],[k+1]})^\top = \eta_{R,[k]}\mathcal{M}_{[k+1]_a}^{-1}\eta_{R,[k]}^\top$$

is invertible.

Thus, following Appendix B.1, we get the Moore–Penrose pseudo-inverse of the matrix $(\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}$ is

$$\begin{aligned}&\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1/2}\big)^{+}\\&\quad=\mathcal{M}_{[k]}^{-1/2}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\big)^{\top}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\mathcal{M}_{[k]}^{-1}\big((\boldsymbol{n}\cdot\boldsymbol{\Lambda})_{[k-1],[k]}\big)^{\top}\big)\big)^{-1}\end{aligned}$$

which is the right inverse of the matrix, and therefore we get the result. □

D.18. Proof of Proposition 6.3.1

Proof. From Definition 2.3.1 we have

$$\begin{aligned}G&=\int_{\mathbb{R}^D}\chi(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})\chi(\boldsymbol{x})^{\top}\\&=\int_{\mathbb{R}^D}\chi(R\boldsymbol{x})\mathrm{d}\mu(R\boldsymbol{x})\chi(R\boldsymbol{x})^{\top}\\&=\int_{\mathbb{R}^D}\eta_R\chi(\boldsymbol{x})\mathrm{d}\mu(\boldsymbol{x})\chi(\boldsymbol{x})^{\top}\eta_R^{\top}\\&=\eta_R G\eta_R^{\top}.\end{aligned}$$

From this formula and the Cholesky factorization we get

$$\begin{aligned}SH(S^{-1})^{\top}&=\eta_R SH(S^{-1})^{\top}\eta_R^{\top}\\&=\eta_R S\eta_R^{-1}\eta_R H\eta_R^{\top}(\eta_R^{\top})^{-1}(S^{-1})^{\top}\eta_R^{\top}\\&=\big(\eta_R S\eta_R^{-1}\big)\big(\eta_R H\eta_R^{\top}\big)\Big(\big(\eta_R S\eta_R^{-1}\big)^{\top}\Big)^{-1},\end{aligned}$$

and given the uniqueness of the Cholesky factorization and that η_R is block diagonal we get the stated result for S and H. The equation for β follows from the equation for S. □

D.19. Proof of Proposition 5.4.1

Proof. For convenience we write here the next couple of equations

$$\begin{aligned}\partial_a W_1&=(\partial_a S+S\Lambda_a)W_0,\\(\partial_{(a,b,c)}-\partial_a\partial_b\partial_c)(W_1)&=\big(\partial_{a,b,c}S-\partial_a\partial_b\partial_c S-\partial_a\partial_b S\Lambda_c-\partial_b\partial_c S\Lambda_a-\partial_c\partial_a S\Lambda_b\\&\qquad-\partial_a S\Lambda_b\Lambda_c-\partial_b S\Lambda_c\Lambda_a-\partial_c S\Lambda_a\Lambda_b\big)W_0.\end{aligned}$$

Taking into account the form of $S=\mathbb{I}+\beta^{(1)}+\beta^{(2)}+\cdots$, $\beta^{(k)}$ being the k-th subdiagonal of S, we can write

$$\partial_a W_1=(\Lambda_a+\beta^{(1)}\Lambda_a)W_0+\mathfrak{l}W_0,\tag{D.19.1}$$

$$(\partial_{(a,b,c)}-\partial_a\partial_b\partial_c)(W_1)=-\big(\overbrace{\partial_a\beta^{(1)}\Lambda_b\Lambda_c+\partial_b\beta^{(1)}\Lambda_c\Lambda_a+\partial_c\beta^{(1)}\Lambda_a\Lambda_b}^{\text{first superdiagonal}}\tag{D.19.2}$$

$$
\begin{aligned}
&+\underbrace{\partial_a\beta^{(2)}\Lambda_b\Lambda_c+\partial_b\beta^{(2)}\Lambda_c\Lambda_a+\partial_c\beta^{(2)}\Lambda_a\Lambda_b}_{\text{diagonal}}\\
&+\underbrace{\partial_a\partial_b\beta^{(1)}\Lambda_c+\partial_b\partial_c\beta^{(1)}\Lambda_a+\partial_c\partial_a\beta^{(1)}\Lambda_b}_{\text{diagonal}}\big)W_0\\
&+\mathfrak{l}W_0.
\end{aligned}
$$

We can use (D.19.1) to move to the RHS of (D.19.2) the contribution on the first super-diagonal so that

$$
\begin{aligned}
&(\partial_{(a,b,c)}-\partial_a\partial_b\partial_c+\partial_a\beta^{(1)}\Lambda_b\partial_c+\partial_b\beta^{(1)}\Lambda_c\partial_a+\partial_c\beta^{(1)}\Lambda_a\partial_b)(W_1)\\
&=-\big(\overbrace{\partial_a\beta^{(2)}\Lambda_b\Lambda_c+\partial_b\beta^{(2)}\Lambda_c\Lambda_a+\partial_c\beta^{(2)}\Lambda_a\Lambda_b+\partial_a\partial_b\beta^{(1)}\Lambda_c+\partial_b\partial_c\beta^{(1)}\Lambda_a+\partial_c\partial_a\beta^{(1)}\Lambda_b}^{\text{diagonal}}\\
&\qquad\underbrace{-\partial_a\beta^{(1)}\Lambda_b\beta^{(1)}\Lambda_c-\partial_b\beta^{(1)}\Lambda_c\beta^{(1)}\Lambda_a-\partial_c\beta^{(1)}\Lambda_a\beta^{(1)}\Lambda_b}_{\text{diagonal}}\big)W_0+\mathfrak{l}W_0,
\end{aligned}
$$

which using Proposition 4.1.4 can be written as[19]

$$
\begin{aligned}
&(\partial_{(a,b,c)}-\partial_a\partial_b\partial_c+\partial_a\beta^{(1)}\Lambda_b\partial_c+\partial_b\beta^{(1)}\Lambda_c\partial_a+\partial_c\beta^{(1)}\Lambda_a\partial_b)(W_1)\\
&=-\big(\overbrace{\partial_a\partial_b\beta^{(1)}\Lambda_c+\partial_b\partial_c\beta^{(1)}\Lambda_a+\partial_c\partial_a\beta^{(1)}\Lambda_b}^{\text{diagonal}}\\
&\quad\underbrace{+(\partial_a\beta^{(1)})\beta^{(1)}\Lambda_b\Lambda_c+(\partial_b\beta^{(1)})\beta^{(1)}\Lambda_c\Lambda_a+(\partial_c\beta^{(1)})\beta^{(1)}\Lambda_a\Lambda_b}_{\text{diagonal}}\\
&\qquad\underbrace{-\partial_a\beta^{(1)}\Lambda_b\beta^{(1)}\Lambda_c-\partial_b\beta^{(1)}\Lambda_c\beta^{(1)}\Lambda_c-\partial_c\beta^{(1)}\Lambda_a\beta^{(1)}\Lambda_b}_{\text{diagonal}}\big)W_0+\mathfrak{l}W_0,
\end{aligned}
$$

which after simplifying and writing $\beta^{(1)}=\beta$

$$
\begin{aligned}
&(\partial_{(a,b,c)}-\partial_a\partial_b\partial_c+\partial_a\beta\Lambda_b\partial_c+\partial_b\beta\Lambda_c\partial_a+\partial_c\beta\Lambda_a\partial_b)(W_1)\\
&\quad=-\big(\partial_a\partial_b\beta\Lambda_c+\partial_b\partial_c\beta\Lambda_a+\partial_c\partial_a\beta\Lambda_b\\
&\qquad+(\partial_a\beta)[\beta,\Lambda_b]\Lambda_c+(\partial_b\beta)[\beta,\Lambda_c]\Lambda_a+(\partial_c\beta)[\beta,\Lambda_a]\Lambda_b\big)W_0+\mathfrak{l}W_0.
\end{aligned}
$$

Hence, from (5.3.1)

$$
\begin{aligned}
R_1:=(&\partial_{(a,b,c)}-\partial_a\partial_b\partial_c+V_{a,b}\partial_c+V_{b,c}\partial_a+V_{c,a}\partial_b\\
&+\partial_c(V_{a,b})+\partial_a(V_{b,c})+\partial_b(V_{c,a})+V_{a,b,c}+V_{b,c,a}+V_{c,b,a})(W_1)\in\mathfrak{l}W_0,
\end{aligned}
$$

and trivially we know that

[19] This could be avoided, depending on whether or not we desired to use $\beta^{(2)}$ in the expressions for $V_{a,b,c}$.

$$
\begin{aligned}
R_2 := (&\partial_{(a,b,c)} - \partial_a\partial_b\partial_c + V_{a,b}\partial_c + V_{b,c}\partial_a + V_{c,a}\partial_b \\
&+ \partial_c(V_{a,b}) + \partial_a(V_{b,c}) + \partial_b(V_{c,a}) + V_{a,b,c} + V_{b,c,a} + V_{c,b,a})(W_2) \in \mathfrak{u},
\end{aligned}
$$

where $R_1 G = R_2$. Consequently, from the *asymptotic module* Proposition 5.1.1 we deduce that $R_1 = R_2 = 0$. Therefore,

$$
\begin{aligned}
\frac{\partial \Psi_i}{\partial t_{(a,b,c)}} = &\frac{\partial^3 \Psi_i}{\partial t_a \partial t_b \partial t_c} - V_{a,b}\frac{\partial \Psi_i}{\partial t_c} - V_{b,c}\frac{\partial \Psi_i}{\partial t_a} - V_{c,a}\frac{\partial \Psi_i}{\partial t_b} \\
&- \Big(\partial_c(V_{a,b}) + \partial_a(V_{b,c}) + \partial_b(V_{c,a}) + V_{a,b,c} + V_{b,c,a} + V_{c,b,a}\Big)\Psi_i. \qquad \square
\end{aligned}
$$

D.20. Proof of Proposition 6.3.2

Proof. To prove (6.3.2) for the MVOPR observe that

$$
\begin{aligned}
P(R\boldsymbol{x}) &= S\eta_R\chi(\boldsymbol{x}) \\
&= \eta_R S\chi(\boldsymbol{x}) \\
&= \eta_R P(\boldsymbol{x})
\end{aligned}
$$

which together with (6.1.2) leads to the result. To check (6.3.3) just follow the next equalities

$$
\begin{aligned}
\eta_R(\boldsymbol{n}\cdot\boldsymbol{J})\eta_R^{-1} &= \eta_R S(\boldsymbol{n}\cdot\boldsymbol{\Lambda})S^{-1}\eta_R^{-1} \\
&= S\eta_R(\boldsymbol{n}\cdot\boldsymbol{\Lambda})\eta_R^{-1}S^{-1} \\
&= S(R\boldsymbol{n}\cdot\boldsymbol{\Lambda})S^{-1} \\
&= R\boldsymbol{n}\cdot\boldsymbol{J}.
\end{aligned}
$$

Equations (6.3.4) are a direct consequence of (2.8.1) and (6.3.1). $\square$

References

[1] M. Adler, P. van Moerbeke, Group factorization, moment matrices and Toda lattices, Int. Math. Res. Not. IMRN 12 (1997) 556–572.
[2] M. Adler, P. van Moerbeke, Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems, Comm. Math. Phys. 207 (1999) 589–620.
[3] M. Adler, P. van Moerbeke, Vertex operator solutions to the discrete KP hierarchy, Comm. Math. Phys. 203 (1999) 185–210.
[4] M. Adler, P. van Moerbeke, The spectrum of coupled random matrices, Ann. of Math. 149 (1999) 921–976.
[5] M. Adler, P. van Moerbeke, Hermitian, symmetric and symplectic random ensembles: PDEs for the distribution of the spectrum, Ann. of Math. 153 (2001) 149–189.
[6] M. Adler, P. van Moerbeke, Darboux transforms on band matrices, weights and associated polynomials, Int. Math. Res. Not. IMRN 18 (2001) 935–984.
[7] M. Adler, P. van Moerbeke, P. Vanhaecke, Moment matrices and multi-component KP, with applications to random matrix theory, Comm. Math. Phys. 286 (2009) 1–38.

[8] M. Álvarez de Morales, L. Fernández, T.E. Pérez, M.A. Piñar, Semiclassical orthogonal polynomials in two variables, J. Comput. Appl. Math. 207 (2007) 323–330.

[9] C. Álvarez Fernández, G. Ariznabarreta, J.C. García-Ardila, M. Mañas, F. Marcellán, Christoffel transformations for matrix orthogonal polynomials in the real line and the non-Abelian 2D Toda lattice hierarchy, Int. Math. Res. Not. IMRN (2016), http://dx.doi.org/10.1093/imrn/rnw027.

[10] C. Álvarez Fernández, G. Ariznabarreta, J.C. García-Ardila, M. Mañas, F. Marcellán, Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line, arXiv:1605.04617.

[11] C. Álvarez-Fernández, M. Mañas, Orthogonal Laurent polynomials on the unit circle, extended CMV ordering and 2D Toda type integrable hierarchies, Adv. Math. 240 (2013) 132–193.

[12] C. Álvarez-Fernández, M. Mañas, On the Christoffel–Darboux formula for generalized matrix orthogonal polynomials, J. Math. Anal. Appl. 418 (2014) 238–247.

[13] C. Álvarez-Fernández, U.F. Prieto, M. Mañas, The multicomponent 2D Toda hierarchy: generalized matrix orthogonal polynomials, multiple orthogonal polynomials and Riemann–Hilbert problems, Inverse Probl. 26 (2010) 055009 (15 pp.).

[14] C. Álvarez-Fernández, U.F. Prieto, M. Mañas, Multiple orthogonal polynomials of mixed type: Gauss–Borel factorization and the multi-component 2D Toda hierarchy, Adv. Math. 227 (2011) 1451–1525.

[15] G. Araznibarreta, M. Mañas, Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems, Adv. Math. 264 (2014) 396–463.

[16] G. Araznibarreta, M. Mañas, A Jacobi type Christoffel–Darboux formula for multiple orthogonal polynomials of mixed type, Linear Algebra Appl. 468 (2015) 154–170.

[17] G. Ariznabarreta, M. Mañas, Darboux transformations for multivariate orthogonal polynomials, arXiv:1503.04786.

[18] G. Ariznabarreta, M. Mañas, Linear spectral transformations for multivariate orthogonal polynomials and multispectral Toda hierarchies, arXiv:1511.09129.

[19] G. Ariznabarreta, M. Mañas, Multivariate orthogonal Laurent polynomials and integrable systems, arXiv:1506.08708.

[20] R. Barrio, J.M. Peña, T. Sauer, Three term recurrence for the evaluation of multivariate orthogonal polynomials, J. Approx. Theory 162 (2010) 407–420.

[21] H.G.W. Begehr, A. Dzhuraev, An Introduction to Several Complex Variables and Partial Differential Equations, Pitman Monographs and Surveys in Pure and Applied Mathematics, vol. 88, Addison–Wesley Longman Limited, Harlow, 1997.

[22] A. Ben-Israel, T.N.E. Greville, Generalized Inverses. Theory and Applications, 2nd edition, Springer-Verlag, New York, 2003.

[23] B. Benouahmane, A. Cuyt, Multivariate orthogonal polynomials, homogeneous Padé approximants and Gaussian cubature, Numer. Algorithms 24 (2000) 1–15.

[24] M.J. Bergvelt, A.P.E. ten Kroode, Tau-functions and zero-curvature equations of Toda-AKNS type, J. Math. Phys. 29 (1988) 1308–1320.

[25] M.J. Bergvelt, A.P.E. ten Kroode, Partitions, vertex operators constructions and multi-component KP equations, Pacific J. Math. 171 (1995) 23–88.

[26] A. Bjerhammar, Application of Calculus of Matrices to Method of Least Squares: With Special References to Geodetic Calculations, Transactions of the Royal Institute of Technology, vol. 49, KTH, Stockholm, 1951.

[27] W.D. Blizard, Multiset theory, Notre Dame J. Form. Log. 30 (1) (1989).

[28] M.I. Bueno, F. Marcellán, Darboux transformation and perturbation of linear functionals, Linear Algebra Appl. 384 (2004) 215–242.

[29] M.I. Bueno, F. Marcellán, Polynomial perturbations of bilinear functionals and Hessenberg matrices, Linear Algebra Appl. 414 (2006) 64–83.

[30] E.A. Carlena, J.S. Geronimo, M. Loss, On the Markov sequence problem for Jacobi polynomials, Adv. Math. 226 (2011) 3426–3466.

[31] G.A. Cassatella-Contra, M. Mañas, Riemann–Hilbert problems, matrix orthogonal polynomials and discrete matrix equations with singularity confinement, Stud. Appl. Math. 128 (2011) 252–274.

[32] G.A. Cassatella-Contra, M. Mañas, P. Tempesta, Singularity confinement for matrix discrete Painlevé equations, Nonlinearity 27 (2014) 2321–2335.

[33] T.F. Chan, P.S. Vassilevski, A framework for block ILU factorizations using block-size reduction, Math. Comp. 64 (1995) 129–156.

[34] E.B. Christoffel, Über die Gaußische Quadratur und eine Verallgemeinerung derselben, J. Reine Angew. Math. 55 (1858) 61–82.

[35] P. Comon, G. Golub, L.-H. Lim, B. Mourrain, Symmetric tensors and symmetric tensor rank, SIAM J. Matrix Anal. Appl. 30 (2008) 1254–1279.
[36] G. Darboux, Sur une proposition relative aux équations linéaires, C. R. Hebdomadaires des Séances Acad. Sci. Paris 94 (1882) 1456–1459.
[37] G. Darboux, Leçons sur la Théorie Générale des Surfaces, 2éme partie, Gauthier-Villars, Paris, 1889.
[38] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Operator approach to the Kadomtsev–Petviashvili equation. Transformation groups for soliton equations. III, J. Phys. Soc. Jpn. 50 (1981) 3806–3812.
[39] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Transformation groups for soliton equations. Euclidean Lie algebras and reduction of the KP hierarchy, Publ. Res. Inst. Math. Sci. 18 (1982) 1077–1110.
[40] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Transformation groups for soliton equations, in: M. Jimbo, T. Miwa (Eds.), Nonlinear Integrable Systems-Classical Theory and Quantum Theory, World Scientific, Singapore, 1983.
[41] A.M. Delgado, J.S. Geronimo, P. Iliev, F. Marcellán, Two variable orthogonal polynomials and structured matrices, SIAM J. Matrix Anal. Appl. 28 (2006) 118–147.
[42] A. Doliwa, P.M. Santini, M. Mañas, Transformations of quadrilateral lattices, J. Math. Phys. 41 (2000) 944–990.
[43] I. Dumitriua, A. Edelmanb, G. Shumanc, MOPS: multivariate orthogonal polynomials (symbolically), J. Symbolic Comput. 42 (2007) 587–620.
[44] C.F. Dunkl, Y. Xu, Orthogonal Polynomials of Several Variables, Encyclopedia of Mathematics and its Applications, vol. 81, Cambridge University Press, Cambridge, 2001.
[45] L.P. Eisenhart, Transformations of Surfaces, Princeton University Press, Hamburg, 1943.
[46] P. Etinghof, I. Gel'fand, S. Gel'fand, V.S. Retakh, Factorization of differential operators, quasideterminants, and nonabelian Toda field equations, Math. Res. Lett. 4 (1997) 413–425.
[47] P. Etinghof, I. Gel'fand, S. Gel'fand, V.S. Retakh, Nonabelian integrable systems, quasideterminants, and Marchenko lemma, Math. Res. Lett. 5 (1998) 1–12.
[48] H. Federer, Geometric Measure Theory, Springer-Verlag, New York, 1969.
[49] R. Felipe, F. Ongay, Algebraic aspects of the discrete KP hierarchy, Linear Algebra Appl. 338 (2001) 1–17.
[50] L. Fernández, T.E. Pérez, M.A. Piñar, Weak classical orthogonal polynomials in two variables, J. Comput. Appl. Math. 178 (2005) 191–203.
[51] L. Fernández, T.E. Pérez, M.A. Piñar, Y. Xu, Krall-type orthogonal polynomials in several variables, J. Comput. Appl. Math. 233 (2010) 1519–1524.
[52] H. Flaschka, The Toda lattice. II. Inverse-scattering solution, Prog. Theor. Exp. Phys. 51 (1974) 703–718.
[53] H. Flaschka, The Toda lattice. II. Existence of integrals, Phys. Rev. B (Solid State) 9 (1974) 1924.
[54] I.M. Gel'fand, S. Gel'fand, V.S. Retakh, R. Wilson, Quasideterminants, Adv. Math. 193 (2005) 56–141.
[55] I.M. Gel'fand, D. Krob, A. Lascoux, B. Leclerc, V.S. Retakh, J.-Y. Thibon, Noncommutative symmetric functions, Adv. Math. 112 (1995) 218–348.
[56] I.M. Gel'fand, V.S. Retakh, Determinants of matrices over noncommutative rings, Funct. Anal. Appl. 25 (1991) 91–102.
[57] I.M. Gel'fand, V.S. Retakh, Theory of noncommutative determinants, and characteristic functions of graphs, Funct. Anal. Appl. 26 (1992) 231–246.
[58] I.M. Gel'fand, V.S. Retakh, Quasideterminants. I, Selecta Math. (N.S.) 3 (1997) 517–546.
[59] C.R. Gilson, S.R. Macfarlane, Dromion solutions of noncommutative Davey–Stewartson equations, J. Phys. A: Math. Theor. 42 (2009) 235202 (20 pp.).
[60] C.R. Gilson, J.J.C. Nimmo, Y. Ohta, Quasideterminant solutions of a non-Abelian Hirota–Miwa equation, J. Phys. A: Math. Theor. 40 (2007) 12607–12617.
[61] C.R. Gilson, J.C. Nimmo, C.M. Sooman, On a direct approach to quasideterminant solutions of a noncommutative modified KP equation, J. Phys. A: Math. Theor. 41 (2008) 085202 (10 pp.).
[62] C. Gu, H. Hu, Z. Zhou, Darboux Transformations in Integrable Systems Theory and Their Applications to Geometry, Springer-Verlag, Amsterdam, 2005.
[63] B. Haider, M. Hassan, Quasideterminant multisoliton solutions of a supersymmetric chiral field model in two dimensions, J. Phys. A: Math. Theor. 43 (2010) 035204.
[64] M. Hamanaka, Noncommutative solitons and quasideterminants, Phys. Scr. 89 (2014) 038006 (11 pp.).
[65] E.V. Haynsworth, On the Schur Complement, Basel Mathematical Notes, vol. 20, University of Basel, June 1968 (17 pp.).

[66] E.V. Haynsworth, Reduction of a matrix using properties of the Schur complement, Linear Algebra Appl. 3 (1970) 23–29.
[67] M. Henón, Integrals of the Toda lattice, Phys. Rev. B 9 (1974) 1921–1923.
[68] A. Heyting, Die Theorie der linearen Gleichungen in einer Zahlenspezies mit nichtkommutativer Multiplikation, Math. Ann. 98 (1928) 465–490.
[69] L. Hörmander, An Introduction to Complex Analysis in Several Variables, North-Holland, Amsterdam, 1990.
[70] P. Iliev, Y. Xu, Discrete orthogonal polynomials and difference equations of several variables, Adv. Math. 212 (2007) 1–36.
[71] M. Jaulent, M.A. Manna, L. Martínez Alonso, Asymptotic modules for solving integrable models, Inverse Probl. 5 (1989) 573–598.
[72] V.G. Kac, J.W. van de Leur, The n-component KP hierarchy and representation theory, J. Math. Phys. 44 (2003) 3245–3293.
[73] M. Kac, P. van Moerbeke, On some periodic Toda lattices (Floquet theory/inverse scattering problems), Proc. Natl. Acad. Sci. USA 72 (1975) 1627–1629.
[74] M. Kac, P. van Moerbeke, A complete solution of the periodic Toda problem (completely integrable systems/Abelian integrals), Proc. Natl. Acad. Sci. USA 72 (1975) 2879–2880.
[75] T. Kailath, A.H. Sayed, B. Hassibi, Linear Estimation, Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2000.
[76] S. Karlin, J. McGregor, Linear growth models with many types and multidimensional Hahn polynomials, in: R.A. Askey (Ed.), Theory and Application of Special Functions, in: Publication of the Mathematics Research Center (University of Wisconsin), vol. 35, Academic Press, 1975, pp. 261–288.
[77] A.W. Knapp, Lie Groups. Beyond an Introduction, Progress in Mathematics, vol. 140, Birkhäuser, Boston, 2002.
[78] G. Knese, Schur–Agler class rational inner functions on the tridisk, Proc. Amer. Math. Soc. 139 (2011) 4063–4072.
[79] S.G. Krantz, Function Theory of Several Complex Variables, AMS Chelsea Publishing, American Mathematical Society, Providence, Rhode Island, 2000.
[80] D. Krob, B. Leclerc, Minor identities for quasi-determinants and quantum determinants, Comm. Math. Phys. 169 (1995) 1–23.
[81] C.X. Li, J.J.C. Nimmo, Darboux transformations for a twisted derivation and quasideterminant solution to the super KdV equation, Proc. R. Soc. A 466 (2010) 2471–2493.
[82] C.X. Li, J.C. Nimmo, S. Shen, Quasideterminant solutions to a noncommutative q-difference two-dimensional Toda lattice equation, arXiv:1406.1330v1 [nlin.SI], 5 Jun 2014.
[83] H. Li, Y. Xu, Discrete Fourier analysis on fundamental domain of A_d lattice and on simplex in d-variables, J. Fourier Anal. Appl. 16 (2010) 383–433.
[84] S.V. Manakov, Complete integrability and stochastization of discrete dynamical systems, J. Exp. Theor. Phys. (Sov. Phys. JETP) 40 (1975) 269–274, Zh. Eksp. Teor. Fiz. 67 (1974) 543–552.
[85] M. Mañas, L. Martínez Alonso, The multicomponent 2D Toda hierarchy: dispersionless limit, Inverse Probl. 25 (2009) 115020 (22 pp.).
[86] M. Mañas, L. Martínez Alonso, C. Álvarez-Fernández, The multicomponent 2D Toda hierarchy: discrete flows and string equations, Inverse Probl. 25 (2009) 065007 (31 pp.).
[87] M. Mañas, L. Martínez Alonso, E. Medina, Dressing methods for geometric nets: I. Conjugate nets, J. Phys. A: Math. Gen. 33 (2000) 2871–2894.
[88] F. Marcellán, J. Hernández, Geronimus spectral transforms and measures on the complex plane, J. Comput. Appl. Math. 219 (2008) 441–456.
[89] V.B. Matveev, M.A. Salle, Differential–difference evolution equations. II: Darboux transformation for the Toda lattice, Lett. Math. Phys. 3 (1979) 425–429.
[90] P.R. Milch, A multi-dimensional linear growth birth and death process, Ann. Math. Stat. 39 (1968) 727–754.
[91] H. Minc, Permanents, Encyclopedia of Mathematics and its Applications, vol. 6, Addison–Wesley Publishing Company, Reading, Massachusetts, 1978.
[92] L. Miranian, Matrix-valued orthogonal polynomials on the real line: some extensions of the classical theory, J. Phys. A: Math. Gen. 38 (2005) 5731–5749.
[93] E.H. Moore, On the reciprocal of the general algebraic matrix, Bull. Amer. Math. Soc. 26 (1920) 394–395.
[94] T. Muir, A Treatise on the Theory of Determinants, Dover Phoenix Editions, Mineola, New York, 1960 & 2003.

[95] M. Mulase, Complete integrability of the Kadomtsev–Petviashvili equation, Adv. Math. 54 (1984) 57–66.
[96] P.J. Olver, On multivariate interpolation, Stud. Appl. Math. 116 (2006) 201–240.
[97] O. Ore, Linear equations in non-commutative fields, Ann. of Math. 32 (1931) 463–477.
[98] R. Penrose, A generalized inverse for matrices, Proc. Cambridge Philos. Soc. 51 (1955) 406–413.
[99] V. Retakh, V. Rubtsov, Noncommutative Toda chains, Hankel quasideterminants and Painlevé II equation, J. Phys. A: Math. Theor. 43 (2010) 505204 (13 pp.).
[100] A.R. Richardson, Hypercomplex determinants, Messenger of Math. 55 (1926) 145–152.
[101] A.R. Richardson, Simultaneous linear equations over a division algebra, Proc. Lond. Math. Soc. 28 (1928) 395–420.
[102] J. Rodal, I. Área, E. Godoy, Structure relations for monic orthogonal polynomials in two discrete variables, J. Math. Anal. Appl. 340 (2008) 825–844.
[103] C. Rogers, W.K. Schief, Bäcklund and Darboux Transformations: Geometry and Modern Applications in Soliton Theory, Cambridge University Press, Cambridge, 2002.
[104] M. Sato, Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds, in: Random Systems and Dynamical Systems, vol. 439, Research Institute for Mathematical Sciences Kokyuroku, 1981, pp. 30–46.
[105] M. Sato, Y. Sato, Soliton equations as dynamical systems on infinite-dimensional Grassmann manifold, in: Nonlinear Partial Differential Equations in Applied Science, Tokyo, 1982, in: Mathematical Studies, vol. 81, North-Holland, Amsterdam, 1983, pp. 259–271.
[106] V. Scheidemann, Introduction to Complex Analysis in Several Variables, Birkhäuser, Basel, 2005.
[107] I. Schur, Über Potenzreihen, die im Innern des Einheitskreises beschrankt sind [I], J. Reine Angew. Math. (Crelle's J.) 147 (1917) 205–232.
[108] R.P. Stanley, Enumerative Combinatorics. Vol. I, second edition, Cambridge University Press, Cambridge, 2011.
[109] G. Szegő, Orthogonal Polynomials, Colloquium Publications, vol. XXIII, American Mathematical Society, Providence, Rhode Island, 1939.
[110] K. Takasaki, T. Takebe, Universal Whitham hierarchy, dispersionless Hirota equations and multi-component KP hierarchy, Phys. D 235 (2007) 109–125.
[111] W. Thirring, Quantum Mechanics of Large Systems, Springer-Verlag, New York, 1983.
[112] M. Toda, Vibration of a chain with nonlinear interaction, J. Phys. Soc. Jpn. 22 (1967) 431–436.
[113] M. Toda, A soliton and two solitons in an exponential lattice and related equations, J. Phys. Soc. Jpn. 34 (1977) 18.
[114] K. Ueno, K. Takasaki, Toda lattice hierarchy. I, Proc. Japan Acad. Ser. A Math. Sci. 59 (1984) 167–170.
[115] K. Ueno, K. Takasaki, Toda lattice hierarchy. II, Proc. Japan Acad. Ser. A Math. Sci. 59 (1984) 215–218.
[116] K. Ueno, K. Takasaki, Toda lattice hierarchy, in: Group Representations and Systems of Differential Equations, in: Advanced Studies in Pure Mathematics, vol. 4, 1984, pp. 1–95.
[117] E.B. Vinberg, A Course in Algebra, Graduate Texts in Mathematics, vol. 56, American Mathematical Society, Providence, Rhode Island, 2003.
[118] S. Waldron, Recursive three-term recurrence relations for the Jacobi polynomials on a triangle, Constr. Approx. 33 (2011) 405–424.
[119] Y. Xu, On multivariate orthogonal polynomials, SIAM J. Math. Anal. 24 (1993) 783–794.
[120] Y. Xu, Multivariate orthogonal polynomials and operator theory, Trans. Amer. Math. Soc. 343 (1994) 193–202.
[121] Y. Xu, Recurrence formulas for multivariate orthogonal polynomials, Math. Comp. 62 (1994) 687–702.
[122] Y. Xu, Common zeros of polynomials in several variables and higher-dimensional quadrature, in: Pitman Research Notes in Mathematics Series, vol. 312, Longman Scientific & Technical, Harlow, 1994.
[123] Y. Xu, Orthogonal polynomials and cubature formulae on balls, simplices, and spheres, J. Comput. Appl. Math. 127 (2001) 349–368.
[124] Y. Xu, Lecture notes on orthogonal polynomials of several variables, in: Inzell Lectures on Orthogonal Polynomials, in: Advances in the Theory of Special Functions and Orthogonal Polynomials, vol. 2, Nova Science Publishers, 2004, pp. 141–196.
[125] Y. Xu, On discrete orthogonal polynomials of several variables, Adv. in Appl. Math. 33 (2004) 615–632.
[126] Y. Xu, Monomial orthogonal polynomials of several variables, J. Approx. Theory 133 (2005) 1–37.

[127] G.J. Yoon, Darboux transforms and orthogonal polynomials, Bull. Korean Math. Soc. 39 (2002) 359–376.
[128] F. Zhang (Ed.), The Schur Complement and Its Applications, Springer, New York, 2005.
[129] X.-M. Zhu, D.-J. Zhang, C.-X. Li, Quasideterminant solutions of a noncommutative nonisospectral Kadomtsev–Petviashvili equation, Commun. Theor. Phys. 55 (2011) 753–759.

3 A Jacobi type Christoffel-Darboux formula for multiple orthogonal polynomials of mixed type

Linear Algebra and its Applications 468 (2015) 154–170

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

A Jacobi type Christoffel–Darboux formula for multiple orthogonal polynomials of mixed type

Gerardo Araznibarreta, Manuel Mañas *

Departamento de Física Teórica II (Métodos Matemáticos de la Física), Universidad Complutense de Madrid, 28040, Madrid, Spain

A R T I C L E I N F O

Article history:
Received 27 October 2013
Accepted 29 April 2014
Available online 28 May 2014
Submitted by P. Psarrakos

MSC:
33C45
42C05
15A23
37K10

Keywords:
Multiple orthogonal polynomials
Christoffel–Darboux formula
Moment matrices
Jacobi type matrices
Gauss decomposition

A B S T R A C T

An alternative expression for the Christoffel–Darboux formula for multiple orthogonal polynomials of mixed type is derived from the *LU* factorization of the moment matrix of a given measure and two sets of weights. We use the action of the generalized Jacobi matrix J, also responsible for the recurrence relations, on the linear forms and their duals to obtain the result.

1. Introduction

In this paper we address a natural question that arises from the *LU* factorization approach to multiple orthogonality [7]. The Gauss decomposition of a Hankel matrix, which plays the role of a moment matrix, leads in the classical case to a natural description

* Corresponding author.
E-mail addresses: gariznab@ucm.es (G. Araznibarreta), manuel.manas@ucm.es (M. Mañas).

http://dx.doi.org/10.1016/j.laa.2014.04.030

of algebraic facts regarding orthogonal polynomials on the real line (OPRL) such as recursion relations and Christoffel–Darboux (CD) formula. In that case we have a chain of orthogonal polynomials $\{P_l(x)\}_{l=0}^{\infty}$ of increasing degree l. In [7] we extended that approach to the multiple orthogonality scenario, and the Gauss decomposition of an appropriate moment matrix led to sequences of families of multiple orthogonal polynomials in the real line (MOPRL), $\{Q^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}\}_{l=0}^{\infty}$ and $\{\bar{Q}^{(\mathrm{I},a_1(l))}_{[\vec{\nu}_2(l);\vec{\nu}_1(l-1)]}\}_{l=0}^{\infty}$. These families happen to be biorthogonal, and therefore we will refer to them as biorthogonal sequences of linear forms. The recursion formulae are relations constructed in terms of the linear forms in these sequences. However, the Daems–Kuijlaars Christoffel–Darboux formula given in Proposition 4 – that was re-deduced in [7] by linear algebraic means (Gauss decomposition) and the use of the ABC theorem – was not expressed in terms of linear forms belonging to the mentioned sequences. This situation is rather different to the OPRL case, in that standard scenario of the CD formula, call it the ABC type CD formula, is expressed in terms of orthogonal polynomials in the sequence. The aim of this paper is to show that, within that scheme, we can deduce an alternative but equivalent MOPRL Christoffel–Darboux formula constructed in terms of linear forms in the sequences $\{Q^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}\}_{l=0}^{\infty}$ and $\{\bar{Q}^{(\mathrm{I},a_1(l))}_{[\vec{\nu}_2(l);\vec{\nu}_1(l-1)]}\}_{l=0}^{\infty}$ as in OPRL situation. Besides we are able to find an OPRL type CD formula, expressed in terms solely of elements in the biorthogonal sequences of MOP of mixed type, there are two prices to pay: first, we need, in general, more terms than in the ABC type CD formula for these MOPs and second, we will need to know the coefficients in the recursion relation; i.e., the Jacobi coefficients. We will refer to these type of CD formulae as Jacobi type CD formulae as they are based on the structure of the Jacobi type matrix associated with the biorthogonal sequences which gives their recursion relations.

We must stress that in the OPRL scenario there are many ways to prove the CD formula [26]. In particular, on the one hand we could prove it using the ABC theorem combined with the moment matrix symmetry and on the other hand using the eigen-value properties of the Jacobi matrix. These two approaches – ABC and Jacobi – lead, in this simple situation, to the same result. However, as already mentioned, in the MOP scenario the two approaches lead to different results: the ABC type CD formula (or Daems–Kuijlaars CD formula) and the Jacobi type CD formula.

1.1. Historical background

Simultaneous rational approximation starts back in 1873 when Hermite proved the transcendence of the Euler number in [21]. Later, K. Mahler delivered at the University of Groningen several lectures [24] where he settled down the foundations of this theory, see also [13] and [22]. Simultaneous rational approximation when expressed in terms of Cauchy transforms leads to multiple orthogonality of polynomials. Given an interval $\Delta \subset \mathbb{R}$ of the real line, let $\mathcal{M}(\Delta)$ denote all the finite positive Borel measures with support containing infinitely many points in Δ. Fix $\mu \in \mathcal{M}(\Delta)$, and let us consider a system of weights $\vec{w} = (w_1, \dots, w_p)$ on Δ, with $p \in \mathbb{N}$; i.e. $w_1, \dots, w_p$ being real integrable functions on Δ which does not change sign on Δ. Fix a multi-index $\vec{\nu} = (\nu_1, \dots, \nu_p) \in \mathbb{Z}_+^p$, $\mathbb{Z}_+ = \{0, 1, 2, \dots\}$, and denote $|\vec{\nu}| = \nu_1 + \dots + \nu_p$. Then, there exist polynomials, $A_1, \dots, A_p$ not all identically equal to zero which satisfy the following orthogonality relations

$$\int_{\Delta} x^j \sum_{a=1}^{p} A_a(x) w_a(x) \mathrm{d}\mu(x) = 0, \qquad \deg A_a \leq \nu_a - 1, \quad j = 0, \dots, |\vec{\nu}| - 2. \tag{1}$$

Analogously, there exists a polynomial B not identically equal to zero, such that

$$\int_{\Delta} x^j B(x) w_b(x) \mathrm{d}\mu(x) = 0, \qquad \deg B \leq |\vec{\nu}|, \quad j = 0, \dots, \nu_b - 1, \; b = 1, \dots, p. \tag{2}$$

These are the so called multiple orthogonal polynomials of type I and type II, respectively, with respect to the combination $(\mu, \vec{w}, \vec{\nu})$ of the measure μ, the systems of weights $\vec{w}$ and the multi-index $\vec{\nu}$. When $p = 1$ both definitions coincide with standard orthogonal polynomials on the real line. Given a measure $\mu \in \mathcal{M}(\Delta)$ and a system of weights $\vec{w}$ on Δ a multi-index $\vec{\nu}$ is called type I or type II normal if $\deg A_a$ must equal to $\nu_a - 1$, $a = 1, \dots, p$, or $\deg B$ must equal to $|\vec{\nu}| - 1$, respectively. When for a pair $(\mu, \vec{w})$ all the multi-indices are type I or type II normal, then the pair is called type I perfect or type II perfect respectively. Multiple orthogonal of polynomials have been employed in several proofs of irrationality of numbers. For example, in [11] F. Beukers shows that Apery's proof [10] of the irrationality of $\zeta(3)$ can be placed in the context of a combination of type I and type II multiple orthogonality which is called mixed type multiple orthogonality of polynomials. More recently, mixed type approximation has appeared in random matrix and non-intersecting Brownian motion theories, [12,15,23]. Sorokin [27] studied a simultaneous rational approximation construction which is closely connected with multiple orthogonal polynomials of mixed type. In [29] a Riemann–Hilbert problem was found that characterizes multiple orthogonality of type I and II, extending in this way the result previously found in [20] for standard orthogonality. In [15] mixed type multiple orthogonality was analyzed from this perspective. For a general study, but not including multiple orthogonality, of Christoffel–Darboux kernels see [26]. In [9] we gave a generalization of CD formulae to matrix generalized orthogonal polynomials. In [16] MOPRL and some CD kernels are used in the study of average characteristic polynomials and in [17] some properties of models of n one-dimensional, nonintersecting Brownian motions with two prescribed starting points at time $t = 0$ and two prescribed ending points at time $t = 1$ in a critical regime are analyzed with the aid of Hermite MOP. Finally, in [19] a large class of MOPRL are shown to fulfill that its zeros on the real line are simple, lie in the interior of the convex hull of the support of the measure and the zeros of consecutive orthogonal polynomials interlace.

1.2. Perfect systems and MOPRL of mixed type

In order to introduce multiple orthogonal polynomials of mixed type we consider two systems of weights $\vec{w}_1 = (w_{1,1}, \dots, w_{1,p_1})$ and $\vec{w}_2 = (w_{2,1}, \dots, w_{2,p_2})$ where $p_1, p_2 \in \mathbb{N}$, and two multi-indices $\vec{\nu}_1 = (\nu_{1,1}, \dots, \nu_{1,p_1}) \in \mathbb{Z}_+^{p_1}$ and $\vec{\nu}_2 = (\nu_{2,1}, \dots, \nu_{2,p_2}) \in \mathbb{Z}_+^{p_2}$ with $|\vec{\nu}_1| = |\vec{\nu}_2| + 1$. There exist polynomials $A_1, \dots, A_{p_1}$, not all identically zero, such that $\deg A_s < \nu_{1,s}$, which satisfy the following relations

$$\int\limits_{\Delta} \sum_{a=1}^{p_1} A_a(x) w_{1,a}(x) w_{2,b}(x) x^j d\mu(x) = 0, \quad j = 0, \dots, \nu_{2,b} - 1,\ b = 1, \dots, p_2. \tag{3}$$

In this paper we say that we have p_1 components of type II and p_2 components of type I. They are called mixed multiple-orthogonal polynomials with respect to the combination $(\mu, \vec{w}_1, \vec{w}_2, \vec{\nu}_1, \vec{\nu}_2)$ of the measure μ, the systems of weights $\vec{w}_1$ and $\vec{w}_2$ and the multi-indices $\vec{\nu}_1$ and $\vec{\nu}_2$. It is easy to show that finding the polynomials $A_1, \dots, A_{p_1}$ is equivalent to solving a system of $|\vec{\nu}_2|$ homogeneous linear equations for the $|\vec{\nu}_1|$ unknown coefficients of the polynomials. Since $|\vec{\nu}_1| = |\vec{\nu}_2| + 1$ the system always has a nontrivial solution. The matrix of this system of equations is the so called moment matrix, and the study of its Gauss decomposition will be the cornerstone of this paper. Observe that when $p_1 = 1$ we are in the type II case and if $p_2 = 1$ in type I case. Hence in general we can find a solution of (3) where there is an $a \in \{1, \dots, p_1\}$ such that $\deg A_a < \nu_{1,a} - 1$. When given a combination $(\mu, \vec{w}_1, \vec{w}_2)$ of a measure $\mu \in \mathcal{M}(\Delta)$ and systems of weights $\vec{w}_1$ and $\vec{w}_2$ on Δ if for each pair of multi-indices $(\vec{\nu}_1, \vec{\nu}_2)$ the conditions (3) determine that $\deg A_a = \nu_{1,a} - 1$, $a = 1, \dots, p_1$, then we say that the combination $(\mu, \vec{w}_1, \vec{w}_2)$ is perfect. In this case we can determine a unique system of mixed type orthogonal polynomials $(A_1, \dots, A_{p_1})$ satisfying (3) requiring for $a_1 \in \{1, \dots, p_1\}$ that A_{a_1} monic. Following [15] we say that we have a type II normalization and denote the corresponding system of polynomials by $A_a^{(\mathrm{II},a_1)}$, $j = 1, \dots, p_1$. Alternatively, we can proceed as follows, since the system of weights is perfect from (3) we deduce that

$$\int x^{\nu_{1,r_1}} \sum_{a=1}^{p_1} A_a(x) w_{1,a}(x) w_{2,b}(x) \mathrm{d}\mu(x) \neq 0.$$

Then, we can determine a unique system of mixed type of multi-orthogonal polynomials $(A_1^{(\mathrm{I},a_2)}, \dots, A_{p_2}^{(\mathrm{I},a_2)})$ imposing that

$$\int x^{\nu_{1,a_2}} \sum_{a=1}^{p_1} A_a^{(\mathrm{I},a_2)}(x) w_{1,a}(x) w_{2,b}(x) \mathrm{d}\mu(x) = 1,$$

which is a type I normalization. We will use the notation $A^{(\mathrm{II},a_1)}_{[\vec{\nu}_1;\vec{\nu}_2],a}$ and $A^{(\mathrm{I},a_2)}_{[\vec{\nu}_1;\vec{\nu}_2],a}$ to denote these multiple orthogonal polynomials with type II and I normalizations, respectively.

A known illustration of perfect combinations $(\mu, \vec{w}_1, \vec{w}_2)$ can be constructed with an arbitrary positive finite Borel measure μ and systems of weights formed with exponentials:

$$\left(\mathrm{e}^{\gamma_1 x}, \dots, \mathrm{e}^{\gamma_p x}\right), \quad \gamma_i \neq \gamma_j,\ i \neq j,\ i, j = 1, \dots, p, \tag{4}$$

or by binomial functions

$$\left((1-z)^{\alpha_1}, \dots, (1-z)^{\alpha_p}\right), \quad \alpha_i - \alpha_j \notin \mathbb{Z},\ i \neq j,\ i, j = 1, \dots, p \tag{5}$$

or combining both classes, see [25]. Recently, in [18] the authors were able to prove perfectness for a wide class of systems of weights. These systems of functions, now called Nikishin systems, were introduced by E.M. Nikishin [25] and initially named MT-systems (after Markov and Tchebycheff).

1.3. Gauss decomposition and multiple orthogonality of mixed type. A reminder

Orthogonal polynomials and the theory of integrable systems have been connected in several ways in the mathematical literature. We are particularly interested in the one based in the Gauss decomposition that was developed in [1–5], and applied further in [6–8]. These papers set the basis for the method we use in this paper to get an alternative CD formula for MOPRL of mixed type.

In the following we extract from [7] the necessary material for the construction of the mentioned alternative Christoffel–Darboux formula. We introduce the moment matrix and recall how the Gauss decomposition leads to multiple orthogonality. Then, we outline how the recursion relations appears by introducing a Jacobi type semi-infinite matrix and recall the reader the CD formula [14,15].

1.3.1. The moment matrix

We now proceed to define the moment matrix. For that aim we need as starting point two systems of weights $\vec{w}_\alpha = (w_{\alpha,1}, \dots, w_{\alpha,p_\alpha})$, $\alpha = 1, 2$ and $p_1, p_2 \in \{1, 2, 3, \dots\}$ and a finite Borel measure $\mathrm{d}\mu$ supported all of them on an interval $\Delta \subset \mathbb{R}$. Given two compositions[1] $\vec{n}_\alpha = (n_{\alpha,1}, \dots, n_{\alpha,p_\alpha})$, $\alpha = 1, 2$, of $|\vec{n}_\alpha| = n_{\alpha,1} + \dots + n_{\alpha,p_\alpha}$ any given $l \in \mathbb{Z}_+ := \{0, 1, 2, \dots\}$ determines uniquely, through Euclidean division, the following non-negative integers $k_\alpha(l) \in \mathbb{Z}_+$, $a_\alpha(l) \in \{1, 2, \dots, p_\alpha\}$ and $r_\alpha(l)$ such that $r_\alpha(l) \in \{0, 1, \dots, n_{\alpha,a_\alpha(l)} - 1\}$ and

[1] Do not confuse with a partition; in Combinatorics, see for example [28], a composition of an integer n is a way of writing n as the sum of a sequence of (strictly) positive integers. Two sequences that differ in the order of their terms define different compositions of their sum, while they are considered to define the same partition of that number. Every integer has finitely many distinct compositions. Given that for the Gauss decomposition description of MOP this order is relevant we have stressed this aspect and preferred the name of composition to that of multi-index, which can be also used.

$$l=\begin{cases}k_\alpha(l)|\vec{n}_\alpha|+r_\alpha(l), & a_\alpha(l)=1,\\ k_\alpha(l)|\vec{n}_\alpha|+n_{\alpha,1}+\cdots+n_{\alpha,a_\alpha(l)-1}+r_\alpha(l), & a_\alpha(l)\neq 1.\end{cases}\tag{6}$$

We define two monomial vectors that may be understood as sequences of monomials according to the composition $\vec{n}_\alpha$, $\alpha=1,2$, introduced previously.

$$\chi_\alpha:=\begin{pmatrix}\chi_{\alpha,[0]}\\ \chi_{\alpha,[1]}\\ \vdots\\ \chi_{\alpha,[k]}\\ \vdots\end{pmatrix}\quad\text{where } \chi_{\alpha,[k]}:=\begin{pmatrix}\chi_{\alpha,[k],1}\\ \chi_{\alpha,[k],2}\\ \vdots\\ \chi_{\alpha,[k],a_\alpha}\\ \vdots\\ \chi_{\alpha,[k],p_\alpha}\end{pmatrix}\text{ and } \chi_{\alpha,[k],a_\alpha}:=\begin{pmatrix}x^{kn_{\alpha,a_\alpha}}\\ x^{kn_{\alpha,a_\alpha}+1}\\ \vdots\\ x^{kn_{\alpha,a_\alpha}+(n_{\alpha,a_\alpha}-1)}\end{pmatrix}.$$

In a similar manner for $\alpha=1,2$ we define the weighted monomial vectors

$$\xi_\alpha:=\begin{pmatrix}\xi_{\alpha,[0]}\\ \xi_{\alpha,[1]}\\ \vdots\\ \xi_{\alpha,[k]}\\ \vdots\end{pmatrix}\quad\text{where } \xi_{\alpha,[k]}:=\begin{pmatrix}w_{\alpha,1}\chi_{\alpha,[k],1}\\ w_{\alpha,2}\chi_{\alpha,[k],2}\\ \vdots\\ w_{\alpha,p_\alpha}\chi_{\alpha,[k],p_\alpha}\end{pmatrix}.$$

For example, let us put $p_1=2$ and $p_2=1$ and set the compositions $n_{1,1}=3$, $n_{1,2}=2$ and $n_{2,1}=1$ with weight vectors $\vec{w}_1=(w_{1,1},w_{1,2})$ and $\vec{w}_2=(w_{2,1})$. Then

$$\begin{aligned}\chi_1&=\big(1,x,x^2,1,x,x^3,x^4,x^5,x^2,x^3,x^6,x^7,x^8,x^4,x^5,\dots\big)^\top,\\ \chi_2&=\big(1,x,x^2,\dots\big)^\top,\\ \xi_1&=\big(w_{1,1},w_{1,1}x,w_{1,1}x^2,w_{1,2},w_{1,2}x,w_{1,1}x^3,w_{1,1}x^4,w_{1,1}x^5,w_{1,2}x^2,w_{1,2}x^3,\dots\big)^\top,\\ \xi_2&=w_{2,1}\big(1,x,x^2,\dots\big)^\top.\end{aligned}$$

We have used two colors, red (light gray in print) and blue (dark gray in print), for $\alpha=1$, to remark the two ($p_1=2$) forms of growth, in steps of 3 for the monomial powers in red component and of 2 for the blue one. For the corresponding MOP of mixed type (in this case are just of type II as we have choose $p_2=1$) these colors are associated, as we will see, to the two components of type II that this example leads to the red and blue components. Observe that for the construction of the monomial vectors χ_α and weighted monomial vectors ξ_α, $\alpha=1,2$, only the two compositions $\vec{n}_\alpha$, $\alpha=1,2$, are needed. However, the l-th entries or coefficients, $\chi_\alpha^{(l)}$ and $\xi_\alpha^{(l)}$, of these semi-infinite vectors can be explicitly expressed in terms of the just introduced Euclidean division

$$\chi_\alpha^{(l)}=x^{\nu_{\alpha,a_\alpha(l)}},\qquad \xi_\alpha^{(l)}=w_{\alpha,a_\alpha(l)}x^{\nu_{\alpha,a_\alpha(l)}}$$

where for any given $l\in\mathbb{Z}_+$ and $a_\alpha:=1,2,\dots,p_\alpha$ we define

$$\nu_{\alpha,a_\alpha}(l) := \begin{cases} k_\alpha(l)|\vec{n}_\alpha| + n_{\alpha,a_\alpha} - 1, & a_\alpha < a_\alpha(l), \\ k_\alpha(l)|\vec{n}_\alpha| + r_\alpha(l), & a_\alpha = a_\alpha(l), \\ k_\alpha(l)|\vec{n}_\alpha| - 1, & a_\alpha > a_\alpha(l). \end{cases}$$

Notice that $\nu_{\alpha,a_\alpha}(l)$ is the highest degree of all the monomials of type a_α up to the component $\chi_\alpha^{(l)}$ included, of the monomial vector.

We stress that for a given positive integer l the number $a_\alpha(l) \in \{1,\dots,p_\alpha\}$ distinguishes to which of the p_α possible components, or different colors (light and dark gray in print) in the previous example (of type II for $\alpha = 1$ and type I for $\alpha = 2$) this integer belongs to. Later on for any positive integer l we will need to know which is the closest integer by defect or by excess in a given component $a_\alpha \in \{1,\dots,p_\alpha\}$, $\alpha \in \{1,2\}$. For that aim we introduce the functions

$$\begin{aligned} [\cdot,\cdot]_\alpha^{\lessgtr} : \mathbb{Z}_+ \times \{1,\dots,p_\alpha\} &\to \mathbb{Z}_+, \\ (l,a) &\mapsto [l,a]_\alpha^{\lessgtr}, \end{aligned}$$

where

$$\begin{aligned} [l,a]_\alpha^{<} &:= \begin{cases} k_\alpha(l)|\vec{n}_\alpha| + \sum_{i=1}^{a} n_{\alpha,i} - 1, & a < a_\alpha(l), \\ l, & a = a_\alpha(l), \\ k_\alpha(l)|\vec{n}_\alpha| - \sum_{i=a+2}^{p_\alpha} n_{\alpha,i} - 1, & a > a_\alpha(l), \end{cases} \\ [l,a]_\alpha^{>} &:= \begin{cases} (k_\alpha(l)+1)|\vec{n}_\alpha| + \sum_{i=1}^{a-1} n_{\alpha,i}, & a < a_\alpha(l), \\ l, & a = a_\alpha(l), \\ (k_\alpha(l)+1)|\vec{n}_\alpha| - \sum_{i=a+1}^{p_\alpha} n_{\alpha,i}, & a > a_\alpha(l). \end{cases} \end{aligned} \tag{7}$$

It can be proven that these are the desired integers; i.e., that $[l,a]_\alpha^{>}$ ($[l,a]_\alpha^{<}$) is the smallest (largest) integer such that $[l,a]_\alpha^{>} \geq l$ ($[l,a]_\alpha^{<} \leq l$) and $a_\alpha([l,a]_\alpha^{>}) = a$ ($a_\alpha([l,a]_\alpha^{<}) = a$).

Finally, given the weighted monomials ξ_α, associated to the compositions $\vec{n}_\alpha$, $\alpha = 1,2$, we introduce the moment matrix in the following manner

Definition 1. The moment matrix is given by

$$g := \int \xi_1(x)\xi_2(x)^\top \mathrm{d}\mu(x). \tag{8}$$

1.3.2. Multiple orthogonality of mixed type: The Gauss decomposition of the moment matrix

Definition 2. For a given perfect combination $(\mu, \vec{w}_1, \vec{w}_2)$ we define

(1) The Gauss decomposition (also known as *LU* factorization) of a semi-infinite moment matrix g, determined by $(\mu, \vec{w}_1, \vec{w}_2)$, is the problem of finding the solution of

$$g = S^{-1}\bar{S}, \qquad S = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ S_{1,0} & 1 & 0 & \cdots \\ S_{2,0} & S_{2,1} & 1 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

$$\bar{S}^{-1} = \begin{pmatrix} \bar{S}'_{0,0} & \bar{S}'_{0,1} & \bar{S}'_{0,2} & \cdots \\ 0 & \bar{S}'_{1,1} & \bar{S}'_{1,2} & \cdots \\ 0 & 0 & \bar{S}'_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{9}$$

where $S_{i,j}, \bar{S}'_{i,j} \in \mathbb{R}$.

(2) In terms of these matrices we construct the polynomials

$$A_a^{(l)} := \sum_i{}' S_{l,i} x^{k_1(i)}, \tag{10}$$

where the sum $\sum'$ is taken for a fixed $a = 1, \ldots, p_1$ over those i such that $a = a_1(i)$ and $i \leq l$. We also construct the dual polynomials

$$\bar{A}_b^{(l)} := \sum_j{}' x^{k_2(j)} \bar{S}'_{j,l}, \tag{11}$$

where the sum $\sum'$ is taken for a given b over those j such that $b = a_2(j)$ and $j \leq l$.

(3) Vectors of linear forms and dual linear forms associated with multiple orthogonal polynomials and their duals are defined by

$$Q := \begin{pmatrix} Q^{(0)} \\ Q^{(1)} \\ \vdots \end{pmatrix} = S\xi_1, \qquad \bar{Q} := \begin{pmatrix} \bar{Q}^{(0)} \\ \bar{Q}^{(1)} \\ \vdots \end{pmatrix} = \left(\bar{S}^{-1}\right)^\top \xi_2. \tag{12}$$

Then – see Propositions 3, 4, 5 and 6 in [7] –

Proposition 1.

(1) *The linear forms and their duals, introduced in Definition 2, are given by*

$$Q^{(l)}(x) := \sum_{a=1}^{p_1} A_a^{(l)}(x) w_{1,a}(x), \qquad \bar{Q}^{(l)}(x) := \sum_{b=1}^{p_2} \bar{A}_b^{(l)}(x) w_{2,b}(x). \tag{13}$$

(2) *The orthogonality relations*

$$\int Q^{(l)}(x) w_{2,b}(x) x^k \mathrm{d}\mu(x) = 0, \quad 0 \leq k \leq \nu_{2,b}(l-1) - 1, \; b = 1, \ldots, p_2,$$

$$\int \bar{Q}^{(l)}(x)w_{1,a}(x)x^k\mathrm{d}\mu(x) = 0, \quad 0 \leq k \leq \nu_{1,a}(l-1)-1, \; a = 1,\dots,p_1, \tag{14}$$

are fulfilled.

(3) *We have the following identifications*

$$A_a^{(l)} = A^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)],a}, \qquad \bar{A}_b^{(l)} = A^{(\mathrm{I},a_1(l))}_{[\vec{\nu}_2(l);\vec{\nu}_1(l-1)],b},$$

in terms of multiple orthogonal polynomials of mixed type with two normalizations I *and* II, *respectively.*

(4) *The following multiple bi-orthogonality relations among linear forms and their duals*

$$\int Q^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}(x)\bar{Q}^{(\mathrm{I},a_1(k))}_{[\vec{\nu}_2(k);\vec{\nu}_1(k-1)]}(x)\mathrm{d}\mu(x) = \delta_{l,k}, \quad l,k \in \mathbb{Z}_+, \tag{15}$$

hold.

Observe that a major difference between the usual approach to MOPRL of mixed type, in which the orthogonality relations (3) are discussed in its own, and the described Gauss decomposition approach is precisely the biorthogonality conditions given by (15). While for standard OPRL both type orthogonal relations – the perpendicularity of each polynomial P_l to $\{1, x, \dots, x^{l-1}\}$ and the orthogonality of the set of polynomials $\{P_l\}_{l=0}^{\infty}$ – are discussed in equal footing this has no parallel before in the MOPRL scenario. Biorthogonality (15) gives such a bridge; i.e., we have two sequences of MORPL – with normalizations of types I and II, respectively – such that its biorthogonality is equivalent to the multiple orthogonality condition of both families.

1.3.3. Jacobi type matrices and recursion relations

The moment matrix has a Hankel type symmetry that implies the recursion relations and the Christoffel–Darboux formula. We consider the shift operators Υ_α defined by

$$(\Upsilon_\alpha)_{l,j} := \delta_{j,[l+1,a_\alpha(l)]^>} \tag{16}$$

which satisfy the following relation

$$\Upsilon_\alpha\chi_\alpha(x) = x\chi_\alpha(x) \quad \Longrightarrow \quad \Upsilon_\alpha\xi_\alpha(x) = x\xi_\alpha(x)$$

In terms of these shift matrices we can describe the particular Hankel symmetries for the moment (see Proposition 12 in [7]) matrix

Proposition 2. *The moment matrix g satisfies the Hankel type symmetry*

$$\Upsilon_1 g = g\Upsilon_2^\top. \tag{17}$$

From this symmetry we see that the following is consistent

Definition 3. We define the matrices

$$J := S\Upsilon_1 S^{-1} = \bar{S}\Upsilon_2^\top \bar{S}^{-1} = J_+ + J_-, \qquad J_+ := \left(S\Upsilon_1 S^{-1}\right)_+, \qquad J_- := \left(\bar{S}\Upsilon_2^\top \bar{S}^{-1}\right)_-, \tag{18}$$

where the sub-indices + and − denote the upper triangular and strictly lower triangular projections.

The matrix J for this MORPL of mixed type is therefore, not a tridiagonal matrix as for the standard OPRL, but more generally a banded matrix with the number of upper and lower diagonal determined by the number of components and compositions.

The recursion relations follow immediately from the eigenvalue property

$$JQ(x) = xQ(x), \qquad \bar{Q}(x)^\top J = x\bar{Q}(x)^\top, \tag{19}$$

which imply for $\{Q^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}(x)\}_{l=0}^\infty$ and $\{\bar{Q}^{(\mathrm{I},a_1(k))}_{[\vec{\nu}_2(k);\vec{\nu}_1(k-1)]}(x)\}_{k=0}^\infty$ recursion relations; i.e., each $xQ^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}(x)$ is expressed as a finite sum of linear forms in $\{Q^{(\mathrm{II},a_1(i))}_{[\vec{\nu}_1(i);\vec{\nu}_2(i-1)]}(x)\}_{i=0}^\infty$ and each $x\bar{Q}^{(\mathrm{I},a_1(k))}_{[\vec{\nu}_2(k);\vec{\nu}_1(k-1)]}(x)$ as a finite combination of dual linear forms in $\{\bar{Q}^{(\mathrm{I},a_1(j))}_{[\vec{\nu}_2(j);\vec{\nu}_1(j-1)]}(x)\}_{j=0}^\infty$.

1.3.4. The ABC type Christoffel–Darboux formula for MOP of mixed type

Definition 4. The l-th Christoffel–Darboux kernel is defined by

$$K^{[l]}(x,y) := \sum_{k=0}^{l-1} Q^{(\mathrm{II},a_1(k))}_{[\vec{\nu}_1(k);\vec{\nu}_2(k-1)]}(y)\bar{Q}^{(\mathrm{I},a_1(k))}_{[\vec{\nu}_2(k);\vec{\nu}_1(k-1)]}(x). \tag{20}$$

Any semi-infinite vector v can be written in block form as follows

$$v = \left(\begin{array}{c} v^{[l]} \\ \hline v^{[\geq l]} \end{array}\right)$$

$v^{[l]}$ is the finite vector formed with the first l coefficients of v and $v^{[\geq l]}$ the semi-infinite vector formed with the remaining coefficients. This decomposition induces the following block structure for any semi-infinite matrix

$$M = \left(\begin{array}{c|c} M^{[l]} & M^{[l,\geq l]} \\ \hline M^{[\geq l,l]} & M^{[\geq l]} \end{array}\right).$$

In Corollary 2 in [7] we found an ABC (Aitken–Berg–Collar) type theorem – this denomination is the one that appears in [26] for the OPRL case –

Proposition 3. *The Christoffel–Darboux kernel can be expressed in terms of the inverse of the truncated moment matrix as follows*

$$K^{[l]}(x,y) = \big(\xi_2^{[l]}(x)\big)^\top \big(g^{[l]}\big)^{-1} \xi_1^{[l]}(y). \tag{21}$$

Finally what we call the ABC type or Kuijlaars–Daems CD formula for MOP of mixed type is (see Proposition 21 in [7])

Proposition 4. *For $l \geq \max(|\vec{n}_1|, |\vec{n}_2|)$ the following*

$$\begin{aligned}(x-y)K^{[l]}(x,y) = \sum_{b=1}^{p_2} \bar{Q}^{(\mathrm{II},b)}_{[\vec{\nu}_2(l-1)+\vec{e}_{2,b};\vec{\nu}_1(l-1)]}(x) Q^{(\mathrm{I},b)}_{[\vec{\nu}_1(l-1);\vec{\nu}_2(l-1)-\vec{e}_{2,b}]}(y) \\ - \sum_{a=1}^{p_1} \bar{Q}^{(\mathrm{I},a)}_{[\vec{\nu}_2(l-1);\vec{\nu}_1(l-1)-\vec{e}_{1,a}]}(x) Q^{(\mathrm{II},a)}_{[\vec{\nu}_1(l-1)+\vec{e}_{1,a};\vec{\nu}_2(l-1)]}(y)\end{aligned} \tag{22}$$

holds.

This ABC type CD formula for MOP of mixed type has been proven by Kuijlaars and Daems using a Riemann–Hilbert problem, see [14,15]. Later on, in [7] it was proven for the first time by algebraic means – not relying on analytic conditions as in [14,15] – using the ABC theorem (21) and the symmetry (17). Here $\{\vec{e}_{i,a}\}_{a=1}^{p_i} \subset \mathbb{R}^{p_i}$ stands for the vectors in the respective canonical basis, $i = 1, 2$. We stress the appearance of $\vec{\nu}_2(l-1) + \vec{e}_{2,b}$, $\vec{\nu}_2(l-1) - \vec{e}_{2,b}$, $\vec{\nu}_1(l-1) - \vec{e}_{1,a}$ and $\vec{\nu}_1(l-1) + \vec{e}_{1,a}$ which are multi-indexes that do not belong to the multi-index sequence associated with the sequence of biorthogonal linear forms $\{Q^{(\mathrm{II},a_1(l))}_{[\vec{\nu}_1(l);\vec{\nu}_2(l-1)]}(x)\}_{l=0}^\infty$ and $\{\bar{Q}^{(\mathrm{I},a_1(k))}_{[\vec{\nu}_2(k);\vec{\nu}_1(k-1)]}(x)\}_{k=0}^\infty$. Our alternative proposal, despite of having a larger number of terms, as we will see below, involves only linear forms in the sequence.

2. Jacobi type Christoffel–Darboux formula for multiple orthogonal polynomials of mixed type

Given any positive integer $l \in \mathbb{Z}_+$ we consider the arithmetic congruence modulo p_α; i.e.

$$l = \bar{l}_\alpha \mod p_\alpha, \qquad \bar{l}_\alpha \in \{0, 1, \dots, p_\alpha - 1\} \cong \mathbb{Z}_{p_\alpha} = \mathbb{Z}/(p_\alpha \mathbb{Z}), \quad \alpha = 1, 2.$$

The result of this paper is the following

Theorem 1. *For $l \geq \max\{|\vec{n}_1|, |\vec{n}_2|\}$ the following Jacobi type Christoffel–Darboux formula holds*

$$(y-x)K^{[l]}(x,y) = \sum_{(i,j)\in\sigma_1[l]} \bar{Q}^{(\mathrm{I},a_1(j))}_{[\vec{\nu}_2(j);\vec{\nu}_1(j-1)]}(x)J_{j,i}Q^{(\mathrm{II},a_1(i))}_{[\vec{\nu}_1(i);\vec{\nu}_2(i-1)]}(y)$$
$$- \sum_{(i,j)\in\sigma_2[l]} \bar{Q}^{(\mathrm{I},a_1(j))}_{[\vec{\nu}_2(j);\vec{\nu}_1(j-1)]}(x)J_{j,i}Q^{(\mathrm{II},a_1(i))}_{[\vec{\nu}_1(i);\vec{\nu}_2(i-1)]}(y),$$

where

$$\sigma_1[l] := \left\{l,\dots,\left[l,\overline{(a_1(l)-1)}_1\right]^{>}_{1}\right\} \times \left\{\left[l-1,\dots,\overline{(a_1(l-1)+1)}_1\right]^{<}_{1}\right\},$$
$$\sigma_2[l] := \left\{\left[l-1,\overline{(a_2(l-1)+1)}_2\right]^{<}_{2},\dots,l-1\right\} \times \left\{l,\dots,\left[l,\overline{(a_2(l)-1)}_2\right]^{>}_{2}\right\}.$$

Proof. Splitting the eigenvalue property (19) into blocks we get

$$JQ(y) = yQ(y) \quad\Longrightarrow\quad J^{[l]}Q(y)^{[l]} + J^{[l,\geq l]}Q(y)^{[\geq l]} = yQ(y)^{[l]}$$
$$\bar{Q}(x)^\top J = x\bar{Q}(x)^\top \quad\Longrightarrow\quad \left[\bar{Q}(x)^\top\right]^{[l]}J^{[l]} + \left[\bar{Q}(x)^\top\right]^{[\geq l]}J^{[\geq l,l]} = x\left[\bar{Q}(x)^\top\right]^{[l]}$$

Multiply the first equation from the left by $[\bar{Q}(x)^\top]^{[l]}$ and the second one from the right by $Q(y)^{[l]}$ substract both results to obtain

$$\left[\bar{Q}(x)^\top\right]^{[l]}J^{[l,\geq l]}Q(y)^{[\geq l]} - \left[\bar{Q}(x)^\top\right]^{[\geq l]}J^{[\geq l,l]}Q(y)^{[l]}$$
$$= (y-x)\left[\bar{Q}(x)^\top\right]^{[l]}\cdot Q(y)^{[l]} = (y-x)K^{[l]}(x,y)$$

A brief study of the shape of J shows that, even though $J^{[l,\geq l]}$ has semi-infinite length rows, most of its elements are 0. Actually it only contains a finite number of nonzero entries that concentrate in the lower left corner of itself. The same reasoning applies to $J^{[\geq l,l]}$. This matrix has semi-infinite length columns but again it only contains a finite number of nonzero terms concentrated in the upper right corner of itself. Of course the number of terms involved in this expression will depend on the value of l. To be more precise we proceed as follows.

After a study of the shape of J we can state

Lemma 1. *For $l \geq \max\{|\vec{n}_1|,|\vec{n}_2|\}$ the only nonzero elements of J along a given row or column are*

$$\begin{array}{ccccccccc}
 & & & & J_{[l-1,\overline{(a_1(l-1)+1)_1}]_1^<,l} & & & & \\
 & & & & * & & & & \\
 & & & & \vdots & & & & \\
 & & & & * & & & & \\
J_{l,[l-1,\overline{(a_2(l-1)+1)_2}]_2^<} & * & \cdots & * & J_{l,l} & * & \cdots & * & J_{l,[l+1,\overline{(a_1(l+1)-1)_2}]_1^>} \\
 & & & & * & & & & \\
 & & & & \vdots & & & & \\
 & & & & * & & & & \\
 & & & & J_{[l+1,\overline{(a_2(l+1)-1)_2}]_2^>,l} & & & &
\end{array}$$

Using this lemma we get the desired result and the proof is complete. □

Remarkably, this Jacobi type CD formula is expressed uniquely in terms of our sequences of biorthogonal linear forms $\{Q^{(\mathrm{II},a_1(l))}_{[\vec\nu_1(l);\vec\nu_2(l-1)]}(x)\}_{l=0}^\infty$ and $\{\bar Q^{(\mathrm{I},a_1(k))}_{[\vec\nu_2(k);\vec\nu_1(k-1)]}(x)\}_{k=0}^\infty$, and does not need of alien multi-indexes to it, as $\vec\nu_2(l-1)+\vec e_{2,b}$, $\vec\nu_2(l-1)-\vec e_{2,b}$, $\vec\nu_1(l-1)-\vec e_{1,a}$ and $\vec\nu_1(l-1)+\vec e_{1,a}$ that appear in the standard CD formula for MOP of mixed type (22). The price we have to pay to have all the terms in the sequence of biorthogonal polynomials is that we will need more terms than in the formula (22) where we have (p_1+p_2) summands (where each summand is strange to the biorthogonal sequence of linear forms).

The number of terms N from the biorthogonal sequence that is needed in this Jacobi type CD formula can be expressed in terms of

$$n_\alpha(l) := l - \big[(l-1), \overline{\big(a_\alpha(l-1)+1\big)_\alpha}\big]_\alpha^<, \quad \alpha = 1, 2$$

as follows

Proposition 5. *We have the following equation*

$$N = \sum_{k=0}^{[l,\overline{(a_1(l)-1)_1}]_1^> - l} \big[n_1(l+k)-k\big] + \sum_{k=0}^{[l,\overline{(a_2(l)-1)_2}]_2^> - l} \big[n_2(l+k)-k\big].$$

The worst situation is reached when the compositions are $\vec n_1 = (1,\dots,1)$ and $\vec n_2 = (1,\dots,1)$; in this case we have that

$$N = \frac{|\vec n_1|(|\vec n_1|+1)}{2} + \frac{|\vec n_2|(|\vec n_2|+1)}{2}.$$

For any other pair of compositions we have less terms. In order to be more clear let us suppose that $p_1 = 3$ and $p_2 = 2$ with $\vec n_1 = (4,3,2)$ and $\vec n_2 = (3,2)$. The corresponding Jacobi type matrix has the following shape

$$
J=\left(\begin{array}{c|c} J^{[12]} & J^{[12,\geq 12]} \\ \hline J^{[\geq 12,12]} & J^{[\geq 12]} \end{array}\right)
$$

$$
=\left(\begin{array}{cccccccccccc|ccccccccccccccc}
\ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & \ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & 1 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \ast & \ast & \ast & \ast & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right),
\tag{23}
$$

where $\ast$ denotes a non-necessarily null real number.

In our example ($p_1 = 3$, $p_2 = 2$, $\vec{n}_1 = (4, 3, 2)$ and $\vec{n}_2 = (3, 2)$) for $l = 12$ we have

$$
(y - x)K^{[12]}(x, y) = \left[\sum_{i=6}^{11}\sum_{j=12}^{16} \bar{Q}(x)^{(i)} J_{i,j} Q(y)^{(j)}\right] - \left[\sum_{i=12}^{13}\sum_{j=9}^{11} \bar{Q}(x)^{(i)} J_{i,j} Q(y)^{(j)}\right].
$$

2.1. Expressing the Jacobi type matrix in terms of factorization factors

As we have seen we can write J in terms of S or of $\bar{S}$, this means that each term of J has two different expressions, giving relations between S with $\bar{S}$. We are not too concerned about these relations since what we want here is the most simple expression we can get for the elements of J. It is easy to realize that this is achieved if we use the expression involving S in order to calculate the upper part of J and the expression involving $\bar{S}$ to calculate the lower part of it. Hence, for every $J_{l,k}$ we will have expressions in terms of the factorization matrices coefficients and the elements of their inverses – thus, in terms of the MOPRL and associated second kind functions. The only terms from the factorization matrices (or their inverses) that will be involved when calculating any $J_{l,k}$ are just those between the main diagonal and the $l - |\vec{n}_1|$ diagonal (both included) of S and those between the main diagonal and the $l + |\vec{n}_2|$ diagonal (both included) of $\bar{S}$. And not even all of them. As we are about to see there are three different kinds of elements in J. The ones along the main diagonal, the ones along the immediate closest diagonals to the main one, and finally all the remaining diagonals. The recursion relation coefficients $J_{k,l}$ are ultimately related to the MOPRL and its associated second kind functions in the following way

Proposition 6. *The elements of the recursion matrix J can be written in terms of products of the entries of the LU factorization matrices and its inverses as follows*

$$\begin{aligned}
J_{l,l} &= S_{l,[(l-1),a(l)]_1^<} + S^{-1}_{[(l+1),a(l)]_1^>,l} + \sum_{\substack{a=1,\dots,p_1\\ a\neq a_1(l)}} S_{l,[(l-1),a]_1^<} S^{-1}_{[(l+1),a]_1^>,l}\\
&= \bar S_{l,[(l+1),a(l)]_2^>} \bar S^{-1}_{l,l} + \bar S_{l,l} \bar S^{-1}_{[(l-1),a(l)]_2^<,l} + \sum_{\substack{a=1,\dots,p_2\\ a\neq a_2(l)}} \bar S_{l,[(l+1),a]_2^>} \bar S^{-1}_{[(l-1),a]_2^<,l}\\
J_{l,l+1} &= S^{-1}_{[(l+1),a(l)]_1^>,l+1} + \sum_{\substack{a=1,\dots,p_1\\ a\neq a_1(l)}} S_{l,[(l-1),a]_1^<} S^{-1}_{[(l+1),a]_1^>,l+1}\\
J_{l+1,l} &= \bar S_{l+1,[(l+1),a(l)]_2^>} \bar S^{-1}_{l,l} + \sum_{\substack{a=1,\dots,p_2\\ a\neq a_2(l)}} \bar S_{l+1,[(l+1),a]_2^>} \bar S^{-1}_{[(l-1),a]_2^<,l}\\
J_{l,l+k} &= \sum_{a=\overline{(a_1(l+k-1)+1)_1}}^{\overline{(a_1(l)-1)_1}\,p_1} S_{l,[(l-1),a]_1^<} S^{-1}_{[(l+1),a]_1^>,l+k},\\
&\quad 2\leq k\leq \big[(l+1),\overline{(a_1(l+1)-1)}_1\big]_1^< - l,\\
J_{l+k,l} &= \sum_{a=\overline{(a_2(l+k-1)+1)_2}}^{\overline{(a_2(l)-1)_2}\,p_2} \bar S_{l+k,[(l+1),a]_2^>} \bar S^{-1}_{[(l-1),a]_2^<,l},\\
&\quad 2\leq k\leq \big[(l+1),\overline{(a_2(l+1)-1)}_2\big]_2^< - l.
\end{aligned}$$

Proof. To prove it we just take the definition (16) of Υ_α, $\alpha=1,2$, and the definition of J given in (18) to compute the different coefficients. $\square$

Where, for $r,r'<p_\alpha$, $\alpha\in\{1,2\}$, we have used

$$\sum_{a=r}^{r'\,p_\alpha} X_a = \begin{cases} \sum_{a=r}^{r'} X_a, & r\leq r',\\ \sum_{a=1}^{r'} X_a + \sum_{a=r}^{p_\alpha} X_a, & r>r'. \end{cases}$$

Acknowledgements

GA thanks economical support from the Universidad Complutense de Madrid Program “Ayudas para Becas y Contratos Complutenses Predoctorales en España 2011”.

MM thanks economical support from the Spanish “Ministerio de Economía y Competitividad” research project MTM2012-36732-C03-01, *Ortogonalidad y aproximacion; Teoria y Aplicaciones*. The authors will like to thank the anonymous work of one the referees, his comments and suggestions have clearly improve the quality of this paper.

References

[1] M. Adler, P. van Moerbeke, Group factorization, moment matrices and Toda lattices, Int. Math. Res. Not. IMRN 12 (1997) 556–572.
[2] M. Adler, P. van Moerbeke, Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems, Comm. Math. Phys. 207 (1999) 589–620.
[3] M. Adler, P. van Moerbeke, Vertex operator solutions to the discrete KP hierarchy, Comm. Math. Phys. 203 (1999) 185–210.
[4] M. Adler, P. van Moerbeke, The spectrum of coupled random matrices, Ann. of Math. 149 (1999) 921–976.
[5] M. Adler, P. van Moerbeke, Darboux transforms on band matrices, weights and associated polynomials, Int. Math. Res. Not. IMRN 18 (2001) 935–984.
[6] C. Álvarez-Fernández, U. Fidalgo, M. Mañas, The multicomponent 2D Toda hierarchy: generalized matrix orthogonal polynomials, multiple orthogonal polynomials and Riemann–Hilbert problems, Inverse Problems 26 (2010), 055009 (17 pp.).
[7] C. Álvarez-Fernández, U. Fidalgo, M. Mañas, Multiple orthogonal polynomials of mixed type: Gauss–Borel factorization and the multi-component 2D Toda hierarchy, Adv. Math. 227 (2011) 1451–1525.
[8] C. Álvarez-Fernández, M. Mañas, Orthogonal Laurent polynomials on the unit circle, extended CMV ordering and 2D Toda type integrable hierarchies, Adv. Math. 240 (2013) 132–193.
[9] C. Álvarez-Fernández, M. Mañas, On the Christoffel–Darboux formula for generalized matrix orthogonal polynomials, J. Math. Anal. Appl. 418 (2014) 238–247.
[10] R. Apery, Irrationalite de $\zeta(2)$ et $\zeta(3)$, Astèrisque 61 (1979) 11–13.
[11] F. Beukers, Padé Approximation in Number Theory, Lecture Notes in Math., vol. 888, Springer Verlag, Berlin, 1981, pp. 90–99.
[12] P.M. Bleher, A.B.J. Kuijlaars, Random matrices with external source and multiple orthogonal polynomials, Int. Math. Res. Not. IMRN 2004 (2004) 109–129.
[13] J. Coates, On the algebraic approximation of functions, I, II, III, Indag. Math. 28 (1966) 421–461.
[14] E. Daems, A.B.J. Kuijlaars, A Christoffel–Darboux formula for multiple orthogonal polynomials, J. Approx. Theory 130 (2004) 188–200.
[15] E. Daems, A.B.J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions, J. Approx. Theory 146 (2007) 91–114.
[16] S. Delvaux, Average characteristic polynomials for multiple orthogonal polynomial ensembles, J. Approx. Theory 162 (5) (2010) 1033–1067.
[17] S. Delvaux, A.B.J. Kuijlaars, L. Zhang, Critical behavior of nonintersecting Brownian motions at a tacnode, Comm. Pure Appl. Math. 64 (2011) 1305–1383.
[18] U. Fidalgo Prieto, G. López Lagomasino, Nikishin systems are perfect, Constr. Approx. 34 (2011) 297–356.
[19] U. Fidalgo, S. Medina Peralta, J. Minguez Ceniceros, Mixed type multiple orthogonal polynomials: perfectness and interlacing properties of zeros, Linear Algebra Appl. 438 (2013) 1229–1239.
[20] A.S. Fokas, A.R. Its, A.V. Kitaev, The isomonodromy approach to matrix models in 2D quantum gravity, Comm. Math. Phys. (1992) 395–430.
[21] Ch. Hermite, Sur la fonction exponentielle, C. R. Acad. Sci. Paris 77 (1873) 18–24, 74–79, 226–233, 285–293; reprinted in his Oeuvres, Tome III, Gauthier–Villars, Paris, 1912, pp. 150–181.
[22] H. Jager, A simultaneous generalization of the Padé table, I–VI, Indag. Math. 26 (1964) 193–249.
[23] A.B.J. Kuijlaars, Multiple orthogonal polynomial ensembles, Contemp. Math. 507 (2010) 155–176.
[24] K. Mahler, Perfect systems, Compos. Math. 19 (1968) 95–166.
[25] E.M. Nikishin, On simultaneous Padé approximants, Mat. Sb. 113 (1980) 499–519 (in Russian); English translation in: Math. USSR Sbornik 41 (1982) 409–425.
[26] B. Simon, The Christoffel–Darboux kernel, in: Perspectives in Partial Differential Equations, Harmonic Analysis and Applications: A Volume in Honor of Vladimir G. Maz'ya's 70th Birthday, in: Proc. Sympos. Pure Math., vol. 79, 2008, pp. 295–336, arXiv:0806.1528.

[27] V.N. Sorokin, On simultaneous approximation of several linear forms, Vestnik Moskov. Univ. Ser. I Mat. Mekh. 1 (1983) 44–47.
[28] R.P. Stanley, Enumerative Combinatorics, Cambridge University Press, Cambridge, 1998.
[29] W. Van Assche, J.S. Geromino, A.B.J. Kuijlaars, Riemann–Hilbert problems for multiple orthogonal polynomials, in: Bustoz, et al. (Eds.), Special Functions 2000: Current Perspectives and Future Directions, Kluwer Academic Publishers, Dordrecht, 2001, pp. 23–59.

4 Matrix Orthogonal Laurent Polynomials on the Unit Circle and Toda type integrable systems

Advances in Mathematics 264 (2014) 396–463

Contents lists available at ScienceDirect

Advances in Mathematics

www.elsevier.com/locate/aim

Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems

Gerardo Ariznabarreta [1], Manuel Mañas [*,2]

Departamento de Física Teórica II (Métodos Matemáticos de la Física), Universidad Complutense de Madrid, 28040-Madrid, Spain

A R T I C L E   I N F O

Article history:
Received 2 December 2013
Accepted 24 June 2014
Available online xxxx
Communicated by Charles Fefferman

MSC:
15A23
33C45
37K10
37L60
42C05
46L55

Keywords:
Matrix orthogonal Laurent polynomials
Borel–Gauss factorization
Christoffel–Darboux kernels
Toda type integrable hierarchies

A B S T R A C T

Matrix orthogonal Laurent polynomials in the unit circle and the theory of Toda-like integrable systems are connected using the Gauss–Borel factorization of two, left and a right, Cantero–Morales–Velázquez block moment matrices, which are constructed using a quasi-definite matrix measure. A block Gauss–Borel factorization problem of these moment matrices leads to two sets of biorthogonal matrix orthogonal Laurent polynomials and matrix Szegő polynomials, which can be expressed in terms of Schur complements of bordered truncations of the block moment matrix. The corresponding block extension of the Christoffel–Darboux theory is derived. Deformations of the quasi-definite matrix measure leading to integrable systems of Toda type are studied. The integrable theory is given in this matrix scenario; wave and adjoint wave functions, Lax and Zakharov–Shabat equations, bilinear equations and discrete flows — connected with Darboux transformations. We generalize the integrable flows of the Cafasso's matrix extension of the Toeplitz lattice for the Verblunsky coefficients of Szegő polynomials. An analysis of the Miwa shifts allows for the finding of interesting connections between Christoffel–Darboux kernels and Miwa shifts of the matrix orthogonal Laurent polynomials.

* Corresponding author.
E-mail addresses: gariznab@ucm.es (G. Ariznabarreta), manuel.manas@ucm.es (M. Mañas).

[1] G.A. thanks economical support from the Universidad Complutense de Madrid Program "Ayudas para Becas y Contratos Complutenses Predoctorales en España 2011".

[2] M.M. thanks economical support from the Spanish "Ministerio de Economía y Competitividad" research project MTM2012-36732-C03-01, *Ortogonalidad y aproximación; teoría y aplicaciones*.

http://dx.doi.org/10.1016/j.aim.2014.06.019

Contents

1. Introduction

In this paper we extend previous results on orthogonal Laurent polynomials in the unit circle (OLPUC) [13] to the matrix realm (MOLPUC). To explain better our aims and results we need a brief account on orthogonal polynomials, Laurent orthogonal polynomials and their matrix extensions, and also some facts about integrable systems.

1.1. Historical background

1.1.1. Szegő polynomials

We will denote the unit circle by $\mathbb{T} := \{z \in \mathbb{C} : |z| = 1\}$ and $\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$ stands for the unit disk; when $z \in \mathbb{T}$ we will use the parametrization $z = \mathrm{e}^{\mathrm{i}\theta}$ with $\theta \in [0, 2\pi)$. In the scalar case, one deals with a complex Borel measure μ supported in $\mathbb{T}$ that is said to be positive definite if it maps measurable sets onto non-negative numbers, that in the absolutely continuous situation (with respect to the Lebesgue measure $\mathrm{d}\theta$) has the form $w(\theta)\mathrm{d}\theta$. For the positive definite situation the orthogonal polynomials in the unit circle (OPUC) or Szegő polynomials are defined as those monic polynomials P_n of degree n that satisfy the following system of equations, called orthogonality relations: $\int_{\mathbb{T}} P_n(z) z^{-k} \mathrm{d}\mu(z) = 0$, for $k = 0, 1, \dots, n-1$ [90]. The connections between orthogonal polynomials on the real line (OPRL) supported in the interval $[-1, 1]$ and OPUC have been explored in the literature, see for example [53,22]. Let us observe that for this analysis the use of spectral theory techniques requires the study of the operator of multiplication by z. Recursion relations for OPRL and OPUC are well known; however, in the real case, the three term recurrence laws provide a tridiagonal matrix, the so-called Jacobi operator, while in the unit circle support case, the problem leads to a Hessenberg matrix [61], being a more involved scenario that the Jacobi one (as it is not a sparse matrix with a finite number of non-vanishing diagonals). In fact, OPUCs recursion relation requires the introduction of reciprocal or reverse Szegő polynomials $P_l^*(z) := z^l \overline{P_l(\bar{z}^{-1})}$ and the reflection or Verblunsky (Schur parameters is another usual name) coefficients $\alpha_l := P_l(0)$. The recursion relations for the Szegő polynomials can be written as $\left(\begin{smallmatrix} P_l \\ P_l^* \end{smallmatrix}\right) = \left(\begin{smallmatrix} z & \alpha_l \\ z\bar{\alpha}_l & 1 \end{smallmatrix}\right)\left(\begin{smallmatrix} P_{l-1} \\ P_{l-1}^* \end{smallmatrix}\right)$. There exist numerous studies on the zeroes of the OPUC [10,16,19,54,58,60,74,80] with interesting applications to signal analysis theory [63,65,82,83]. Despite the mentioned advances for the OPUC theory, the corresponding state of the art in the OPRL context is still much more developed. An issue to stress here is that Szegő polynomials are, in general, not a dense set in the Hilbert space $L^2(\mathbb{T}, \mu)$; Szegő's theorem implies for a non-trivial probability measure $\mathrm{d}\mu$ on $\mathbb{T}$ with Verblunsky coefficients $\{\alpha_n\}_{n=0}^{\infty}$ that the corresponding Szegő's polynomials are dense in $L^2(\mathbb{T}, \mu)$ if and only if $\prod_{n=0}^{\infty}(1 - |\alpha_n|^2) = 0$. For an absolutely continuous probability measure Kolmogorov's density theorem ensures that density in $L^2(\mathbb{T}, \mu)$ of the OPUC holds iff the so-called Szegő's condition $\int_{\mathbb{T}} \log(w(\theta))\mathrm{d}\theta = -\infty$ is fulfilled [89]. We refer the reader to Barry Simon's books [85] and [86] for a very detailed studied of OPUC.

1.1.2. Orthogonal Laurent polynomials

Orthogonal Laurent polynomials on the real line (OLPRL) were introduced in [66, 67] in the context of the strong Stieltjes moment problem. When this moment problem has a solution, there exist polynomials $\{Q_n\}$, known as Laurent polynomials, such that $\int_{\mathbb{R}} x^{-n+j} Q_n(x)\mathrm{d}\mu(x) = 0$ for $j = 0, \ldots, n-1$. The theory of Laurent polynomials on the real line was developed in parallel with the theory of orthogonal polynomials, see [33,43,64] and [81]. The theory of orthogonal Laurent polynomials was carried from the real line to the circle [91] and subsequent works broadened the matter (e.g. [37,29,35,36]), treating subjects like recursion relations, Favard's theorem, quadrature problems, and Christoffel–Darboux formulae. The Cantero–Moral–Velázquez (CMV) [29] representation is a hallmark in the study of certain aspects of Szegő polynomials, as we mentioned already while the OLPUC are always dense in $L^2(\mathbb{T}, \mu)$, this is not true in general for the OPUC [25,37]. The bijection between OLPUC in the CMV representation and the ordinary Szegő polynomials implies the replacement of complicated recursion relations with five term relations similar to the OPRL situation. Other papers have reviewed and broadened the study of CMV matrices, see for example [87,68]; in particular alternative or generic orders in the base used to span the space of OLPUC can be found in [36]. In particular, the reading of Simon's account of the CMV theory [87] is illuminating. In fact, the discovery of the advantages of the CMV ordering goes back to the previous work [93].

1.1.3. Matrix orthogonal polynomials

Orthogonal polynomials with matrix coefficients on the real line were considered in detail by Krein [69,70] in 1949, and thereafter were studied sporadically until the last decade of the XX-th century. Some relevant papers on this subject are [20,56,17]; in particular, in [17] the scattering problem is solved for a kind of discrete Sturm–Liouville operators that are equivalent to the recursion equation for scalar orthogonal polynomials. They found that polynomials that satisfy a relation of the form

$$xP_k(x) = A_k P_{k+1}(x) + B_k P_k(x) + A^*_{k-1} P_{k-1}(x), \quad k = 0, 1, \dots,$$

are orthogonal with respect to a positive definite measure. This is a matrix version of Favard's theorem for scalar orthogonal polynomials. Then, in the 1990s and the 2000s some authors found that matrix orthogonal polynomials (MOPs) satisfy in certain cases some properties that satisfy scalar-valued orthogonal polynomials; for example, Laguerre, Hermite and Jacobi polynomials, i.e., the scalar-type Rodrigues' formula [47,48,34] and a second order differential equation [44,46,24]. Later on, it has been proven [45] that operators of the form $D = \partial^2 F_2(t) + \partial^1 F_1(t) + \partial^0 F_0$ have as eigen-functions different infinite families of MOPs. Moreover, in [24] a new family of MOPs satisfying second order differential equations, whose coefficients do not behave asymptotically as the identity matrix, was found; see also [30]. In [31] the Riemann–Hilbert problem for this matrix situation and the appearance of non-Abelian discrete versions of Painlevé I were explored, showing singularity confinement — see [32]; for Riemann–Hilbert problems see also [62]. Let us mention that in [75,76] and [27] the MOPs are expressed in terms of Schur complements that play the role of determinants in the standard scalar case. For a survey on matrix orthogonal polynomials, we refer the reader to [39].

1.1.4. Integrable hierarchies and the Gauss–Borel factorization

The seminal paper of M. Sato [84] and further developments performed by the Kyoto school [40–42] settled the Lie-group theoretical description of the integrable hierarchies. It was Mulase [78] the one who made the connection between factorization problems, dressing procedures and integrability. In this context, Ueno and Takasaki [92] performed an analysis of the Toda type hierarchies and their soliton-like solutions. Adler and van Moerbeke [4–8,3,9] have clarified the connection between the Lie-group factorization, applied to Toda type hierarchies — what they call discrete Kadomtsev–Petviashvilii (KP) — and the Gauss–Borel factorization applied to a moment matrix that comes from orthogonality problems; thus, the corresponding orthogonal polynomials are closely related to specific solutions of the integrable hierarchy. See [21,52,71,11] for further developments in relation with the factorization problem, multicomponent Toda lattices and generalized orthogonality. In [3] a profound study of the OPUC and the Toda type associated lattice, called the Toeplitz lattice (TL), was performed. A relevant reduction of the equations of the TL has been found by Golinskii [59] in the context of Schur flows when the measure is invariant under conjugation (also studied in [88] and [49]), another interesting paper on this subject is [77]. The Toeplitz lattice was proven to be equivalent to the Ablowitz–Ladik lattice (ALL) [1,2], and that work has been generalized to the link between matrix orthogonal polynomials and the non-Abelian ALL in [27]. Both of them have to deal with the Hessenberg operator for the multiplication by z. Research about the integrable structure of Schur flows and its connection with ALL has been done (in recent and not so recent works) from a Hamiltonian point of view in [79], and other works also introduce connections with Laurent polynomials and τ-functions, like [50,51,23].

1.2. *Preliminary material*

1.2.1. *Semi-infinite block matrices*

For the matrix extension considered in the present work we need to deal with block matrices and block Gauss–Borel factorizations. For each $m \in \mathbb{N}$, the directed set of natural numbers, we consider ring of the complex $m \times m$ matrices $\mathbb{M}_m := \mathbb{C}^{m\times m}$, and its direct limit $\mathbb{M}_\infty := \varinjlim \mathbb{M}_m$, the ring of semi-infinite complex matrices. We will denote by $\operatorname{diag}_m \subset \mathbb{M}_m$ the set of diagonal matrices. For any $A \in \mathbb{M}_\infty$, $A_{ij} \in \mathbb{C}$ denotes the (i,j)-th element of A, while $(A)_{ij} \in \mathbb{M}_m$ denotes the (i,j)-th block of it when subdivided into $m \times m$ blocks. We will denote by G_∞ the group of invertible semi-infinite matrices of $\mathbb{M}_\infty$. In this paper two important subgroups are $\mathscr{U}$, the invertible upper triangular — by blocks — matrices, and $\mathscr{L}$, the lower triangular — by blocks — matrices with the identity matrix along their block diagonal. The corresponding restriction on invertible upper triangular block matrices is denoted by $\widehat{\mathscr{U}}$. Block diagonal matrices will be denoted by $\mathscr{D} = \{D \in \mathbb{M}_\infty : (D)_{i,j} = d_i \cdot \delta_{i,j} \text{ with } d_i \in \mathbb{M}_m\}$. Given a semi-infinite matrix $A \in \mathbb{M}_\infty$, we consider its l-th block leading submatrix

$$A^{[l]} = \begin{pmatrix} (A)_{0,0} & (A)_{0,1} & \dots & (A)_{0,l-1} \\ (A)_{1,0} & (A)_{1,1} & \dots & (A)_{1,l-1} \\ \vdots & & & \vdots \\ (A)_{l-1,0} & (A)_{l-1,1} & \dots & (A)_{l-1,l-1} \end{pmatrix} \in \mathbb{M}_{ml}, \quad (A)_{i,j} \in \mathbb{M}_m,$$

and we write

$$A = \left(\begin{array}{c|c} A^{[l]} & A^{[l,\geq l]} \\ \hline A^{[\geq l,l]} & A^{[\geq l]} \end{array}\right), \tag{1}$$

for the corresponding block partition of a matrix A where, for example, $A^{[l,\geq l]}$ denotes all the $(A)_{i,j}$-th blocks of the matrix A with $i < l, j \geq l$. Very much related to the block partition of a matrix M are the Schur complements. The Schur complement with respect to the upper left block of the block partition

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathbb{M}_{p+q}, \quad A = (a_{i,j}) \in \mathbb{M}_p, \ D \in \mathbb{M}_q,$$

is

$$M\diagup A := \mathrm{SC}\left(\begin{array}{c|c}(a_{i,j}) & B \\ \hline C & D\end{array}\right) := D - CA^{-1}D,$$

where we have assumed that A is an invertible matrix.

1.2.2. Quasi-definiteness

Let us recall the reader that measures and linear functionals are closely connected; given a linear functional $\mathcal{L}$ on $\Lambda_{[\infty]}$, the set of Laurent polynomials on the circle — or polynomial loops $L_{\mathrm{pol}}\mathbb{C}$, we define the corresponding moments of $\mathcal{L}$ as $c_n := \mathcal{L}[z^n]$ for all the possible integer values of $n \in \mathbb{Z}$. The functional $\mathcal{L}$ is said to be Hermitian whenever $c_{-n} = \overline{c_n}$, $\forall n \in \mathbb{Z}$. Moreover, the functional $\mathcal{L}$ is defined as quasi-definite (positive definite) when the principal submatrices of the Toeplitz moment matrix $(\Delta_{i,j})$, $\Delta_{i,j} := c_{i-j}$, associated to the sequence c_n, are non-singular (positive definite), i.e., $\forall n \in \mathbb{Z}$, $\Delta_n := \det(c_{i-j})_{i,j=0}^n \neq 0$ (> 0). Some aspects on quasi-definite functionals and their perturbations are studied in [15,26]. It is known [57] that when the linear functional $\mathcal{L}$ is Hermitian and positive definite, there exists a finite positive Borel measure with a support lying on $\mathbb{T}$ such that $\mathcal{L}[f] = \int_{\mathbb{T}} f \mathrm{d}\mu$, $\forall f \in \Lambda_{[\infty]}$. In addition, a Hermitian positive definite linear functional $\mathcal{L}$ defines a sesquilinear form $\langle\cdot,\cdot\rangle_{\mathcal{L}} : \Lambda_{[\infty]} \times \Lambda_{[\infty]} \mapsto \mathbb{C}$ as $\langle f, g\rangle_{\mathcal{L}} = \mathcal{L}[f\bar{g}]$, $\forall f, g \in \Lambda_{[\infty]}$. Two Laurent polynomials $\{f, g\} \subset \Lambda_{[\infty]}$ are said to be orthogonal with respect to $\mathcal{L}$ if $\langle f, g\rangle_{\mathcal{L}} = 0$. From the properties of $\mathcal{L}$ it is easy to see that $\langle\cdot,\cdot\rangle_{\mathcal{L}}$ is a scalar product and if μ is the positive finite Borel measure associated to $\mathscr{L}$ we are led to the corresponding Hilbert space $L^2(\mathbb{T}, \mu)$, the closure of $\Lambda_{[\infty]}$. The more general setting, when $\mathcal{L}$ is just quasi-definite is associated to a corresponding quasi-definite complex measure μ, see [55]. As before, a sesquilinear form $\langle\cdot,\cdot\rangle_{\mathcal{L}}$ is defined for any such linear functional $\mathcal{L}$; thus, we just have the linearity (in the first entry) and skew-linearity (in the second entry) properties. However, we have no symmetry allowing the interchange of the two arguments. We formally broaden the notion of orthogonality and say that f is orthogonal to g if $\langle f, g\rangle_{\mathcal{L}} = 0$, but we must be careful as in this general situation it could happen that $\langle f, g\rangle_{\mathcal{L}} = 0$ but $\langle g, f\rangle_{\mathcal{L}} \neq 0$.

1.2.3. Matrix Laurent polynomials and orthogonality

A matrix-valued measure $\mu = (\mu_{i,j})$ supported on $\mathbb{T}$ is said to be Hermitian and/or positive definite, if for every Borel subset $\mathscr{B}$ of $\mathbb{T}$, the matrix $\mu(\mathscr{B})$ is a Hermitian and/or positive definite matrix. When the scalar measures $\mu_{i,j}$, $i, j = 1, \ldots, m$, are absolutely continuous with respect to the Lebesgue measure on the circle $\mathrm{d}\theta$, according to the Radon–Nikodym theorem, it can be always expressed using complex weight (density or Radon–Nikodym derivative of the measure) functions $w_{i,j}$, $i, j = 1, \ldots, m$, so that $\mathrm{d}\mu_{i,j}(\theta) = w_{i,j}(\theta)\mathrm{d}\theta$, $\theta \in [0, 2\pi)$. If, in addition, the matrix measure μ is Hermitian and positive definite, then the matrix $(w_{i,j}(\theta))$ is a positive definite Hermitian matrix. For the sake of notational simplicity we will use, whenever it is convenient, the complex notation $\mathrm{d}\mu(z) = \mathrm{i}\mathrm{e}^{\mathrm{i}\theta}\mathrm{d}\mu(\theta)$.

The moments of the matrix measure μ are

$$c_n := \frac{1}{2\pi}\oint_{\mathbb{T}} z^{-n}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z} = \frac{1}{2\pi}\int_0^{2\pi} \mathrm{e}^{-\mathrm{i}n\theta}\mathrm{d}\mu(\theta) \in \mathbb{M}_m,$$

while the Fourier series of the measure is

$$F_\mu(u) := \sum_{n=-\infty}^{\infty} c_n u^n, \tag{2}$$

that for absolutely continuous measures, $\mathrm{d}\mu(\theta) = w(\theta)\mathrm{d}\theta$ satisfies $F_\mu(\theta) = w(\theta)$. Let $D(0; r, R) = \{z \in \mathbb{C} : r < |z| < R\}$ denote the annulus around $z = 0$ with interior and exterior radii r and R, $R_{ij,\pm} := (\limsup_{n\to\infty} \sqrt[n]{|c_{ij,\pm n}|}\,)^{\mp 1}$ and $R_+ = \min_{i,j=1,\dots,m} R_{ij,+}$ and $R_{ij,-} = \max_{i,j=1,\dots,m} R_{ij,-}$. Then, according to the Cauchy–Hadamard theorem, the series $F_\mu(z)$ converges uniformly in any compact set K, $K \subset D(0; R_-, R_+)$.

The space $\Lambda_{m,[p,q]} := \mathbb{M}_m\{\mathbb{I}z^{-p}, \mathbb{I}z^{-p+1}, \dots, \mathbb{I}z^q\}$ (where $\mathbb{I} \in \mathbb{M}_m$ is the identity matrix) of complex Laurent polynomials with $m \times m$ matrix coefficients and the corresponding restrictions on their degrees is an $\mathbb{M}_m$ free module of rank $p+q+1$. We denote by $L_{\mathrm{pol}}\mathbb{M}_m$ the infinite set of Laurent matrix polynomials or polynomial loops in $\mathbb{M}_m$.

Given a matrix measure μ, we introduce the following left and right matrix-valued sesquilinear forms in the loop space $L\mathbb{M}_m$ considered as left and right modules for the ring $\mathbb{M}_m$, respectively,

$$\langle\!\langle f, g\rangle\!\rangle_L := \oint_{\mathbb{T}} g(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z} f(z)^\dagger \in \mathbb{M}_m, \tag{3}$$

$$\langle\!\langle f, g\rangle\!\rangle_R := \oint_{\mathbb{T}} f(z)^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z} g(z) \in \mathbb{M}_m. \tag{4}$$

The sesquilinearity of these forms means that the following two properties hold:

(1) $\langle\!\langle f_1 + f_2, g\rangle\!\rangle_H = \langle\!\langle f_1, g\rangle\!\rangle_H + \langle\!\langle f_2, g\rangle\!\rangle_H$ and $\langle\!\langle f, g_1 + g_2\rangle\!\rangle_H = \langle\!\langle f, g_1\rangle\!\rangle_H + \langle\!\langle f, g_2\rangle\!\rangle_H$ for all $f, f_1, f_2, g, g_1, g_2 \in L\mathbb{M}_m$ and $H = L, R$.

(2) $\langle\!\langle mf, g\rangle\!\rangle_L = \langle\!\langle f, g\rangle\!\rangle_L m^\dagger$, $\langle\!\langle f, mg\rangle\!\rangle_L = m\langle\!\langle f, g\rangle\!\rangle_L$, $\langle\!\langle fm, g\rangle\!\rangle_R = m^\dagger\langle\!\langle f, g\rangle\!\rangle_R$ and $\langle\!\langle f, gm\rangle\!\rangle_R = \langle\!\langle f, g\rangle\!\rangle_R m$, for all $f, g \in L\mathbb{M}_m$ and $m \in \mathbb{M}_m$.

Moreover, if the matrix measure is Hermitian, then so are these forms; i.e.,

$$\langle\!\langle f, g\rangle\!\rangle_H^\dagger = \langle\!\langle g, f\rangle\!\rangle_H, \quad H = L, R.$$

Actually, from these sesquilinear forms, for a positive definite Hermitian measure, we can derive the corresponding scalar products

$$\langle f,g\rangle_H^{\dagger}=\langle f,g\rangle_H:=\operatorname{Tr}\big[\langle\!\langle f,g\rangle\!\rangle_H\big],\qquad \|f\|_H^2=\langle f,f\rangle_H,\quad H=L,R,$$

and corresponding Hilbert spaces $\mathcal{H}^H$ with a norm — of Frobenius type — given by

$$\|f\|_H=+\sqrt{\langle f,f\rangle_H},\quad H=L,R.$$

A set $\{p_l\}_{l=0}^N\subset\mathcal{H}^H$, $H=L,R$, is an orthogonal set if and only if

$$\langle\!\langle p_l^H,p_j^H\rangle\!\rangle_K=\delta_{ij}h_j,\quad h_j\in\mathbb{M}_m.$$

1.3. On the content of the paper

In previous papers we have approached the study of the link between orthogonality and integrability within an algebraic/group theoretical point of view. Our keystone relies on the fact that a number of facets of orthogonality and integrability can be described with the aid of the Gauss–Borel factorization of an infinite matrix. This approach was applied in [12] for the analysis of multiple orthogonal polynomials of mixed type, allowing for an algebraic proof of the Christoffel–Darboux formula, alternative to the analytic one, based on the Riemann–Hilbert problem (and constrained therefore by convenient analytic conditions) given in [38]. This approach was also used successfully in [13], where a CMV ordering of the Fourier basis gave, for a given measure on the unitary circle, a moment matrix whose Gauss–Borel factorization leads to OLPUC. Recursion relations and Christoffel–Darboux formula appeared also in a straightforward manner. Also continuous and discrete deformations and τ-function theory were extended to the circular case under the suitable choice of moment matrices and shift operators. In this last paper we only requested to the measure to be quasi-definite, condition that implies the existence of the Gauss–Borel factorization. Let us mention that we have applied this method in the finding of Christoffel–Darboux type formulae in other situations, see [18,14].

In this paper we consider two semi-infinite block matrices, whose coefficients (matrices in $\mathbb{C}^{m\times m}$) are left and right matrix moments, ordered in a Cantero–Morales–Velázquez style, of a matrix measure on the circle. The corresponding block Gauss–Borel factorization of these CMV block moment matrices leads to MOLPUC. To be more precise, we get the right and left versions of two biorthogonal families of matrix Laurent polynomials and corresponding Szegő polynomials. When the matrix measure is Hermitian, these two families happen to be proportional resulting in two families of MOLPUC. Following [75,27] we express them as Schur complements of bordered truncated moment matrices. We also prove, in an algebraic manner using the Gauss–Borel factorization, the five term recursion relations and the Christoffel–Darboux formula. Let us stress that in this paper we introduce an intertwining operator η not used in [13] that clarifies the appearance of reciprocal polynomials and simplifies the algebraic proofs. The recursion relations indicate which deformations of the quasi-definite matrix measure lead to integrable systems of Toda type. Thus, we discuss the following elements: wave and adjoint wave functions, Lax and Zakharov–Shabat equations, bilinear equations and discrete flows — connected with Darboux transformations. In this context we find a generalization of the matrix Cafasso's extension of the Toeplitz lattice for the Verblunsky coefficients of Szegő polynomials. The Cafasso flows correspond to what we call total flows, which are only a part of the integrable flows associated to MOLPUC. We unsuccessfully tried to get a matrix τ theory, but despite this failure, we get interesting byproducts. We analyze the role of Miwa shifts in this context and, as a collateral effect, nicely connect them with the Christoffel–Darboux kernels. These formulae suggest a link of these kernels with the Cauchy propagators that in the Grassmannian $\bar{\partial}$ approach to multicomponent KP hierarchy was used in [72,73]. This identification allows us to give in Theorem 6 expressions of the MOLPUC in terms of products of their Miwa shifted and non-shifted quasi-norms. Despite that these expressions lead to the τ-function representation in the scalar case, this is not the case within the matrix context.

Let us mention that the submodules of matrix Laurent polynomials considered in this paper have the higher and lower powers constrained to be of some particular form, implied by the chosen CMV ordering. In [13] this limitation was overcome by the introduction of extended CMV orderings of the Fourier basis, which allowed for general subspaces of Laurent polynomials. A similar procedure can be performed in this matrix situation; but, as its development follows very closely the ideas of [13], we prefer to avoid its inclusion here.

The layout of this paper is as follows. Section 2 is devoted to orthogonality theory, in particular in Section 2.1 we consider the left and right block CMV moment matrices and perform corresponding block Gaussian factorizations in Section 2.2, getting the associated families of right and left MOLPUC and matrix Szegő polynomials and their biorthogonality relations. We also get the recursion relations and Schur complement expressions of them in terms of bordered truncations of the moment matrices. Then, in Section 2.3 we introduce the matrix second kind functions that are connected with the Fourier series of the measure and that will be relevant later on for the adjoint Baker functions. The reconstruction of the recursion relations from the Gauss–Borel factorizations is performed in Section 2.4; the Christoffel–Darboux formulae for this non-Abelian scenario are given in Section 2.5. Observe that in this case, the projection operators are projectors in a module over the ring $\mathbb{C}^{m\times m}$, that in the Hermitian definite positive situation lead to orthogonal projections in the standard geometrical sense. The integrability aspects are treated in Section 3. Given adequate deformations of the moment matrices, we find wave functions, Lax equations and Zakharov–Shabat equations in Section 3.1; here we also consider a generalization of the Cafasso's Toeplitz lattice and the bilinear equations formulation of the hierarchy. Finally, we extend to this matrix context the discrete flows for the Toeplitz lattice, intimately related to Darboux transformations in Section 3.2 and also derive the bilinear equations fulfilled by the MOLPUC in Section 3.1.3. Finally, in Section 3.3 we consider the action of Miwa transformations and get the previously mentioned results. We conclude the paper with a series of appendices that serve as support of certain sections.

Finally, let us stress that this paper is not just an extension of the results of [13] to the matrix realm but we also have introduced important elements not discussed there, which also hold in that scalar case, as the η operator, a different proof of the Christoffel–Darboux formula with no need of associated polynomials and new relations between Christoffel–Darboux kernels and Miwa shifted MOLPUC.

2. Matrix orthogonality and block Gauss–Borel factorization

In this section, inspired by the CMV construction [29] and the previous work [13], for a given matrix measure, we introduce an appropriate block moment matrix that, when factorized as a product of lower and upper block matrices, gives a set of biorthogonal matrix Laurent polynomials on the unit circle. This Borel–Gauss factorization problem also allows us to derive the recursion relations and the Christoffel–Darboux theory.

2.1. The CMV right and left moment matrices for quasi-definite matrix measures

The following $m \times m$ matrix-valued vectors will be relevant in the construction of biorthogonal families of MOLPUC

Definition 1. The CMV vectors are given by

$$\begin{aligned}
\chi_1(z) &:= \left(\mathbb{I}, 0, \mathbb{I}z, 0, \mathbb{I}z^2, \dots\right)^\top, \\
\chi_2(z) &:= \left(0, \mathbb{I}, 0, \mathbb{I}z, 0, \mathbb{I}z^2, \dots\right)^\top, \\
\chi_a^*(z) &:= z^{-1}\chi_a\left(z^{-1}\right), \quad a = 1, 2, \\
\chi(z) &:= \chi_1(z) + \chi_2^*(z) = \left(\mathbb{I}, \mathbb{I}z^{-1}, \mathbb{I}z, \mathbb{I}z^{-2}, \mathbb{I}z^2, \dots\right)^\top.
\end{aligned}$$

In the sequel, the matrix $\chi^{(l)}$ will denote the l-th component of the matrix vector χ

$$\chi = \left(\chi^{(0)}, \chi^{(1)}, \chi^{(2)}, \dots\right)^\top.$$

Definition 2. The CMV left and right moment matrices of the measure μ are

$$g^L := \oint_{\mathbb{T}} \chi(z) \frac{\mathrm{d}\mu(z)}{\mathrm{i}z} \left(\chi(z)\right)^\dagger = 2\pi \begin{pmatrix} c_0 & c_{-1} & c_1 & c_{-2} & \cdots \\ c_1 & c_0 & c_2 & c_{-1} & \cdots \\ c_{-1} & c_{-2} & c_0 & c_{-3} & \cdots \\ c_2 & c_1 & c_3 & c_0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \tag{5}$$

$$g^R := \oint_{\mathbb{T}} \big(\chi(z)^\top\big)^\dagger \frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^\top = 2\pi \begin{pmatrix} c_0 & c_1 & c_{-1} & c_2 & \dots \\ c_{-1} & c_0 & c_{-2} & c_1 & \dots \\ c_1 & c_2 & c_0 & c_3 & \\ c_{-2} & c_{-1} & c_{-3} & c_0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{6}$$

Notice that when $\mathrm{d}\mu(\theta)$ is Hermitian, so are the moment matrices g^L and g^R.

In the scalar case [13], the only requirement that the moment matrix needs to meet is to be Gaussian factorable; i.e., all the principal minors of the matrix are requested to be not degenerated. The measure from which this moment matrix is constructed receives the name of quasi-definite measure. This condition is related to the existence of *biorthogonal* polynomials of all degrees — also called non-triviality of the measure. In the matrix case, the requirement is a bit more relaxed.

Definition 3. The matrix measure μ is said to be *quasi-definite* if its truncated moment matrices satisfy

$$\det\big(\big(g^H\big)^{[l]}\big) \neq 0 \quad \text{for } H = R, L \text{ and } l = 1, 2, 3, \dots .$$

Notice that $(g^H)^{[l]} \in \mathbb{M}_{ml}$; a quite different situation from the scalar case in which all the principal minors had to be non-degenerate, while in the matrix case only the ml-order principal minors should meet this requirement. Actually, this is the only restriction — besides having compact support on $\mathbb{T}$ — that from hereon the matrix measures must satisfy, since when this condition holds

Proposition 1. *The moment matrices g^H, $H = L, R$, of a matrix quasi-definite measure μ admit a block Gauss–Borel factorization.*

Proof. See Appendix A. $\square$

2.1.1. The generalized matrix Szegő polynomials

Definition 4. Given a matrix quasi-definite measure μ, the set of monic matrix polynomials $\{P^L_{i,l}\}_{l=0}^\infty$, $\{P^R_{i,l}\}_{l=0}^\infty$, $i = 1, 2$, with $\deg P^H_{i,l} = l$, $H = L, R$, satisfying

$$\big\langle\big\langle z^j\mathbb{I}, P^L_{1,l}(z)\big\rangle\big\rangle_L = \oint_{\mathbb{T}} P^L_{1,l}(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-j} = 0, \quad j = 0, \dots, l-1,$$

$$\big\langle\big\langle P^L_{2,l}(z), z^j\mathbb{I}\big\rangle\big\rangle_L = \oint_{\mathbb{T}} z^j\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\big[P^L_{2,l}(z)\big]^\dagger = 0, \quad j = 0, \dots, l-1,$$

$$\big\langle\big\langle P^R_{2,l}(z), z^j\mathbb{I}\big\rangle\big\rangle_R = \oint_{\mathbb{T}} \big[P^R_{2,l}(z)\big]^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^j = 0, \quad j = 0, \dots, l-1,$$

$$\langle\!\langle z^j\mathbb{I}, P^R_{1,l}(z)\rangle\!\rangle_R = \oint_{\mathbb{T}} z^{-j}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}P^R_{1,l}(z) = 0, \quad j = 0,\dots,l-1,$$

are said to be Szegő polynomials.

Proposition 2. *The matrix Szegő polynomials introduced in Definition 4 for the quasi-definite situation exist and are unique. Moreover, there exist matrices $h^H_r \in \mathbb{M}_m$, $H = L, R$, such that the biorthogonality conditions are fulfilled*

$$\delta_{r,j}h^H_r := \langle\!\langle P^H_{2,r}, P^H_{1,j}\rangle\!\rangle_H, \quad H = R, L.$$

Now we introduce the matrix extension of the Verblunsky coefficients.

Definition 5. The Verblunsky matrices of a matrix quasi-definite measure are

$$\alpha^H_{i,l} := P^H_{i,l}(0), \quad i = 1,2, \ l = 1,2,3,\dots, \ H = L, R,$$

and the reciprocal or reversed Szegő matrix polynomials are given by

$$\big(P^H_l\big)^*(z) := z^l\big(P^H_l\big(\bar z^{-1}\big)\big)^\dagger, \quad H = L, R.$$

Notice that in the Hermitian positive definite case, the matrices h^H_l, $H = L, R$, $l = 0,1,2,\dots$, can be interpreted as a kind of “matrix-valued norms” for the matrix Szegő polynomials, as the square-root of their traces is a norm indeed.

2.2. The CMV matrix Laurent polynomials

We consider now the $m\times m$ block LU factorization of the moment matrices (5) and (6); in fact, there are two block Gauss–Borel factorizations, for both the right and left moment matrices, to consider

$$g^L := S_1^{-1}D^L\widehat{S_2} = S_1^{-1}S_2, \quad S_1\in\mathscr{L},\ S_2\in\mathscr{U},\ \widehat{S_2}\in\widehat{\mathscr{U}},\ D_L\in\mathscr{D}, \tag{7}$$

$$g^R := Z_2D^R\widehat{Z_1}^{-1} = Z_2Z_1^{-1}, \quad Z_2\in\mathscr{L},\ Z_1\in\mathscr{U},\ \widehat{Z_1}\in\widehat{\mathscr{U}},\ D_R\in\mathscr{D}. \tag{8}$$

For the entries of the block diagonal matrices, we use the notation

$$D^H = \operatorname{diag}\big(D^H_0, D^H_1,\dots\big), \quad H = L, R. \tag{9}$$

The reader should notice that in the Hermitian case, the two normalized matrices of the factorization are related

$$S_1^\dagger = \widehat{S_2}^{-1}, \qquad Z_2^\dagger = \widehat{Z_1}^{-1}, \tag{10}$$

and the block diagonal matrices are Hermitian; $(D^H)^\dagger = D^H$, $H = L, R$.

Definition 6. We introduce the following partial CMV matrix Laurent polynomials

$$\begin{aligned}
\phi_{1,1}^L &:= S_1\chi_1(z), & \phi_{1,2}^L &:= S_1\chi_2^*(z),\\
\phi_{2,1}^L &:= \left(S_2^{-1}\right)^\dagger\chi_1(z), & \phi_{2,2}^L &:= \left(S_2^{-1}\right)^\dagger\chi_2^*(z),\\
\phi_{1,1}^R &:= \chi_1^\top(z)Z_1, & \phi_{1,2}^R &:= \left[\chi_2^*\right]^\top(z)Z_1,\\
\phi_{2,1}^R &:= \chi_1^\top(z)\left(Z_2^{-1}\right)^\dagger, & \phi_{2,2}^R &:= \left[\chi_2^*\right]^\top(z)\left(Z_2^{-1}\right)^\dagger,
\end{aligned}$$

and CMV matrix Laurent polynomials

$$\phi_1^L := \phi_{1,1}^L+\phi_{1,2}^L = S_1\chi(z), \qquad \phi_2^L := \phi_{2,1}^L+\phi_{2,2}^L = \left(S_2^{-1}\right)^\dagger\chi(z), \tag{11}$$

$$\phi_1^R := \phi_{1,1}^R+\phi_{1,2}^R = \chi^\top(z)Z_1, \qquad \phi_2^R := \phi_{2,1}^R+\phi_{2,2}^R = \chi^\top(z)\left(Z_2^{-1}\right)^\dagger. \tag{12}$$

Notice that these semi-infinite vectors with matrix coefficients $(\varphi_j^H)^{(l)}(z)$, $l=0,1,\dots,$ can be written as

$$\phi_j^L =: \begin{pmatrix} (\varphi_j^L)^{(0)}(z)\\ (\varphi_j^L)^{(1)}(z)\\ \vdots \end{pmatrix}, \qquad \phi_j^R =: \left(\left(\varphi_j^R\right)^{(0)}(z), \left(\varphi_j^R\right)^{(1)}(z),\dots\right), \quad j=1,2.$$

For the Hermitian case, we have

$$\begin{aligned}
&\left(\varphi_2^L\right)^{(l)}(z) = \left(D_l^L\right)^{-1}\left(\varphi_1^L\right)^{(l)}(z), \qquad \left(\varphi_2^R\right)^{(l)}(z) = \left(\varphi_1^R\right)^{(l)}(z)D_l^R,\\
&l=0,1,\dots.
\end{aligned} \tag{13}$$

2.2.1. Biorthogonality

From the Gaussian factorization, whose existence is ensured for quasi-definite matrix measures, we infer that these matrix Laurent polynomials satisfy biorthogonal type relations.

Theorem 1. *The matrix Laurent polynomials* $\{(\varphi^H)_1^{(l)}\}_{l=0}^\infty$ *and* $\{(\varphi^H)_2^{(l)}\}_{l=0}^\infty$, $H=L,R$, *introduced in (11) and (12), are biorthogonal on the unit circle*

$$\left\langle\!\left\langle \left(\varphi_2^H\right)^{(j)}, \left(\varphi_1^H\right)^{(k)}\right\rangle\!\right\rangle_H = \mathbb{I}\delta_{j,k}, \quad H=L,R,\ j,k=0,1,\dots. \tag{14}$$

Proof. It is straightforward to check that

$$\oint_{\mathbb{T}} \phi_1^L(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\left(\phi_2^L(z)\right)^\dagger = S_1\left(\oint_{\mathbb{T}}\chi(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^\dagger\right)S_2^{-1} = S_1 g^L S_2^{-1} = \mathbb{I},$$

$$\oint_{\mathbb{T}} \left(\phi_2^R(z)\right)^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\phi_1^R(z) = Z_2^{-1}\left(\oint_{\mathbb{T}}\overline{\chi(z)}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^\top\right)Z_1 = Z_2^{-1}g^R Z_1 = \mathbb{I}. \qquad \square$$

In order to relate the CMV matrix Laurent polynomials to the Szegő polynomials, we rewrite the *quasi-orthogonality* conditions from Theorem 1

$$\begin{aligned}
&\oint_{\mathbb{T}} \big(\varphi_1^L\big)^{(2l)}(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-k}=0, \quad k=-l,\dots,l-1,\\
&\oint_{\mathbb{T}} \big(\varphi_1^L\big)^{(2l+1)}(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-k}=0, \quad k=-l,\dots,l,\\
&\oint_{\mathbb{T}} z^k\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\big[\big(\varphi_2^L\big)^{(2l)}(z)\big]^\dagger=0, \quad k=-l,\dots,l-1,\\
&\oint_{\mathbb{T}} z^k\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\big[\big(\varphi_2^L\big)^{(2l+1)}(z)\big]^\dagger=0, \quad k=-l,\dots,l, \qquad (15)\\
&\oint_{\mathbb{T}} \big[\big(\varphi_2^R\big)^{(2l)}(z)\big]^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{k}=0, \quad k=-l,\dots,l-1,\\
&\oint_{\mathbb{T}} \big[\big(\varphi_2^R\big)^{(2l+1)}(z)\big]^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{k}=0, \quad k=-l,\dots,l,\\
&\oint_{\mathbb{T}} z^{-k}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\big(\varphi_1^R\big)^{(2l)}(z)=0, \quad k=-l,\dots,l-1,\\
&\oint_{\mathbb{T}} z^{-k}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\big(\varphi_1^R\big)^{(2l+1)}(z)=0, \quad k=-l,\dots,l. \qquad (16)
\end{aligned}$$

Proposition 3. *For a quasi-definite matrix measure μ, the matrix Szegő polynomials and the CMV matrix Laurent polynomials are related in the following way for the left case*

$$\begin{aligned}
z^l\big(\varphi_1^L\big)^{(2l)}(z)&=P_{1,2l}^L(z),\\
z^{l+1}\big(\varphi_1^L\big)^{(2l+1)}(z)&=\big(P_{2,2l+1}^R\big)^*(z),\\
z^l\big(D_{2l}^L\big)^\dagger\big(\varphi_2^L\big)^{(2l)}(z)&=P_{2,2l}^L(z),\\
z^{l+1}\big(D_{2l+1}^L\big)^\dagger\big(\varphi_2^L\big)^{(2l+1)}(z)&=\big(P_{1,2l+1}^R\big)^*(z), \qquad (17)
\end{aligned}$$

and

$$\begin{aligned}
z^l\big(\varphi_2^R\big)^{(2l)}(z)&=P_{2,2l}^R(z),\\
z^{l+1}\big(\varphi_2^R\big)^{(2l+1)}(z)&=\big(P_{1,2l+1}^L\big)^*(z),\\
z^l\big(\varphi_1^R\big)^{(2l)}(z)D_{2l}^R&=P_{1,2l}^R(z),\\
z^{l+1}\big(\varphi_1^R\big)^{(2l+1)}(z)D_{2l+1}^R&=\big(P_{2,2l+1}^L\big)^*(z) \qquad (18)
\end{aligned}$$

for the right case.

Proof. Taking the differences between the RHS and LHS of the equalities, we get matrix polynomials, of degree $d = 2l - 1, 2l$, that when paired via $\langle\!\langle\cdot,\cdot\rangle\!\rangle_H$, $H = L, R$, to all the powers z^j, $j = 0, \ldots, q$ cancels. Therefore, as we have a quasi-definite matrix measure, with moment matrices having non-null principal block minors, the only possibility for the difference is to be 0. □

The last identifications together with (4) define some of the entries of the Gaussian factorization matrices.

Proposition 4. *The* matrix quasi-norms h_k^H *introduced in Definition 4 and the coefficients* D_k^H *given in (9) satisfy*

$$\begin{aligned} h_{2l}^L &= D_{2l}^L, & h_{2l+1}^L &= D_{2l+1}^R, \\ h_{2l}^R &= D_{2l}^R, & h_{2l+1}^R &= D_{2l+1}^L. \end{aligned}$$

For the first non-trivial block diagonal of the factors in the Gauss–Borel factorization, we get

Proposition 5. *The matrices of the block LU factorization can be written more explicitly in terms of the Verblunsky coefficients as follows*

$$S_1 = \begin{pmatrix} \mathbb{I} & 0 & 0 & 0 & 0 & \dots \\ [\alpha_{2,1}^R]^\dagger & \mathbb{I} & 0 & 0 & 0 & \dots \\ * & \alpha_{1,2}^L & \mathbb{I} & 0 & 0 & \dots \\ * & * & [\alpha_{2,3}^R]^\dagger & \mathbb{I} & 0 & \dots \\ * & * & * & \alpha_{1,4}^L & \mathbb{I} & \\ \vdots & \vdots & \vdots & & \ddots & \ddots \end{pmatrix},$$

$$\widehat{S_2}^{-1} = \begin{pmatrix} \mathbb{I} & \alpha_{1,1}^R & * & * & * & \dots \\ 0 & \mathbb{I} & [\alpha_{2,2}^L]^\dagger & * & * & \dots \\ 0 & 0 & \mathbb{I} & \alpha_{1,3}^R & * & \dots \\ 0 & 0 & 0 & \mathbb{I} & [\alpha_{2,4}^L]^\dagger & \\ 0 & 0 & 0 & 0 & \mathbb{I} & \ddots \\ \vdots & \vdots & \vdots & \vdots & & \ddots \end{pmatrix},$$

$$Z_2^{-1} = \begin{pmatrix} \mathbb{I} & 0 & 0 & 0 & 0 & \dots \\ \alpha_{1,1}^L & \mathbb{I} & 0 & 0 & 0 & \dots \\ * & [\alpha_{2,2}^R]^\dagger & \mathbb{I} & 0 & 0 & \dots \\ * & * & \alpha_{1,3}^L & \mathbb{I} & 0 & \dots \\ * & * & * & [\alpha_{2,4}^R]^\dagger & \mathbb{I} & \\ \vdots & \vdots & \vdots & & \ddots & \ddots \end{pmatrix},$$

$$\widehat{Z_1}=\begin{pmatrix}\mathbb{I} & [\alpha^L_{2,1}]^\dagger & * & * & * & \dots\\ 0 & \mathbb{I} & \alpha^R_{1,2} & * & * & \dots\\ 0 & 0 & \mathbb{I} & [\alpha^L_{2,3}]^\dagger & * & \dots\\ 0 & 0 & 0 & \mathbb{I} & \alpha^R_{1,4} & \\ \vdots & \vdots & \vdots & \vdots & \ddots & \ddots\end{pmatrix}.$$

This gives the following structure for the MOLPUC

Proposition 6.

(1) *The MOLPUC are of the form*

$$\begin{aligned}
(\phi^L_1)^{(2l)} &= \alpha^L_{1,2l}z^{-l}+\dots+z^l,\\
(\phi^L_1)^{(2l+1)} &= z^{-l-1}+\dots+(\alpha^R_{2,2l+1})^\dagger z^l,\\
(\phi^L_2)^{(2l)} &= \big((h^L_{2l})^\dagger\big)^{-1}\big(\alpha^L_{2,2l}z^{-l}+\dots+z^l\big),\\
(\phi^L_2)^{(2l+1)} &= \big((h^R_{2l+1})^\dagger\big)^{-1}\big(z^{-l-1}+\dots+(\alpha^R_{1,2l+1})^\dagger z^l\big),\\
(\phi^R_1)^{(2l)} &= \big(\alpha^R_{1,2l}z^{-l}+\dots+z^l\big)\big(h^R_{2l}\big)^{-1},\\
(\phi^R_1)^{(2l+1)} &= \big(z^{-l-1}+\dots+(\alpha^L_{2,2l+1})^\dagger z^l\big)\big(h^L_{2l+1}\big)^{-1},\\
(\phi^R_2)^{(2l)} &= \alpha^R_{2,2l}z^{-l}+\dots+z^l,\\
(\phi^R_2)^{(2l+1)} &= z^{-l-1}+\dots+\alpha^L_{1,2l+1}z^l.
\end{aligned}$$

(2) *The "quasi-norms" and the MOLPUC fulfill*

$$\begin{aligned}
h^R_{2l+1} &= \oint_{\mathbb{T}}(\phi^L_1)^{(2l+1)}(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{l+1}, & h^L_{2l} &= \oint_{\mathbb{T}}(\phi^L_1)^{(2l)}(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-l},\\
h^R_{2l} &= \oint_{\mathbb{T}}\big((\phi^R_2)^{(2l)}(z)\big)^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{l}, & h^L_{2l+1} &= \oint_{\mathbb{T}}\big((\phi^R_2)^{(2l+1)}(z)\big)^\dagger\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-l-1}.
\end{aligned} \tag{19}$$

Proof.

(1) Use (7), (8) and Propositions 4 and 5.
(2) Consider the biorthogonality (14) together with the explicit expressions of the first item in this proposition and orthogonality relations (15) and (16). □

Recalling (10), we conclude from Proposition 5 that in the Hermitian context, we have

$$\alpha^H_{1,l} = \alpha^H_{2,l}, \quad H = L, R, \ l = 0, 1, \ldots,$$
$$\left(D^H_l\right)^\dagger = D^H_l, \quad H = L, R, \ l = 0, 1, \ldots.$$

It is not difficult to see comparing the previous result with the proof of the Gaussian factorization (A.1) that in terms of Schur complements, we have

Proposition 7.

(1) *The matrices $D^H_l \in \mathbb{C}^{m\times m}$, $H = L, R$, $l = 0, 1, \ldots$, from the diagonal block of the bock LU factorization can be written as the following Schur complements*

$$D^H_l = \left(g^H\right)^{[l+1]} \big/ \left(g^H\right)^{[l]}, \quad H = L, R, \ l = 0, 1, \ldots. \tag{20}$$

(2) *The Verblunsky matrices can be expressed as*

$$\alpha^L_{1,2k} = -\sum_{i=0}^{2k-1} \left(g^L\right)_{2k,i} \left(\left(\left(g^L\right)^{[2k]}\right)^{-1}\right)_{i,2k-1},$$
$$\left[\alpha^R_{2,2k+1}\right]^\dagger = -\sum_{i=0}^{2k} \left(g^L\right)_{2k+1,i} \left(\left(\left(g^L\right)^{[2k+1]}\right)^{-1}\right)_{i,2k},$$
$$\left[\alpha^R_{2,2k}\right]^\dagger = -\sum_{i=0}^{2k-1} \left(g^R\right)_{2k,i} \left(\left(\left(g^R\right)^{[2k]}\right)^{-1}\right)_{i,2k-1},$$
$$\alpha^L_{1,2k+1} = -\sum_{i=0}^{2k} \left(g^R\right)_{2k+1,i} \left(\left(\left(g^R\right)^{[2k+1]}\right)^{-1}\right)_{i,2k},$$
$$\left[\alpha^L_{2,2k}\right]^\dagger = -\sum_{i=0}^{2k-1} \left(\left(\left(g^L\right)^{[2k]}\right)^{-1}\right)_{2k-1,i} \left(g^L\right)_{i,2k},$$
$$\alpha^R_{1,2k+1} = -\sum_{i=0}^{2k} \left(\left(\left(g^L\right)^{[2k+1]}\right)^{-1}\right)_{2k,i} \left(g^L\right)_{i,2k+1},$$
$$\alpha^R_{1,2k} = -\sum_{i=0}^{2k-1} \left(\left(\left(g^R\right)^{[2k]}\right)^{-1}\right)_{2k-1,i} \left(g^R\right)_{i,2k},$$
$$\left[\alpha^L_{2,2k+1}\right]^\dagger = -\sum_{i=0}^{2k} \left(\left(\left(g^R\right)^{[2k+1]}\right)^{-1}\right)_{2k,i} \left(g^R\right)_{i,2k+1}.$$

2.2.2. Alternative ways to express the CMV matrix Laurent polynomials

For later use, we now present some alternative expressions for the MOLPUC $(\varphi^H_i)^{(l)}(z)$, $H = L, R$, $l = 0, 1, \ldots$ in terms of Schur complements of bordered truncated matrices

Lemma 1. *The next expressions hold true*

$$
\begin{aligned}
\big(\varphi_1^L\big)^{(l)}(z) &= (S_2)_{ll}\begin{pmatrix}0 & 0 & \dots & 0 & \mathbb{I}\end{pmatrix}\big(\big(g^L\big)^{[l+1]}\big)^{-1}\chi^{[l+1]}\\
&= \chi^{(l)} - \begin{pmatrix}(g^L)_{l,0} & (g^L)_{l,1} & \cdots & (g^L)_{l,l-1}\end{pmatrix}\big(\big(g^L\big)^{[l]}\big)^{-1}\chi^{[l]}\\
&= \mathrm{SC}\left(\begin{array}{cccc|c}
(g^L)_{0,0} & (g^L)_{0,1} & \dots & (g^L)_{0,l-1} & \chi(z)^{(0)}\\
(g^L)_{1,0} & (g^L)_{1,1} & \dots & (g^L)_{1,l-1} & \chi(z)^{(1)}\\
 & \vdots & & & \vdots\\
(g^L)_{l-1,0} & (g^L)_{l-1,1} & \dots & (g^L)_{l-1,l-1} & \chi(z)^{(l-1)}\\
\hline
(g^L)_{l,0} & (g^L)_{l,1} & \dots & (g^L)_{l,l-1} & \chi(z)^{(l)}
\end{array}\right),
\end{aligned}
\tag{21}
$$

$$
\begin{aligned}
\big[\big(\varphi_2^L\big)^{(l)}(z)\big]^{\dagger} &= \big(\chi^{[l+1]}\big)^{\dagger}\big(\big(g^L\big)^{[l+1]}\big)^{-1}\begin{pmatrix}0\\0\\\vdots\\0\\\mathbb{I}\end{pmatrix}\\
&= \left(\big(\chi^{(l)}\big)^{\dagger} - \big(\chi^{[l]}\big)^{\dagger}\big((g_L)^{[l]}\big)^{-1}\begin{pmatrix}(g^L)_{0,l}\\(g^L)_{1,l}\\\vdots\\(g^L)_{l-1,l}\end{pmatrix}\right)\big(D^L\big)_l\\
&= \mathrm{SC}\left(\begin{array}{cccc|c}
(g^L)_{0,0} & (g^L)_{0,1} & \dots & (g^L)_{0,l-1} & (g^L)_{0,l}\\
(g^L)_{1,0} & (g^L)_{1,1} & \dots & (g^L)_{1,l-1} & (g^L)_{1,l}\\
 & \vdots & & & \vdots\\
(g^L)_{l-1,0} & (g^L)_{l-1,1} & \dots & (g^L)_{l-1,l-1} & (g^L)_{l-1,l}\\
\hline
[\chi(z)^{\dagger}]^{(0)} & [\chi(z)^{\dagger}]^{(1)} & \dots & [\chi(z)^{\dagger}]^{(l-1)} & [\chi(z)^{\dagger}]^{(l)}
\end{array}\right)\big(D^L\big)_l
\end{aligned}
\tag{22}
$$

and

$$
\begin{aligned}
\big(\varphi_1^R\big)^{(l)}(z) &= \big[\chi^{[l+1]}\big]^{\top}\big(\big(g^R\big)^{[l+1]}\big)^{-1}\begin{pmatrix}0\\0\\\vdots\\0\\\mathbb{I}\end{pmatrix}\\
&= \big[\chi^{(l)}\big]^{\top} - \big[\chi^{[l]}\big]^{\top}\big(\big(g^R\big)^{[l]}\big)^{-1}\begin{pmatrix}(g^R)_{0,l}\\(g^R)_{1,l}\\\vdots\\(g^R)_{l-1,l}\end{pmatrix}\big(\big(D^R\big)^{-1}\big)_l
\end{aligned}
$$

$$= \mathrm{SC}\left(\begin{array}{cccc|c} (g^R)_{0,0} & (g^R)_{0,1} & \cdots & (g^R)_{0,l-1} & (g^R)_{0,l} \\ (g^R)_{1,0} & (g^R)_{1,1} & \cdots & (g^R)_{1,l-1} & (g^R)_{1,l} \\ \vdots & \vdots & & \vdots & \vdots \\ (g^R)_{l-1,0} & (g^R)_{l-1,1} & \cdots & (g^R)_{l-1,l-1} & (g^R)_{l-1,l} \\ \hline [\chi(z)^\top]^{(0)} & [\chi(z)^\top]^{(1)} & \cdots & [\chi(z)^\top]^{(l-1)} & [\chi(z)^\top]^{(l)} \end{array}\right)\big((D^R)^{-1}\big)_l,$$

$$\big[(\varphi_2^R)^{(l)}(z)\big]^\dagger = (D^R)_l\,(0 \quad \ldots \quad 0 \quad \mathbb{I})\big((g^R)^{[l+1]}\big)^{-1}\big[(\chi^{[l+1]})^\top\big]^\dagger$$

$$= \big[(\chi^{(l)})^\top\big]^\dagger - \big((g^R)_{l,0} \quad \cdots \quad (g^R)_{l,l-1}\big)\big((g^R)^{[l]}\big)^{-1}\big[(\chi^{[l]})^\top\big]^\dagger$$

$$= \mathrm{SC}\left(\begin{array}{cccc|c} (g^R)_{0,0} & (g^R)_{0,1} & \ldots & (g^R)_{0,l-1} & [(\chi(z)^\top)^\dagger]^{(0)} \\ (g^R)_{1,0} & (g^R)_{1,1} & \ldots & (g^R)_{1,l-1} & [(\chi(z)^\top)^\dagger]^{(1)} \\ & & \vdots & & \vdots \\ (g^R)_{l-1,0} & (g^R)_{l-1,1} & \ldots & (g^R)_{l-1,l-1} & [(\chi(z)^\top)^\dagger]^{(l-1)} \\ \hline (g^R)_{l,0} & (g^R)_{l,1} & \ldots & (g^R)_{l,l-1} & [(\chi(z)^\top)^\dagger]^{(l)} \end{array}\right).$$

Proof. See Appendix A. □

Following [27] we give expressions in terms of Schur complements for the matrix Szegő polynomials, in terms of bordered truncated matrices of the right and left block CMV moment matrices, extending though similar expressions given in [27] in terms of standard block moment matrices.

Corollary 1. *The left matrix Szegő polynomials can be rewritten as the following Schur complements of bordered truncated CMV moment matrices*

$$P^L_{1,2l}(z) = z^l\,\mathrm{SC}\left(\begin{array}{ccc|c} & (g^L)^{[2l]} & & \chi(z)^{[2l]} \\ \hline (g^L)_{2l,0} & \ldots & (g^L)_{2l,2l-1} & \chi(z)^{(2l)} \end{array}\right),$$

$$P^L_{1,2l+1}(z) = z^{l+1}\,\mathrm{SC}\left(\begin{array}{ccc|c} & (g^R)^{[2l+1]} & & \chi^*(z)^{[2l+1]} \\ \hline (g^R)_{2l+1,0} & \ldots & (g^R)_{2l+1,2l-1} & \chi^*(z)^{(2l+1)} \end{array}\right),$$

$$\big[P^L_{2,2l}(z)\big]^\dagger = \bar{z}^l\,\mathrm{SC}\left(\begin{array}{c|c} (g^L)^{[2l]} & \begin{matrix}(g^L)_{0,2l}\\ \vdots \\ (g^L)_{2l-1,2l}\end{matrix} \\ \hline (\chi(z)^\dagger)^{[2l]} & (\chi(z)^\dagger)^{(2l)} \end{array}\right),$$

$$\big[P^L_{2,2l+1}(z)\big]^\dagger = \bar{z}^{l+1}\,\mathrm{SC}\left(\begin{array}{c|c} (g^R)^{[2l+1]} & \begin{matrix}(g^R)_{0,2l+1}\\ \vdots \\ (g^R)_{2l,2l+1}\end{matrix} \\ \hline (\chi^*(z)^\dagger)^{[2l+1]} & (\chi^*(z)^\dagger)^{(2l+1)} \end{array}\right),$$

while for the right polynomials we have

$$P^R_{1,2l}(z) = z^l \operatorname{SC}\left(\begin{array}{c|c} & (g^R)_{0,2l} \\ (g^R)^{[2l]} & \vdots \\ & (g^R)_{2l-1,2l} \\ \hline (\chi(z)^\top)^{[2l]} & (\chi(z)^\top)^{(2l)} \end{array}\right),$$

$$P^R_{1,2l+1}(z) = z^{l+1} \operatorname{SC}\left(\begin{array}{c|c} & (g^L)_{0,2l+1} \\ (g^L)^{[2l+1]} & \vdots \\ & (g^L)_{2l,2l+1} \\ \hline (\chi^*(z)^\top)^{[2l+1]} & (\chi^*(z)^\top)^{(2l+1)} \end{array}\right),$$

$$\left[P^R_{2,2l}(z)\right]^\dagger = \bar{z}^l \operatorname{SC}\left(\begin{array}{ccc|c} & (g^R)^{[2l]} & & (\chi(z)^\dagger)^{[2l]} \\ \hline (g^R)_{2l,0} & \dots & (g^R)_{2l,2l-1} & (\chi(z)^\dagger)^{(2l)} \end{array}\right),$$

$$\left[P^R_{2,2l+1}(z)\right]^\dagger = \bar{z}^{l+1} \operatorname{SC}\left(\begin{array}{ccc|c} & (g^L)^{[2l+1]} & & (\chi^*(z)^\dagger)^{[2l+1]} \\ \hline (g^R)_{2l+1,0} & \dots & (g^R)_{2l+1,2l-1} & (\chi^*(z)^\dagger)^{(2l+1)} \end{array}\right).$$

Proof. These relations appear when one introduces in (17) and (18) the expressions of the CMV polynomials in terms of Schur complements. □

2.3. Matrix second kind functions

The following matrix fashion of rewriting previous left objects

$$\begin{pmatrix} \phi^L_{1,1} & \phi^L_{1,2} \\ \phi^L_{2,1} & \phi^L_{2,2} \end{pmatrix} = \begin{bmatrix} S_1 & 0 \\ 0 & [S_2^{-1}]^\dagger \end{bmatrix} \begin{pmatrix} \chi_1 & \chi_2^* \\ \chi_1 & \chi_2^* \end{pmatrix} = \begin{pmatrix} S_1\chi_1 & S_1\chi_2^* \\ [S_2^{-1}]^\dagger\chi_1 & [S_2^{-1}]^\dagger\chi_2^* \end{pmatrix},$$

$$\begin{pmatrix} \phi^L_1 \\ \phi^L_2 \end{pmatrix} = \begin{pmatrix} \phi^L_{1,1} & \phi^L_{1,2} \\ \phi^L_{2,1} & \phi^L_{2,2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\begin{bmatrix} [g^L]^\dagger & 0 \\ 0 & g^L \end{bmatrix} = \begin{bmatrix} [S_2]^\dagger [S_1^{-1}]^\dagger & 0 \\ 0 & S_1^{-1} S_2 \end{bmatrix} = \begin{bmatrix} [S_2]^\dagger & 0 \\ 0 & S_1^{-1} \end{bmatrix} \begin{bmatrix} [S_1^{-1}]^\dagger & 0 \\ 0 & S_2 \end{bmatrix}$$

and the right ones

$$\begin{pmatrix} \phi^R_{1,1} & \phi^R_{2,1} \\ \phi^R_{1,2} & \phi^R_{2,2} \end{pmatrix} = \begin{pmatrix} \chi_1^\top & \chi_1^\top \\ [\chi_2^*]^\top & [\chi_2^*]^\top \end{pmatrix} \begin{bmatrix} Z_1 & 0 \\ 0 & [Z_2^{-1}]^\dagger \end{bmatrix},$$

$$\begin{pmatrix} \phi^R_1 & \phi^R_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} \phi^R_{1,1} & \phi^R_{2,1} \\ \phi^R_{1,2} & \phi^R_{2,2} \end{pmatrix},$$

$$\begin{bmatrix} [g^R]^\dagger & 0 \\ 0 & g^R \end{bmatrix} = \begin{bmatrix} [Z_1^{-1}]^\dagger & 0 \\ 0 & Z_2 \end{bmatrix} \begin{bmatrix} [Z_2]^\dagger & 0 \\ 0 & Z_1^{-1} \end{bmatrix}$$

inspires the next

Definition 7. The partial matrix CMV second kind sequences are given by

$$\begin{pmatrix} C^L_{1,1} & C^L_{1,2} \\ C^L_{2,1} & C^L_{2,2} \end{pmatrix} := \begin{bmatrix} [S_1^{-1}]^\dagger & 0 \\ 0 & S_2 \end{bmatrix} \begin{pmatrix} \chi_1^* & \chi_2 \\ \chi_1^* & \chi_2 \end{pmatrix} = \begin{pmatrix} [S_1^{-1}]^\dagger \chi_1^* & [S_1^{-1}]^\dagger \chi_2 \\ S_2 \chi_1^* & S_2 \chi_2 \end{pmatrix},$$

$$\begin{pmatrix} C^R_{1,1} & C^R_{2,1} \\ C^R_{1,2} & C^R_{2,2} \end{pmatrix} := \begin{pmatrix} [\chi_1^*]^\top & [\chi_1^*]^\top \\ \chi_2^\top & \chi_2^\top \end{pmatrix} \begin{bmatrix} [Z_1^{-1}]^\dagger & 0 \\ 0 & Z_2 \end{bmatrix} = \begin{pmatrix} [\chi_1^*]^\top [Z_1^{-1}]^\dagger & [\chi_1^*]^\top Z_2 \\ \chi_2^\top [Z_1^{-1}]^\dagger & \chi_2^\top Z_2 \end{pmatrix},$$

and the corresponding matrix CMV second kind sequences are

$$\begin{pmatrix} C^L_1 \\ C^L_2 \end{pmatrix} = \begin{pmatrix} C^L_{1,1} & C^L_{1,2} \\ C^L_{2,1} & C^L_{2,2} \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} C^R_1 & C^R_2 \end{pmatrix} = \begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} C^R_{1,1} & C^R_{2,1} \\ C^R_{1,2} & C^R_{2,2} \end{pmatrix}.$$

Complementary to the above definition

Definition 8. The associated CMV Fourier series are

$$\begin{pmatrix} \Gamma^L_{2,1} & \Gamma^L_{2,2} \\ \Gamma^L_{1,1} & \Gamma^L_{1,2} \end{pmatrix} := \begin{bmatrix} [g^L]^\dagger & 0 \\ 0 & g^L \end{bmatrix} \begin{pmatrix} \chi_1^* & \chi_2 \\ \chi_1^* & \chi_2 \end{pmatrix} = \begin{pmatrix} [g^L]^\dagger \chi_1^* & [g^L]^\dagger \chi_2 \\ g^L \chi_1^* & g^L \chi_2 \end{pmatrix},$$

$$\begin{pmatrix} \Gamma^R_{2,1} & \Gamma^R_{1,1} \\ \Gamma^R_{2,2} & \Gamma^R_{1,2} \end{pmatrix} := \begin{pmatrix} [\chi_1^*]^\top & [\chi_1^*]^\top \\ \chi_2^\top & \chi_2^\top \end{pmatrix} \begin{bmatrix} [g^R]^\dagger & 0 \\ 0 & g^R \end{bmatrix} = \begin{pmatrix} [\chi_1^*]^\top [g^R]^\dagger & [\chi_1^*]^\top g^R \\ \chi_2^\top [g^R]^\dagger & \chi_2^\top g^R \end{pmatrix},$$

for which we have:

Proposition 8.

(1) *The elements Γ^H and C^H, $H = L, R$, are related in the following way*

$$\begin{pmatrix} \Gamma^L_2 \\ \Gamma^L_1 \end{pmatrix} = \begin{bmatrix} [S_2]^\dagger & 0 \\ 0 & S_1^{-1} \end{bmatrix} \begin{pmatrix} C^L_1 \\ C^L_2 \end{pmatrix}, \qquad \begin{pmatrix} \Gamma^R_2 & \Gamma^R_1 \end{pmatrix} = \begin{pmatrix} C^R_1 & C^R_2 \end{pmatrix} \begin{bmatrix} [Z_2]^\dagger & 0 \\ 0 & Z_1^{-1} \end{bmatrix}.$$

(2) *The second kind functions can be expressed as Schur complements as follows*

$$\left[\left(C^L_1\right)^{(l)}(z)\right]^\dagger = \operatorname{SC}\left(\begin{array}{c|c} \multirow{3}{*}{$(g^L)^{[l]}$} & (g^L)_{0,l} \\ & \vdots \\ & (g^L)_{l-1,l} \\ \hline (\Gamma^L_2(z)^\dagger)^{[l]} & (\Gamma^L_2(z)^\dagger)^{(l)} \end{array}\right) \left(D_L^{-1}\right)_l,$$

$$\left(C^L_2\right)^{(l)}(z) = \operatorname{SC}\left(\begin{array}{ccc|c} & (g^L)^{[l]} & & (\Gamma_1(z))^{[l]} \\ \hline (g^L)_{l,0} & \dots & (g^L)_{l,l-1} & (\Gamma_1(z))^{(l)} \end{array}\right),$$

$$\big[(C_1^R)^{(l)}(z)\big]^\dagger = \mathrm{SC}\left(\begin{array}{ccc|c} & (g^R)^{[l]} & & (\Gamma_2^R(z)^\dagger)^{[l]} \\ \hline (g^R)_{l,0} & \dots & (g^R)_{l,l-1} & (\Gamma_2^R(z)^\dagger)^{(l)} \end{array}\right),$$

$$(C_2^R)^{(l)}(z) = \mathrm{SC}\left(\begin{array}{c|c} & (g^R)_{0,l} \\ (g^R)^{[l]} & \vdots \\ & (g^R)_{l-1,l} \\ \hline (\Gamma_1^R(z))^{[l]} & (\Gamma_1^R(z))^{(l)} \end{array}\right)(D_R^{-1})_l.$$

(3) *In terms of the matrix Laurent orthogonal polynomials and the Fourier series of the matrix measure, we have*

$$\begin{aligned}
(C_1^L)^{(l)}(z) &= 2\pi z^{-1}(\varphi_2^L)^{(l)}(z^{-1})F_\mu^\dagger(z),\\
(C_2^L)^{(l)}(z) &= 2\pi z^{-1}(\varphi_1^L)^{(l)}(z^{-1})F_\mu(z^{-1}),\\
(C_1^R)^{(l)}(z) &= 2\pi z^{-1}F_\mu^\dagger(z)(\varphi_2^R)^{(l)}(z^{-1}),\\
(C_2^R)^{(l)}(z) &= 2\pi z^{-1}F_\mu(z^{-1})(\varphi_1^R)^{(l)}(z^{-1}).
\end{aligned} \tag{23}$$

Proof. The first part of the proposition follows directly from comparison of the structure of the relations from the previous lemma with the definitions of the CMV matrix polynomials. For example

$$\Gamma_1^L = S_1^{-1}C_2^L \;\Rightarrow\; C_2^L = S_1\Gamma_1^L \text{ same structure as } \phi_1^L = S_1\chi \text{ replacing } \Gamma_1^L \longleftrightarrow \chi.$$

For the second part of the proposition, we shall only prove one of the cases since the rest of them can be proven following the same procedure. First, from the definition of the second kind functions, we have

$$\begin{aligned}
C_1^R(z) &= \big(\chi^*(z)\big)^\top\big[Z_1^{-1}\big]^\dagger\\
&= \big(\chi^*(z)\big)^\top\big[g^R\big]^\dagger\big(Z_2^{-1}\big)^\dagger\\
&= \big[\chi^*(z)\big]^\top\oint_{\mathbb{T}}\big[\chi^\top(u)\big]^\dagger\left[\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\right]^\dagger\big[\chi(u)\big]^\top Z_2^{-1}\\
&= \big[\chi^*(z)\big]^\top\oint_{\mathbb{T}}\big[\chi^\top(u)\big]^\dagger\left[\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\right]^\dagger\phi_2^R(u).
\end{aligned}$$

Taking the l-th component of this vector of matrices, we get

$$(C_1^R)^{(l)}(z) = \int_0^{2\pi}\sum_{n=-\infty} z^{n-1}\mathrm{e}^{\mathrm{i}n\theta}\big[\mathrm{d}\mu(\theta)\big]^\dagger(\varphi_2^R)^{(l)}(\mathrm{e}^{\mathrm{i}\theta})$$

$$= \sum_{k,n=-\infty} z^{n-1} \int_0^{2\pi} e^{i(n+k)\theta} \left[d\mu(\theta)\right]^\dagger \left(\varphi_{2,k}^R\right)^{(l)}$$
$$= 2\pi z^{-1} \left(\sum_{k=-\infty} \left(\varphi_{2,k}^R\right)^{(l)} z^{-k} \right) \left(\sum_{n=-\infty} c_{n+k}^\dagger z^{n+k} \right)$$
$$= 2\pi z^{-1} F_\mu^\dagger(z) \left(\varphi_2^R\right)^{(l)} \left(z^{-1}\right). \qquad \square$$

Recalling the previously stated relation between the Γ^H and the C^H, it follows from Proposition 8 that

Proposition 9. *The associated CMV Fourier series satisfy*

$$\Gamma_{1,j}^L = 2\pi z^{-1} F_\mu\left(z^{-1}\right) \chi^{(j)}\left(z^{-1}\right), \qquad \Gamma_{2,j}^L = 2\pi z^{-1} F_\mu^\dagger(z) \chi^{(j)}\left(z^{-1}\right),$$
$$\Gamma_{1,j}^R = 2\pi z^{-1} F_\mu\left(z^{-1}\right) \chi^{(j)}\left(z^{-1}\right), \qquad \Gamma_{2,j}^R = 2\pi z^{-1} F_\mu^\dagger(z) \chi^{(j)}\left(z^{-1}\right).$$

Another interesting representation of these functions is

Proposition 10. *The second kind functions have the following Cauchy integral type formulae*

$$\left[C_{1,1}^L(z)\right]^\dagger = \oint_{\mathbb{T}} \left[z^{-1} \frac{u}{u - z^{-1}}\right]^\dagger \frac{d\mu(u)}{iu} \left[\phi_2^L(u)\right]^\dagger,$$
$$C_{2,1}^L(z) = \oint_{\mathbb{T}} \phi_1^L(u) \frac{d\mu(u)}{iu} \left[z^{-1} \frac{u}{u - z^{-1}}\right], \quad |z| > 1,$$
$$\left[C_{1,2}^L(z)\right]^\dagger = \oint_{\mathbb{T}} \left[-z^{-1} \frac{u}{u - z^{-1}}\right]^\dagger \frac{d\mu(u)}{iu} \left[\phi_2^L(u)\right]^\dagger,$$
$$C_{2,2}^L(z) = \oint_{\mathbb{T}} \phi_1^L(u) \frac{d\mu(u)}{iu} \left[-z^{-1} \frac{u}{u - z^{-1}}\right], \quad |z| < 1,$$
$$\left[C_{1,1}^R(z)\right]^\dagger = \oint_{\mathbb{T}} \left[\phi_2^R(u)\right]^\dagger \frac{d\mu(u)}{iu} \left[z^{-1} \frac{u}{u - z^{-1}}\right]^\dagger,$$
$$\left[C_{2,1}^R(z)\right] = \oint_{\mathbb{T}} \left[z^{-1} \frac{u}{u - z^{-1}}\right] \frac{d\mu(u)}{iu} \phi_1^R(u), \quad |z| > 1,$$
$$\left[C_{1,2}^R(z)\right]^\dagger = \oint_{\mathbb{T}} \left[\phi_2^R(u)\right]^\dagger \frac{d\mu(u)}{iu} \left[-z^{-1} \frac{u}{u - z^{-1}}\right]^\dagger,$$
$$\left[C_{2,1}^R(z)\right] = \oint_{\mathbb{T}} \left[-z^{-1} \frac{u}{u - z^{-1}}\right] \frac{d\mu(u)}{iu} \phi_1^R(u), \quad |z| < 1.$$

Proof. Direct substitution leads to

$$\big[C^L_{1,1}\big](z)^\dagger = \oint_{\mathbb{T}} \Big[\sum_{n=0} z^{-1}(uz)^{-n}\Big]^\dagger \frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\big[\phi^L_2(u)\big]^\dagger,$$

$$C^L_{2,1}(z) = \oint_{\mathbb{T}} \phi^L_1(u)\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\Big[\sum_{n=0} z^{-1}(uz)^{-n}\Big],$$

$$\big[C^L_{1,2}(z)\big]^\dagger = \oint_{\mathbb{T}} \Big[\sum_{n=0} u(uz)^{n}\Big]^\dagger \frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\big[\phi^L_2(u)\big]^\dagger,$$

$$C^L_{2,2}(z) = \oint_{\mathbb{T}} \phi^L_1(u)\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\Big[\sum_{n=0} u(uz)^{n}\Big],$$

$$\big[C^R_{1,1}(z)\big]^\dagger = \oint_{\mathbb{T}} \big[\phi^R_2(u)\big]^\dagger\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\Big[\sum_{n=0} z^{-1}(uz)^{-n}\Big]^\dagger,$$

$$C^R_{2,1}(z) = \oint_{\mathbb{T}} \Big[\sum_{n=0} z^{-1}(uz)^{-n}\Big]\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\phi^R_1(u),$$

$$\big[C^R_{1,2}(z)\big]^\dagger = \oint_{\mathbb{T}} \big[\phi^R_2(u)\big]^\dagger\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\Big[\sum_{n=0} u(uz)^{n}\Big]^\dagger,$$

$$C^R_{2,1}(z) = \oint_{\mathbb{T}} \Big[\sum_{n=0} u(uz)^{-n}\Big]\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}\phi^R_1(u).$$

But these are the series expansions of the functions of the proposition. We will not deal here with convergence problems since their discussion follows the ideas of [13]. □

2.4. Recursion relations

In order to get the recursion relations we introduce the following

Definition 9. For each pair $i, j \in \mathbb{Z}_+$, we consider the block semi-infinite matrix $E_{i,j}$ whose only non-zero $m \times m$ block is the (i, j)-th block where the identity of $\mathbb{M}_m$ appears. Then, we define the projectors

$$\Pi_1 := \sum_{j=0}^{\infty} E_{2j,2j}, \qquad \Pi_2 := \sum_{j=0}^{\infty} E_{2j+1,2j+1},$$

and the following matrices

$$\Lambda_1 := \sum_{j=0}^{\infty} E_{2j,2+2j}, \qquad \Lambda_2 := \sum_{j=0}^{\infty} E_{1+2j,3+2j}, \qquad \Lambda := \sum_{j=0}^{\infty} E_{j,j+1},$$
$$\Upsilon := \Lambda_1 + \Lambda_2^\top + E_{1,1}\Lambda^\top.$$

The matrix Υ, which can be written more explicitly as follows

$$\Upsilon=\left(\begin{array}{cc|cc|cc|cc|ccc}
0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
\mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
\hline
0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & \cdots\\
0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & \cdots\\
0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & \cdots\\
0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & \cdots\\
\hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots\\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & \cdots\\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right),$$

satisfies

$$\Upsilon^{\dagger}=\Upsilon^{-1}=\Upsilon^{\top}$$

and has the following properties

Proposition 11. *The next* eigen-value *type relations hold true*

$$\Upsilon\chi(z)=z\chi(z), \qquad \Upsilon^{-1}\chi(z)=z^{-1}\chi(z), \tag{24}$$
$$\chi(z)^{\top}\Upsilon^{-1}=z\chi(z)^{\top}, \qquad \chi(z)^{\top}\Upsilon=z^{-1}\chi(z)^{\top}. \tag{25}$$

Proof. It follows from the relations

$$\begin{aligned}
&\Lambda_1\chi(z)=z\Pi_1\chi(z), && \Lambda_2\chi(z)=z^{-1}\Pi_2\chi(z),\\
&\Lambda_1^{\top}\chi(z)=\bigl(z^{-1}\Pi_1-E_{0,0}\Lambda\bigr)\chi(z), && \Lambda_2^{\top}\chi(z)=\bigl(z\Pi_2-E_{1,1}\Lambda^{\top}\bigr)\chi(z). \qquad \square
\end{aligned}$$

From these, the following *symmetry* relations are obtained

Proposition 12. *The moment matrices commute with Υ; i.e.,*

$$\Upsilon g^{H}=g^{H}\Upsilon, \quad H=L,R. \tag{26}$$

Proof. It is a consequence of

$$\begin{aligned}
\Upsilon g^{L}&=\oint_{\mathbb{T}} z\chi(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^{\dagger}=\oint_{\mathbb{T}}\chi(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\bigl(z^{-1}\chi(z)\bigr)^{\dagger}=g^{L}\Upsilon,\\
\Upsilon g^{R}&=\oint_{\mathbb{T}}\overline{z\chi(z)}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^{\top}=\oint_{\mathbb{T}}\overline{\chi(z)}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}z^{-1}\chi(z)^{\top}=g^{R}\Upsilon. \qquad \square
\end{aligned}$$

We now introduce another important matrix in the CMV theory

Definition 10. The intertwining matrix η is

$$\eta := \left(\begin{array}{c|cc|cc|cc c}
\mathbb{I} & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & \mathbb{I} & 0 & 0 & 0 & 0 & \cdots \\
0 & \mathbb{I} & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & \mathbb{I} & 0 & 0 & \cdots \\
0 & 0 & 0 & \mathbb{I} & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & \mathbb{I} & \cdots \\
0 & 0 & 0 & 0 & 0 & \mathbb{I} & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right),$$

which, as the reader can easily check, has the following properties

$$\eta^{-1} = \eta, \qquad \eta\chi(z) = \chi\left(z^{-1}\right), \qquad \chi(z)^{\top}\eta = \chi\left(z^{-1}\right)^{\top}.$$

When $z \in \mathbb{T}$, we have that $\eta\chi = \bar{\chi}$ and $\chi^{\top}\eta = \chi^{\dagger}$ which lead to the intertwining property

Proposition 13. *The left and right moment matrices satisfy the intertwining type property*

$$\eta g^R = g^L \eta.$$

Proof. It is straightforward to realize that

$$\begin{aligned}
\eta g^L \eta &= \oint_{\mathbb{T}} \eta\chi(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^{\dagger}\eta = \oint_{\mathbb{T}} \chi(\bar{z})\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi\left(\bar{z}^{-1}\right)^{\top} \\
&= \oint_{\mathbb{T}} \overline{\chi(z)}\frac{\mathrm{d}\mu(z)}{\mathrm{i}z}\chi(z)^{\top} = g^R. \qquad \square
\end{aligned}$$

Proposition 14. *The matrices Υ and η are related by*

$$\eta\Upsilon = \Upsilon^{-1}\eta.$$

Now we proceed to the *dressing* of Υ and η. We first notice that

Proposition 15. *The following equations hold*

$$\begin{aligned}
S_1\Upsilon S_1^{-1} &= S_2\Upsilon S_2^{-1}, \\
Z_1^{-1}\Upsilon Z_1 &= Z_2^{-1}\Upsilon Z_2, \\
Z_2^{-1}\eta\Upsilon^p S_1^{-1} &= Z_1^{-1}\eta\Upsilon^p S_2^{-1}, \quad p \in \mathbb{Z}.
\end{aligned}$$

Those equations allow us to define

Definition 11. Let us define

$$J^L := S_1 \Upsilon S_1^{-1} = S_2 \Upsilon S_2^{-1}, \qquad J^R := Z_1^{-1} \Upsilon Z_1 = Z_2^{-1} \Upsilon Z_2, \tag{27}$$

and for any $p \in \mathbb{Z}$, introduce

$$C_{[p]} = Z_2^{-1} \eta \Upsilon^p S_1^{-1} = Z_1^{-1} \eta \Upsilon^p S_2^{-1}. \tag{28}$$

Observations.

(1)

$$C_{[-|p|]} = Z_2^{-1} \eta \Upsilon^{-|p|} S_1^{-1} = Z_2^{-1} \Upsilon^{|p|} \eta S_1^{-1}.$$

(2) In the Hermitian case,

$$\begin{aligned} &C_{[0]} = Z_2^{-1} \eta S_1^{-1} = Z_1^{-1} \eta S_2^{-1} = D^R \widehat{Z_1}^{-1} \eta \widehat{S_2}^{-1} D^L = D^R Z_2^{\dagger} \eta S_1^{\dagger} D^L \\ &\implies \quad C_{[0]}^{\dagger} = D^L C_{[0]}^{-1} D^R. \end{aligned}$$

Proposition 16. *Powers of J_H can be expressed as follows*

$$\left(J^R\right)^{l-p} = C_{[p]} [C_{[l]}]^{-1}, \qquad \left(J^L\right)^{p-l} = [C_{[l]}]^{-1} C_{[p]}.$$

Now we give the schematic shape of some of these matrices

$$J^H = \left(\begin{array}{cc|cc|cc|cc|cc|c}
* & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
 & * & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & * & * & * & * & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & * & * & * & * & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & * & * & * & * & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & * & * & * & * & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & * & * & * & * & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & * & * & * & * & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & * & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right), \qquad \left(J^H\right)^{-1} = \left(\begin{array}{cc|cc|cc|cc|cc|c}
* & * & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
 & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ \hline
* & * & * & * & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & * & * & * & * & 0 & 0 & 0 & 0 & \cdots \\ \hline
0 & 0 & * & * & * & * & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & * & * & * & * & 0 & 0 & \cdots \\ \hline
0 & 0 & 0 & 0 & * & * & * & * & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & \cdots \\ \hline
0 & 0 & 0 & 0 & 0 & 0 & * & * & * & * & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & * & * & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right),$$

$$
C_{[0]} = \left(\begin{array}{c|cc|cc|ccc}
* & * & 0 & 0 & 0 & 0 & 0 & \cdots \\
\hline
* & * & * & 0 & 0 & 0 & 0 & \cdots \\
0 & * & * & * & 0 & 0 & 0 & \cdots \\
\hline
0 & 0 & * & * & * & 0 & 0 & \cdots \\
0 & 0 & 0 & * & * & * & 0 & \cdots \\
\hline
0 & 0 & 0 & 0 & * & * & * & \cdots \\
0 & 0 & 0 & 0 & 0 & * & * & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right), \qquad
C_{[-1]} = \left(\begin{array}{c|cc|cc|ccc}
* & * & 0 & 0 & 0 & 0 & 0 & \cdots \\
\hline
* & * & * & 0 & 0 & 0 & 0 & \cdots \\
0 & * & * & * & 0 & 0 & 0 & \cdots \\
\hline
0 & 0 & * & * & * & 0 & 0 & \cdots \\
0 & 0 & 0 & * & * & * & 0 & \cdots \\
\hline
0 & 0 & 0 & 0 & * & * & * & \cdots \\
0 & 0 & 0 & 0 & 0 & * & * & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{array}\right).
$$

Here the $*$ are non-zero $m \times m$ blocks that, thanks to the factorization problem, can be written in terms of the Verblunsky coefficients as we will see later. The shape of each matrix is a consequence of the two possible definitions (in terms of upper or lower block-triangular matrices). For the explicit form of these matrices, see Appendix B.

A first consequence is the following relations among Verblunsky coefficients and the *matrix quasi-norms* of the Szegő polynomials

Proposition 17. *The following relations are fulfilled*

$$
\begin{aligned}
h_m^R\left[\alpha_{2,m}^L\right]^\dagger &= \left[\alpha_{2,m}^R\right]^\dagger h_m^L, & h_m^L\alpha_{1,m+1}^R &= \alpha_{1,m+1}^L h_m^R, \\
\alpha_{1,m}^L h_m^R &= h_m^L \alpha_{1,m}^R, & \left[\alpha_{2,m+1}^R\right]^\dagger h_m^L &= h_m^R\left[\alpha_{2,m+1}^L\right]^\dagger, \\
h_k^L &= \left(\mathbb{I} - \alpha_{1,k}^L\left[\alpha_{2,k}^R\right]^\dagger\right)h_{k-1}^L, & h_k^R &= \left(\mathbb{I} - \left[\alpha_{2,k}^R\right]^\dagger\alpha_{1,k}^L\right)h_{k-1}^R, \\
h_k^L &= h_{k-1}^L\left(\mathbb{I} - \alpha_{1,k}^R\left[\alpha_{2,k}^L\right]^\dagger\right), & h_k^R &= h_{k-1}^R\left(\mathbb{I} - \left[\alpha_{2,k}^L\right]^\dagger\alpha_{1,k}^R\right),
\end{aligned}
$$

Proof. Just compare the two possible definitions of $C_{[0]}^{\pm 1}$ and $C_{[-1]}^{\pm 1}$. $\square$

Notice that the two relations in each column coincide in the Hermitian case.

Proposition 18. *The next* eigen-value *properties hold*

$$
\begin{aligned}
J^L\Phi_1^L &= z\Phi_1^L, & \left(J^L\right)^{-1}\Phi_1^L &= z^{-1}\Phi_1^L, \\
\left[J^L\right]^\dagger\Phi_2^L &= z^{-1}\Phi_2^L, & \left(\left[J^L\right]^\dagger\right)^{-1}\Phi_2^L &= z\Phi_2^L, \\
\Phi_2^R\left[J^R\right]^\dagger &= z\Phi_2^R, & \Phi_2^R\left(\left[J^R\right]^\dagger\right)^{-1} &= z^{-1}\Phi_2^R, \\
\Phi_1^R J^R &= z^{-1}\Phi_1^R, & \Phi_1^R\left(J^R\right)^{-1} &= z\Phi_1^R,
\end{aligned}
$$

and the following properties are fulfilled

$$
C_{[p]}\Phi_1^L(z) = z^p\left(\Phi_2^R\left(\bar{z}^{-1}\right)\right)^\dagger, \qquad \Phi_1^R(z)C_{[p]} = z^p\left(\Phi_2^L\left(\bar{z}^{-1}\right)\right)^\dagger.
$$

Proof. The results follow directly from the action of $\Upsilon^{\pm 1}$ and η on χ and the definitions of $J^H, C_{[p]}$ and Φ_i^H. For example

$$\begin{aligned}
J^L \Phi_1^L &= S_1 \Upsilon S_1^{-1} S_1 \chi(z) = S_1 \Upsilon \chi(z) = z S_1 \chi(z) = z \Phi_1^L, \\
C_{[p]} \Phi_1^L &= Z_2^{-1} \eta \Upsilon^p S_1^{-1} S_1 \chi(z) = Z_2^{-1} \eta z^p \chi(z) = z^p Z_2^{-1} \chi\big(z^{-1}\big) = z^p \big(\Phi_2^R\big(\bar{z}^{-1}\big)\big)^{\dagger}.
\end{aligned}$$

For the remaining relations, one proceeds in a similar way. □

This last proposition implies

Proposition 19. *The following recursion relations for the left Laurent polynomials hold*

$$\begin{aligned}
z\big(\varphi_1^L\big)^{(2k)} &= -\alpha_{1,2k+1}^L \big(\mathbb{I} - \big[\alpha_{2,2k}^R\big]^{\dagger} \alpha_{1,2k}^L\big) \big(\varphi_1^L\big)^{(2k-1)} - \alpha_{1,2k+1}^L \big[\alpha_{2,2k}^R\big]^{\dagger} \big(\varphi_1^L\big)^{(2k)} \\
&\quad - \alpha_{1,2k+2}^L \big(\varphi_1^L\big)^{(2k+1)} + \big(\varphi_1^L\big)^{(2k+2)}, \\
z\big(\varphi_1^L\big)^{(2k+1)} &= \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^{\dagger} \alpha_{1,2k+1}^L\big) \big(\mathbb{I} - \big[\alpha_{2,2k}^R\big]^{\dagger} \alpha_{1,2k}^L\big) \big(\varphi_1^L\big)^{(2k-1)} \\
&\quad + \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^{\dagger} \alpha_{1,2k+1}^L\big) \big[\alpha_{2,2k}^R\big]^{\dagger} \big(\varphi_1^L\big)^{(2k)} \\
&\quad - \big[\alpha_{2,2k+1}^R\big]^{\dagger} \alpha_{1,2k+2}^L \big(\varphi_1^L\big)^{(2k+1)} + \big[\alpha_{2,2k+1}^R\big]^{\dagger} \big(\varphi_1^L\big)^{(2k+2)}, \\
z\big(\varphi_2^L\big)^{(2k)}(z) &= -\alpha_{2,2k+1}^R \big(\varphi_2^L\big)^{(2k-1)}(z) - \alpha_{2,2k+1}^R \big[\alpha_{1,2k}^L\big]^{\dagger} \big(\varphi_2^L\big)^{(2k)} \\
&\quad - \big(\mathbb{I} - \alpha_{2,2k+1}^R \big[\alpha_{1,2k+1}^L\big]^{\dagger}\big) \big(\varphi_2^L\big)^{(2k+1)}(z) \\
&\quad + \big(\mathbb{I} - \alpha_{2,2k+1}^R \big[\alpha_{1,2k+1}^L\big]^{\dagger}\big) \big(\mathbb{I} - \alpha_{2,2k+2}^R \big[\alpha_{1,2k+2}^L\big]^{\dagger}\big) \big(\varphi_2^L\big)^{(2k+2)}(z), \\
z\big(\varphi_2^L\big)^{(2k+1)}(z) &= \big(\varphi_2^L\big)^{(2k-1)}(z) + \big[\alpha_{1,2k}^L\big]^{\dagger} \big(\varphi_2^L\big)^{(2k)}(z) - \big[\alpha_{1,2k+1}^L\big]^{\dagger} \alpha_{2,2k+2}^R \big(\varphi_2^L\big)^{(2k+1)}(z) \\
&\quad + \big[\alpha_{1,2k+1}^L\big]^{\dagger} \big(\mathbb{I} - \alpha_{2,2k+2}^R \big[\alpha_{1,2k+2}^L\big]^{\dagger}\big) \big(\varphi_2^L\big)^{(2k+2)}(z),
\end{aligned}$$

while for the right polynomials, relations are

$$\begin{aligned}
z\big(\varphi_1^R\big)^{(2k)} &= -\big(\varphi_1^R\big)^{(2k-1)} \alpha_{1,2k+1}^L - \big(\varphi_1^R\big)^{(2k)} \big[\alpha_{2,2k}^R\big]^{\dagger} \alpha_{1,2k+1}^L \\
&\quad - \big(\varphi_1^R\big)^{(2k+1)} \alpha_{1,2k+2}^L \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^{\dagger} \alpha_{1,2k+1}^L\big) \\
&\quad + \big(\varphi_1^R\big)^{(2k+2)} \big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^{\dagger} \alpha_{1,2k+2}^L\big) \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^{\dagger} \alpha_{1,2k+1}^L\big), \\
z\big(\varphi_1^R\big)^{(2k+1)} &= \big(\varphi_1^R\big)^{(2k-1)} + \big(\varphi_1^R\big)^{(2k)} \big[\alpha_{2,2k}^R\big]^{\dagger} - \big(\varphi_1^R\big)^{(2k+1)} \alpha_{1,2k+2}^L \big[\alpha_{2,2k+1}^R\big]^{\dagger} \\
&\quad + \big(\varphi_1^R\big)^{(2k+2)} \big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^{\dagger} \alpha_{1,2k+2}^L\big) \big[\alpha_{2,2k+1}^R\big]^{\dagger}, \\
z\big(\varphi_2^R\big)^{(2k)} &= -\big(\varphi_2^R\big)^{(2k-1)} \big(\mathbb{I} - \alpha_{2,2k}^R \big[\alpha_{1,2k}^L\big]^{\dagger}\big) \alpha_{2,2k+1}^R - \big(\varphi_2^R\big)^{(2k)} \big[\alpha_{1,2k}^L\big]^{\dagger} \alpha_{2,2k+1}^R \\
&\quad - \big(\varphi_2^R\big)^{(2k+1)} \alpha_{2,2k+2}^R + \big(\varphi_2^R\big)^{(2k+2)}, \\
z\big(\varphi_2^R\big)^{(2k+1)} &= \big(\varphi_2^R\big)^{(2k-1)} \big(\mathbb{I} - \alpha_{2,2k}^R \big[\alpha_{1,2k}^L\big]^{\dagger}\big) \big(\mathbb{I} - \alpha_{2,2k+1}^R \big[\alpha_{1,2k+1}^L\big]^{\dagger}\big)
\end{aligned}$$

$$+\left(\varphi_2^R\right)^{(2k)}\left[\alpha_{1,2k}^L\right]^\dagger\left(\mathbb{I}-\alpha_{2,2k+1}^R\left[\alpha_{1,2k+1}^L\right]^\dagger\right)$$
$$-\left(\varphi_2^R\right)^{(2k+1)}\alpha_{2,2k+1}^R\left[\alpha_{1,2k+1}^L\right]^\dagger+\left(\varphi_2^R\right)^{(2k+2)}\left[\alpha_{1,2k+1}^L\right]^\dagger.$$

We have written down just the recursion relations for z and not those for z^{-1}, which can be derived similarly to these ones. For the complete recursion expressions, see Appendix C.

Proposition 20. *The following relations hold true*

$$\begin{aligned}
&\left(\left(\varphi_2^R\right)^{(2k)}\left(\bar z^{-1}\right)\right)^\dagger=\left(\mathbb{I}-\left[\alpha_{2,2k}^R\right]^\dagger\alpha_{1,2k}^L\right)\left(\varphi_1^L\right)^{(2k-1)}(z)+\left[\alpha_{2,2k}^R\right]^\dagger\left(\varphi_1^L\right)^{(2k)}(z),\\
&\left(\left(\varphi_2^R\right)^{(2k+1)}\left(\bar z^{-1}\right)\right)^\dagger=-\alpha_{1,2k+2}^L\left(\varphi_1^L\right)^{(2k+1)}(z)+\left(\varphi_1^L\right)^{(2k+2)}(z),\\
&\left(\left(\varphi_2^L\right)^{(2k)}\left(\bar z^{-1}\right)\right)^\dagger=\left(\varphi_1^R\right)^{(2k-1)}(z)+\left(\varphi_1^R\right)^{(2k)}(z)\left[\alpha_{2,2k}^R\right]^\dagger,\\
&\left(\left(\varphi_2^L\right)^{(2k+1)}\left(\bar z^{-1}\right)\right)^\dagger\\
&\quad=-\left(\varphi_1^R\right)^{(2k+1)}(z)\alpha_{1,2k+2}^L+\left(\varphi_1^R\right)^{(2k+2)}(z)\left(\mathbb{I}-\left[\alpha_{2,2k+2}^R\right]^\dagger\alpha_{1,2k+2}^L\right),\\
&\frac{1}{z}\left(\left(\varphi_2^R\right)^{(2k)}\left(\bar z^{-1}\right)\right)^\dagger=-\left[\alpha_{2,2k+1}^R\right]^\dagger\left(\varphi_1^L\right)^{(2k)}(z)+\left(\varphi_1^L\right)^{(2k+1)}(z),\\
&\frac{1}{z}\left(\left(\varphi_2^R\right)^{(2k+1)}\left(\bar z^{-1}\right)\right)^\dagger\\
&\quad=\left(\mathbb{I}-\alpha_{1,2k+1}^L\left[\alpha_{2,2k+1}^R\right]^\dagger\right)\left(\varphi_1^L\right)^{(2k)}(z)+\alpha_{1,2k+1}^L\left(\varphi_1^L\right)^{(2k+1)}(z),\\
&\frac{1}{z}\left(\left(\varphi_2^L\right)^{(2k+1)}\left(\bar z^{-1}\right)\right)^\dagger=\left(\varphi_1^R\right)^{(2k)}(z)+\left(\varphi_1^R\right)^{(2k)}(z)\alpha_{1,2k+1}^L,\\
&\frac{1}{z}\left(\left(\varphi_2^L\right)^{(2k)}\left(\bar z^{-1}\right)\right)^\dagger\\
&\quad=-\left(\varphi_1^R\right)^{(2k)}(z)\left[\alpha_{2,2k+1}^R\right]^\dagger+\left(\varphi_1^R\right)^{(2k+1)}(z)\left(\mathbb{I}-\alpha_{1,2k+1}^L\left[\alpha_{2,2k+1}^R\right]^\dagger\right).
\end{aligned}$$

Proof. These relations appear just by substituting into (18) the expressions of the blocks of $(J_H)^{\pm1}$, $C_{[0]}$, $C_{[-1]}$. □

Using Proposition 19 and the matrix CMV recursion relations in Proposition 20, one derives the recursion relations for the matrix Szegő polynomials:

$$\begin{aligned}
zP_{1,2l+1}^L(z)-P_{1,2l+2}^L(z)&=-\alpha_{1,2l+2}^L\left(P_{2,2l+1}^R(z)\right)^*\\
\left(P_{2,2l}^R(z)\right)^*-\left(\mathbb{I}-\left(\alpha_{2,2l}^R\right)^\dagger\alpha_{1,2l}^L\right)\left(P_{2,2l-1}^R(z)\right)^*&=\left(\alpha_{2,2l}^R\right)^\dagger P_{1,2l}^L(z)\\
\left(P_{2,2l}^L(z)\right)^*-\left(P_{2,2l-1}^L(z)\right)^*\left(\mathbb{I}-\alpha_{1,2l}^R\left(\alpha_{2,2l}^L\right)^\dagger\right)&=P_{1,2l}^R(z)\left(\alpha_{2,2l}^L\right)^\dagger\\
zP_{1,2l+1}^R(z)-P_{1,2l+2}^R(z)&=-\left(P_{2,2l+1}^L(z)\right)^*\alpha_{1,2l+2}^L\\
\left(P_{2,2l+1}^R(z)\right)^*-\left(P_{2,2l}^R(z)\right)^*&=\left(\alpha_{2,2l+1}^R\right)^\dagger zP_{1,2l}^L(z)
\end{aligned}$$

$$\begin{aligned}
P^L_{1,2l+1}(z) - \big(\mathbb{I} - \alpha^L_{1,2l+1}\big(\alpha^R_{2,2l+1}\big)^\dagger\big) z P^L_{1,2l}(z) &= \alpha^L_{1,2l+1}\big(P^R_{2,2l+1}(z)\big)^* \\
P^R_{1,2l+1}(z) - z P^R_{1,2l}(z)\big(\mathbb{I} - \big(\alpha^L_{2,2l+1}\big)^\dagger \alpha^R_{1,2l+1}\big) &= \big(P^L_{2,2l+1}(z)\big)^* \alpha^R_{1,2l+1} \\
\big(P^L_{2,2l}(z)\big)^* - \big(P^L_{2,2l+1}(z)\big)^* &= -z P^R_{1,2l}(z)\big(\alpha^L_{2,2l+1}\big)^\dagger
\end{aligned}$$

which after the prescription

$$\begin{aligned}
x^l_N &:= \alpha^L_{1,N} & x^r_N &:= \alpha^R_{1,N} \\
y^l_N &:= \big(\alpha^L_{2,N}\big)^\dagger & y^r_N &:= \big(\alpha^R_{2,N}\big)^\dagger
\end{aligned}$$

coincide with the formulae in [27].

2.5. *Christoffel–Darboux theory*

To conclude this section, we show how the Gaussian factorization leads to the Christoffel–Darboux theorem for the matrix Laurent polynomials on the unit circle context. In this particular situation we must consider two different cases. As we are working in a non-Abelian situation, we first have projections in the corresponding modules, "orthogonal" in the ring (our blocks) context. Secondly, when the matrix measures are Hermitian and positive definite, we will have a scalar product, and the projections to consider are orthogonal indeed.

2.5.1. *Projections in modules*

Given a right or left $\mathbb{M}_m$ module M, any idempotent endomorphism $\pi \in \operatorname{End}_{\mathbb{M}_m}(M)$, $\pi^2 = \pi$, is called a projection. For any given projection π, we have $\operatorname{Ker}\pi = \operatorname{Im}(1-\pi)$, $\operatorname{Ker}(1-\pi) = \operatorname{Im}\pi$, and the following direct decomposition holds: $M = \operatorname{Im}\pi \oplus \operatorname{Im}(1-\pi)$. Two projections π and π' are said to be orthogonal if $\pi\pi' = 0$; observe that $(1-\pi)$ is idempotent and moreover orthogonal to π. Orthogonality is not related here to any inner product so far, it is just a construction in the module. In particular, in our discussion of matrix Laurent polynomials, we introduce the following free modules

$$\Lambda_{[l]} := \mathbb{M}_m\big\{\chi^{(j)}\big\}_{j=0}^{l} = \begin{cases} \Lambda_{m,[-k,k]}, & l = 2k, \\ \Lambda_{m,[-k-1,k]}, & l = 2k+1. \end{cases}$$

That we can consider as a left free module, when multiplied by the left, and denoted by $V_{[l+1]}$, or as a right free module (when multiplication by matrices is performed by the right) and denoted by $W_{[l+1]}$. We will denote by $V = \varinjlim V_{[l]}$ and $W = \varinjlim W_{[l]}$ the corresponding direct limits, the left and right modules of matrix Laurent polynomials. The bilinear form

$$G(f,g) = \big\langle\!\big\langle g^\dagger, f\big\rangle\!\big\rangle_L = \big\langle\!\big\langle f^\dagger, g\big\rangle\!\big\rangle_R = \oint_{\mathbb{T}} f(z)\frac{\mathrm{d}\mu(z)}{\mathrm{i}z} g(z),$$

$$G_{i,j} := \oint_{\mathbb{T}} \chi^{(i)}(z) \frac{\mathrm{d}\mu(z)}{\mathrm{i}z} \big(\chi^{(j)}(z)\big)^{\top},$$

fulfills

$$G = \eta g^R = g^L \eta.$$

This can be understood as a change of basis in the left and right modules $W_{[l]}$ and $V_{[l]}$; the left moment matrix can be understood as the matrix of the bilinear form G when on the left module $W_{[l]}$ we apply the isomorphism or change of basis represented by the η matrix. Similarly, the right moment matrix can be understood as the matrix of the bilinear form G when on the right module $V_{[l]}$ we apply the isomorphism represented by the η matrix. Observe that the G dual vectors introduced in Appendix D are of the form

$$\begin{aligned} \big(\big(\varphi_1^L\big)_j\big)^* &= \big(\varphi_2^L\big)_j^{\dagger}, & \big(\big(\varphi_2^L\big)_j^{\dagger}\big)^* &= \big(\varphi_1^L\big)_j, \\ \big(\big(\varphi_1^R\big)_j\big)^* &= \big(\varphi_2^R\big)_j^{\dagger}, & \big(\big(\varphi_2^R\big)_j^{\dagger}\big)^* &= \big(\varphi_1^R\big)_j. \end{aligned}$$

Thus, following Appendix D, we consider the ring of G projections in these left and right modules

Definition 12.

(1) The Christoffel–Darboux projectors

$$\pi_L^{[l]} : V \longrightarrow V_{[l]}, \qquad \pi_R^{[l]} : W \longrightarrow W_{[l]},$$

are the ring left and right projections associated to the bilinear form G.

(2) The matrix Christoffel–Darboux kernels are

$$\begin{aligned} K^{L,[l]}\big(z,z'\big) &:= \sum_{k=0}^{l-1} \big[\big(\varphi_2^L\big)^{(k)}(z)\big]^{\dagger} \big(\varphi_1^L\big)^{(k)}\big(z'\big), \\ K^{R,[l]}\big(z,z'\big) &:= \sum_{k=0}^{l-1} \big(\varphi_1^R\big)^{(k)}(z) \big[\big(\varphi_2^R\big)^{(k)}\big(z'\big)\big]^{\dagger}. \end{aligned} \tag{29}$$

Proposition 21. *For the projections and matrix Christoffel–Darboux kernels introduced in Definition 12, we have the following relations*

$$\big(\pi_L^{[l]} f\big)(z) = \int_{\mathbb{T}} f\big(z'\big) \frac{\mathrm{d}\mu(z')}{\mathrm{i}z'} K^{L,[l]}\big(z',z\big) = \sum_{k=0}^{l-1} \big\langle\!\big\langle \big(\varphi_2^L\big)^{(k)}, f \big\rangle\!\big\rangle_L \big(\varphi_1^L\big)^{(k)}(z),$$

$$\forall f \in V,$$

$$\big[\big(\pi_L^{[l]}f\big)(z)\big]^\dagger=\oint_{\mathbb{T}}K^{L,[l]}\big(z,z'\big)\frac{\mathrm{d}\mu(z')}{\mathrm{i}z'}\big[f\big(z'\big)\big]^\dagger=\sum_{k=0}^{l-1}\big[\big(\varphi_2^L\big)^{(k)}(z)\big]^\dagger\big\langle\!\big\langle f,\big(\varphi_1^L\big)^{(k)}\big\rangle\!\big\rangle_L,$$
$$\forall f\in V,$$
$$\big(\pi_R^{[l]}f\big)(z)=\oint_{\mathbb{T}}K^{R,[l]}\big(z,z'\big)\frac{\mathrm{d}\mu(z')}{\mathrm{i}z'}f\big(z'\big)=\sum_{k=0}^{l-1}\big(\varphi_1^R\big)^{(k)}(z)\big\langle\!\big\langle\big(\varphi_2^R\big)^{(k)},f\big\rangle\!\big\rangle_R,$$
$$\forall f\in W,$$
$$\big[\big(\pi_R^{[l]}f\big)(z)\big]^\dagger=\oint_{\mathbb{T}}\big[f\big(z'\big)\big]^\dagger\frac{\mathrm{d}\mu(z')}{\mathrm{i}z}K^{R,[l]}\big(z',z\big)=\sum_{k=0}^{l-1}\big\langle\!\big\langle f,\big(\varphi_1^R\big)^{(k)}\big\rangle\!\big\rangle_R\big[\big(\varphi_2^R\big)^{(k)}(z)\big]^\dagger,$$
$$\forall f\in W.$$

Proposition 22. *The Christoffel–Darboux kernels have the reproducing property*

$$K^{H,[l]}(z,y)=\oint_{\mathbb{T}}K^{H,[l]}\big(z,z'\big)\frac{\mathrm{d}\mu(z')}{\mathrm{i}z'}K^{H,[l]}\big(z',y\big),\quad H=L,R. \tag{30}$$

Proof. This follows from the idempotency property of the π's. □

Moreover,

Proposition 23. *If the matrix measure μ is Hermitian, then*

(1) *the followings expansions are satisfied*

$$\big(\pi_L^{[l]}f\big)(z)=\sum_{k=0}^{l-1}\big\langle\!\big\langle\big(\varphi_1^L\big)^{(k)},f\big\rangle\!\big\rangle_L\big(h_k^L\big)^{-1}\big(\varphi_1^L\big)^{(k)}(z),\quad\forall f\in V,$$
$$\big(\pi_R^{[l]}f\big)(z)=\sum_{k=0}^{l-1}\big(\varphi_1^R\big)^{(k)}(z)\big(h_k^R\big)^{-1}\big\langle\!\big\langle\big(\varphi_1^R\big)^{(k)},f\big\rangle\!\big\rangle_R,\quad\forall f\in W;$$

(2) *the following Hermitian type property holds for the projectors*

$$\big\langle\!\big\langle\pi_H^{[l]}f,g\big\rangle\!\big\rangle_H=\big\langle\!\big\langle f,\pi_H^{[l]}g\big\rangle\!\big\rangle_H,\quad H=R,L.$$

When the matrix measure is Hermitian and positive definite, we have a standard scalar product and a complex Hilbert space, and the projections $\pi_H^{[l]}$ are orthogonal projections — not only in the module but in the geometrical sense as well — to the subspaces of truncated matrix Laurent polynomials; notice that there are two different, however equivalent, scalar products and distances involved. In this situation, as is well known, these projections give the best approximation within the truncated Laurent polynomials and the corresponding left and right distances.

2.5.2. The Christoffel–Darboux type formulae

Theorem 2. *For $\bar{z}z' \neq 1$, the matrix Christoffel–Darboux kernels fulfill*

$$
\begin{aligned}
&K^{L,[2l]}(z,z')(1-\bar{z}z') \\
&\quad= (\varphi_1^R)^{(2l)}(\bar{z}^{-1})h_{2l}^R(h_{2l-1}^R)^{-1}(\varphi_1^L)^{(2l-1)}(z') - (\varphi_1^R)^{(2l-1)}(\bar{z}^{-1})(\varphi_1^L)^{(2l)}(z') \\
&\quad= \big[\big[z(\varphi_2^L)^{(2l+1)}(z)\big]^\dagger h_{2l+1}^R - \big[z(\varphi_2^L)^{(2l)}(z)\big]^\dagger h_{2l}^L\alpha_{1,2l+1}^R\big](h_{2l-1}^R)^{-1}(\varphi_1^L)^{(2l-1)}(z') \\
&\qquad - \big[\big[z(\varphi_2^L)^{(2l-2)}(z)\big]^\dagger + \big[z(\varphi_2^L)^{(2l-1)}(z)\big]^\dagger\big[\alpha_{2,2l-1}^R\big]^\dagger\big](\varphi_1^L)^{(2l)}(z'), \\
&K^{L,[2l+1]}(z,z')(1-\bar{z}z') \\
&\quad= \big[z(\varphi_2^L)^{(2l+1)}(z)\big]^\dagger h_{2l+1}^R(h_{2l}^R)^{-1}\big[(\varphi_2^R)^{(2l)}\big]^\dagger(z'^{-1}) \\
&\qquad - \big[z(\varphi_2^L)^{(2l)}(z)\big]^\dagger\big[(\varphi_2^R)^{(2l+1)}\big]^\dagger(z'^{-1}) \\
&\quad= \big[z(\varphi_2^L)^{(2l+1)}(z)\big]^\dagger h_{2l+1}^R\big[(h_{2l-1}^R)^{-1}(\varphi_1^L)^{(2l-1)}(z') \\
&\qquad + \big[\alpha_{2,2l}^L\big]^\dagger(h_{2l}^L)^{-1}(\varphi_1^L)^{(2l)}(z')\big] \\
&\qquad - \big[z(\varphi_2^L)^{(2l)}(z)\big]^\dagger\big[(\varphi_1^L)^{(2l+2)}(z') - \alpha_{1,2l+2}^L(\varphi_1^L)^{(2l+1)}(z')\big],
\end{aligned}
\tag{31}
$$

$$
\begin{aligned}
&K^{R,[2l]}(z,z')(1-\bar{z}'z) \\
&\quad= \big[(\varphi_2^L)^{(2l)}\big]^\dagger(z^{-1})h_{2l}^L(h_{2l-1}^L)^{-1}\big[(\varphi_2^R)^{(2l-1)}(z')\big]^\dagger \\
&\qquad - \big[(\varphi_2^L)^{(2l-1)}\big]^\dagger(z^{-1})\big[(\varphi_2^R)^{(2l)}(z')\big]^\dagger \\
&\quad= z\big[(\varphi_1^R)^{(2l+1)}(z)h_{2l+1}^L - (\varphi_1^R)^{(2l)}(z)h_{2l}^R\big[\alpha_{2,2l+1}^L\big]^\dagger\big](h_{2l+1}^L)^{-1}\big[(\varphi_2^R)^{(2l-1)}(z')\big]^\dagger \\
&\qquad -z\big[(\varphi_1^R)^{(2l-2)}(z) - (\varphi_1^R)^{(2l-1)}(z)\alpha_{1,2l-1}^L\big]\big[(\varphi_2^R)^{(2l)}(z')\big]^\dagger, \\
&K^{R,[2l+1]}(z,z')(1-\bar{z}'z) \\
&\quad= z(\varphi_1^R)^{(2l+1)}(z)h_{2l+1}^L(h_{2l}^L)^{-1}(\varphi_1^L)^{(2l-1)}(\bar{z}'^{-1}) - z(\varphi_1^R)^{(2l)}(z)(\varphi_1^L)^{(2l+1)}(\bar{z}'^{-1}) \\
&\quad= z(\varphi_1^R)^{(2l+1)}(z)h_{2l+1}^L\big[(h_{2l-1}^L)^{-1}(\varphi_2^R)^{(2l-1)}(z')\big]^\dagger + \alpha_{1,2l}^R(h_{2l}^R)^{-1}\big[(\varphi_2^R)^{(2l)}(z')\big]^\dagger \\
&\qquad - z(\varphi_1^R)^{(2l)}(z)\big[(\varphi_2^R)^{(2l+2)}(z')\big]^\dagger - \big[\alpha_{2,2l+2}^R\big]^\dagger\big[(\varphi_2^R)^{(2l+1)}(z')\big]^\dagger.
\end{aligned}
\tag{32}
$$

Proof. See Appendix A. □

In terms of the matrix Szegő polynomials, we have

Corollary 2. *The matrix Christoffel–Darboux kernels can be expressed in terms of the matrix Szegő polynomials as follows*

$$
\begin{aligned}
K^{L,[2l]}(z,z') = \frac{(\bar{z}z'^{-1})^l}{1-\bar{z}z'}\big[&P_{1,2l}^R(\bar{z}^{-1})(h_{2l-1}^R)^{-1}(P_{2,2l-1}^R)^*(z') \\
&- (P_{2,2l-1}^L)^*(\bar{z}^{-1})(h_{2l-1}^L)^{-1}P_{1,2l}^L(z')\big],
\end{aligned}
$$

$$
\begin{aligned}
K^{L,[2l+1]}(z,z') &= \frac{\bar z^{l+1} z'^{-l}}{1-\bar z z'}\big[P^R_{1,2l+1}(\bar z^{-1})(h^R_{2l})^{-1}(P^R_{2,2l})^*(z') \\
&\quad - (P^L_{2,2l})^*(z^{-1})(h^L_{2l})^{-1}P^L_{1,2l+1}(z')\big], \\
K^{R,[2l]}(z,z') &= \frac{(z\bar z'^{-1})^l}{1-\bar z' z}\big[\big[P^L_{2,2l}(\bar z^{-1})\big]^\dagger (h^L_{2l-1})^{-1}\big[(P^L_{1,2l-1})^*(z')\big]^\dagger \\
&\quad - \big[(P^R_{1,2l-1})^*(\bar z^{-1})\big]^\dagger (h^R_{2l-1})^{-1}\big[P^R_{2,2l}(z')\big]^\dagger\big], \\
K^{R,[2l+1]}(z,z') &= \frac{z^{l+1}\bar z'^{-l}}{1-\bar z' z}\big[\big[P^L_{2,2l+1}(\bar z^{-1})\big]^\dagger (h^L_{2l})^{-1}\big[(P^L_{1,2l})^*(z')\big]^\dagger \\
&\quad - \big[(P^R_{1,2l})^*(\bar z^{-1})\big]^\dagger (h^R_{2l})^{-1}\big[P^R_{1,2l+1}(z')\big]^\dagger\big],
\end{aligned}
$$

where we assume that $\bar z z' \neq 1$.

As we have just seen, letting an operator act to the left or to the right and comparing the two results has been very successful with J_K. Actually we still have the operators C_0, C_{-1} to which we can also apply the same procedure to get some other interesting relations for the CD kernels.

Proposition 24. *The next relations between K^L and K^R hold*

$$
\begin{aligned}
&K^{R,[2l+1]}\left(z,\frac{1}{\bar z'}\right) = K^{L,[2l+1]}\left(\frac{1}{\bar z}, z'\right), \\
&K^{R,[2l+2]}\left(z,\frac{1}{\bar z'}\right) - K^{L,[2l+2]}\left(\frac{1}{\bar z}, z'\right) \\
&\quad = \left(\varphi^R_1\right)^{(2k+1)}(z)\left(\varphi^L_1\right)^{(2k+2)}(z') \\
&\qquad - \left(\varphi^R_1\right)^{(2k+2)}(z)\left(\mathbb{I} - \left[\alpha^R_{2,2k+2}\right]^\dagger \alpha^L_{1,2k+2}\right)\left(\varphi^L_1\right)^{(2k+1)}(z'), \\
&\frac{1}{z'}K^{R,[2l+1]}\left(z,\frac{1}{\bar z'}\right) - \frac{1}{z}K^{L,[2l+1]}\left(\frac{1}{\bar z}, z'\right) \\
&\quad = \left(\varphi^R_1\right)^{(2k)}(z)\left(\varphi^L_1\right)^{(2k+1)}(z') \\
&\qquad - \left(\varphi^R_1\right)^{(2k+1)}(z)\left(\mathbb{I} - \alpha^L_{1,2k+1}\left[\alpha^R_{2,2k+1}\right]^\dagger\right)\left(\varphi^L_1\right)^{(2k)}(z'), \\
&\frac{1}{z'}K^{R,[2l+2]}\left(z,\frac{1}{\bar z'}\right) = \frac{1}{z}K^{L,[2l+2]}\left(\frac{1}{\bar z}, z'\right).
\end{aligned}
$$

Proof. The first two relations arise when comparing the action of C_0 to the left or to the right in

$$
\phi^R_1(z)C_0\phi^L_1(z')
$$

and truncating the expressions up to $[2k+1]$ (first relation) or $[2k+2]$ (second one). The other two relations are obtained proceeding in the same way but using C_{-1} instead. □

3. MOLPUC and two dimensional Toda type hierarchies

Once we have explored how the Gauss–Borel factorization of block CMV moment matrices leads to the algebraic theory of MOLPUC, we are ready to show how this approach also connects these polynomials to integrable hierarchies of Toda type. We first introduce convenient deformations of the moment matrices, that as we will show correspond to deformations of the matrix measure. With these we will construct wave functions, Lax equations, Zakharov–Shabat equations, discrete flows and Darboux transformations and Miwa transformations. These last transformations will lead to interesting relations between the matrix Christoffel–Darboux kernels, Miwa shifted MOLPUC and their "norms". The integrable equations that we derive are a non-Abelian version of the Toeplitz lattice or non-Abelian ALL equations that extend, in the partial flows case, those of [27] — appearing these last ones in what we denominate total flows.

3.1. 2D Toda continuous flows

In order to construct deformation matrices which will act on the moment matrices (resulting in a deformation of the matrix measure) we first introduce some definitions.

Definition 13.

(1) Given the diagonal matrices $t^H_j = \operatorname{diag}(t^H_{j,1}, \dots, t^H_{j,m}) \in \operatorname{diag}_m$, $j = 0, 1, 2, \dots$, $H = L, R$ and $t^H_{j,a} \in \mathbb{C}$, we introduce

$$t^L := \big(t^L_0, t^L_1, t^L_2, \dots\big), \qquad t^H := \big(t^R_0, t^R_1, t^R_2, \dots\big)^\top,$$

we also impose $t^H_0 = 0$.

(2) We also consider the CMV ordered Fourier monomial vector but evaluated in $\varUpsilon$

$$\big[\chi(\varUpsilon)\big]^\top = \big(\mathbb{I}, \varUpsilon, \varUpsilon^{-1}, \varUpsilon^2, \varUpsilon^{-2}, \dots\big).$$

(3) With this we introduce

$$t^L * \chi(\varUpsilon) := \sum_{j=0}^\infty t^L_j \chi(\varUpsilon)^{(j)}, \qquad \big[\chi(\varUpsilon)\big]^\top * t^R := \sum_{j=0}^\infty \big[\chi(\varUpsilon)^{(j)}\big]^\top t^R_j.$$

The products in the above expressions are by blocks; i.e., the factors in $\mathbb{M}_m$ multiply $\mathbb{M}_m$ block of the $\mathbb{M}_\infty$ block matrix.

(4) The deformation matrices are

$$W_0\big(t^L\big) := \exp\big(t^L * \chi(\varUpsilon)\big), \qquad V_0\big(t^R\big) := \exp\big(\big[\chi(\varUpsilon)\big]^\top * t^R\big).$$

(5) The t-dependent deformation of moment matrices, $g^H(t)$, $H = L, R$, and their Gauss–Borel factorization are considered

$$g^L(t) := W_0\big(t^L\big)g^L\big[V_0\big(-\eta t^R\big)\big]^{-1}, \qquad g^L(t) = \big(S_1(t)\big)^{-1}S_2(t), \tag{33}$$
$$g^R(t) := \big[W_0\big(-t^L\eta\big)\big]^{-1}g^R V_0\big(t^R\big), \qquad g^R(t) = Z_2(t)\big(Z_1(t)\big)^{-1}. \tag{34}$$

Proposition 25.

(1) *The deformed moment matrices can be understood as the moment matrices for a deformed (t-dependent) measure given by*

$$\mathrm{d}\mu(z,t) := \exp\big(t_L\chi(z)\big)\mathrm{d}\mu(z)\exp\big(\chi(z)^\top t_R\big),$$

with the deformed Fourier series of the evolved matrix measure given by

$$F_{\mu(t)}(z) := \exp\big(t_L\chi(z)\big)F(z)\exp\big(\chi(z)^\top t_R\big). \tag{35}$$

(2) *The Hermitian character of the matrix measure is preserved by the deformation whenever* $t^L = (t^R)^\dagger\eta$.

Observe that in this paper we introduce a slightly different set of flows or deformations of the measure than those in the scalar case [13]. Despite that in that scalar situation both definitions give the very same flows that is not the case in this non-Abelian scenario, as in this case we have deformation matrices multiplying at the left and right of the initial matrix measure, and the order is relevant.

3.1.1. The Gauss–Borel approach to integrability

We consider the elements that enable us to construct the integrable hierarchy

Definition 14.

(1) Left and right wave matrices

$$\begin{aligned} W_1^L(t) &:= S_1(t)W_0\big(t^L\big), & W_2^L(t) &:= S_2(t)V_0\big(-\eta t^R\big), \\ W_1^R(t) &:= V_0\big(t^R\big)Z_1(t), & W_2^R(t) &:= W_0\big(-t^L\eta\big)Z_2(t). \end{aligned} \tag{36}$$

(2) Left and right wave and adjoint wave functions

$$\begin{aligned} \Psi_1^L(z,t) &= W_1^L(t)\chi(z), & \big(\Psi_1^L\big)^*(z,t) &= \big[\big(W_1^L\big)^{-1}(t)\big]^\dagger\chi^*(z), \\ \Psi_2^L(z,t) &= W_2^L(t)\chi^*(z), & \big(\Psi_2^L\big)^*(z,t) &= \big[\big(W_2^L\big)^{-1}(t)\big]^\dagger\chi(z), \\ \Psi_1^R(z,t) &= \chi(z)^\top W_1^R(t), & \big(\Psi_1^R\big)^*(z,t) &= \chi^*(z)^\top\big[\big(W_1^R\big)^{-1}(t)\big]^\dagger, \\ \Psi_2^R(z,t) &= \chi^*(z)^\top W_2^R(t), & \big(\Psi_2^R\big)^*(z,t) &= \chi(z)^\top\big[\big(W_2^R\big)^{-1}(t)\big]^\dagger. \end{aligned}$$

(3) Left and right Jacobi vector of matrices (using our previous notation)

$$\chi\big(J_H(t)\big) := \begin{pmatrix} \mathbb{I} \\ (J^H(t))^{-1} \\ J^H(t) \\ (J^H(t))^{-2} \\ (J^H(t))^2 \\ \vdots \end{pmatrix}, \quad H = L, R.$$

(4) Projection operators, $a = 1, \dots, m$

$$P_a^{(H,H')} := \begin{cases} S_1 E_{aa} S_1^{-1}, & H = L,\ H' = L, \\ S_2 E_{aa} S_2^{-1}, & H = R,\ H' = L, \\ Z_2^{-1} E_{aa} Z_2, & H = L,\ H' = R, \\ Z_1^{-1} E_{aa} Z_1, & H = R,\ H' = R. \end{cases} \tag{37}$$

(5) Left and right Lax matrices

$$L_1(t) := S_1(t)\Upsilon S_1(t)^{-1} = S_2(t)\Upsilon S_2(t)^{-1} = J^L(t),$$
$$R_1(t) := Z_1(t)^{-1}\Upsilon^{-1} Z_1(t) = Z_2(t)^{-1}\Upsilon^{-1} Z_2(t) = J^R(t), \tag{38}$$
$$L_2(t) := S_2(t)\Upsilon^{-1} S_2(t)^{-1} = S_1(t)\Upsilon^{-1} S_1(t)^{-1} = \big(J^L(t)\big)^{-1}$$
$$R_2(t) := Z_2(t)^{-1}\Upsilon Z_2(t) = Z_1(t)^{-1}\Upsilon Z_1(t) = \big(J^R(t)\big)^{-1}. \tag{39}$$

(6) Zakharov–Shabat matrices

$$B_{j,a}^{(H,H')} := \begin{cases} (S_1 E_{aa}(\chi(\Upsilon))^{(j)} S_1^{-1})_+, & H = L,\ H' = L, \\ -(S_2 E_{aa}(\chi(\Upsilon))^{(j)} S_2^{-1})_-, & H = R,\ H' = L, \\ -(Z_2^{-1} E_{aa}(\chi(\Upsilon^{-1}))^{(j)} Z_2)_+, & H = L,\ H' = R, \\ (Z_1^{-1} E_{aa}(\chi(\Upsilon^{-1}))^{(j)} Z_1)_-, & H = R,\ H' = R, \end{cases}$$
$$B_j^{(H,H')} := \begin{cases} ((\chi(J^L))^{(j)})_+, & H = L,\ H' = L, \\ -((\chi(J^L))^{(j)})_-, & H = R,\ H' = L, \\ -(\chi((J^R)^{-1})^{(j)})_+, & H = L,\ H' = R, \\ (\chi((J^R)^{-1})^{(j)})_-, & H = R,\ H' = R. \end{cases} \tag{40}$$

(7) A time dependent intertwining operator

$$C_{[p]}(t) = Z_2(t)^{-1}\eta\Upsilon^p S_1(t)^{-1} = Z_1(t)^{-1}\eta\Upsilon^p S_2(t)^{-1}. \tag{41}$$

Observe that

$$g^L = \big(W_1^L(t)\big)^{-1} W_2^L(t), \qquad g^R = W_2^R(t)\big(W_1^R\big)^{-1}(t). \tag{42}$$

Definition 15. For $H = R, L$ we introduce the total derivatives

$$\partial_{H,j} := \sum_{a=1}^{m} \frac{\partial}{\partial t_{j,a}^{H}}.$$

We now present the linear systems, Lax equations and Zakharov–Shabat equations that characterize integrability

Proposition 26. *The following equations hold:*

(1) *Linear systems for the wave matrices*

$$\frac{\partial W_i^L}{\partial t_{j,a}^H} = B_{j,a}^{H,L} W_i^L, \qquad \partial_{H,j} W_i^L = B_j^{H,L} W_i^L,$$
$$\frac{\partial W_i^R}{\partial t_{j,a}^H} = W_i^R B_{j,a}^{H,R}, \qquad \partial_{H,j} W_i^R = W_i^R B_j^{H,R},$$

for $i = 1, 2,\ H = L, R,\ a = 1, \dots, m,\ j = 0, 1, \dots$.

(2) *Lax equations*

$$\frac{\partial J^{H'}}{\partial t_{j,a}^H} = \left[B_{j,a}^{H,H'}, J^{H'}\right], \qquad \partial_{H,j} J^{H'} = \left[B_j^{H,H'}, J^{H'}\right],$$
$$\frac{\partial P_b^{H',H''}}{\partial t_{j,a}^H} = \left[B_{j,a}^{H,H''}, P_b^{H',H''}\right], \qquad \partial_{H,j} P_b^{H',H''} = \left[B_j^{H,H''}, P_b^{H',H''}\right]$$

with $H, H', H'' = L, R,\ a, b = 1, \dots, m$ *and* $j = 0, 1, \dots$.

(3) *Evolution of the dressed intertwining operator*

$$\frac{\partial C_{[p]}}{\partial t_{j,a}^H} = -B_{j,a}^{H,R} C_{[p]} - C_{[p]} B_{j,a}^{H,L}, \qquad \frac{\partial C_{[p]}}{\partial t_j^H} = -B_j^{H,R} C_{[p]} - C_{[p]} B_j^{H,L}.$$

(4) *Zakharov-Shabat equations*

$$\frac{\partial B_{j_1,b_1}^{H_1,H'}}{\partial t_{j_2,a_2}^{H_2}} - \frac{\partial B_{j_2,b_2}^{H_2,H'}}{\partial t_{j_1,a_1}^{H_1}} + \left[B_{j_1,b_1}^{H_1,H'}, B_{j_2,b_2}^{H_2,H'}\right] = 0.$$

Proof. See Appendix A □

From the definitions of the wave functions, the action of Υ on χ, the expression (35), and the relations (23), it follows that

Proposition 27. *The wave functions are linked to the CMV polynomials and the Fourier series of the measure as follows*

$$\begin{aligned}
\Psi_1^L(z,t) &= \phi_1^L(z,t)\exp\big(t^L\chi(z)\big),\\
\big(\Psi_1^L\big)^*(z,t) &= 2\pi z^{-1}\phi_2^L\big(z^{-1},t\big)F_\mu^\dagger(z)\exp\big(-\overline{t^L}\chi(z)\big),\\
\Psi_2^L(z,t) &= 2\pi z^{-1}\phi_1^L\big(z^{-1},t\big)F_\mu\big(z^{-1}\big)\exp\big(-\chi\big(z^{-1}\big)^\top t^R\big),\\
\big(\Psi_2^L\big)^*(z,t) &= \phi_2^L(z,t)\exp\big(\chi\big(z^{-1}\big)^\top\overline{t^R}\big),\\
\Psi_1^R(z,t) &= \exp\big(\chi(z)^\top t^R\big)\phi_1^R(z,t),\\
\big(\Psi_1^R\big)^*(z,t) &= 2\pi z^{-1}\exp\big(-\chi(z)^\top\overline{t^R}\big)F_\mu^\dagger(z)\big(\phi_2^R\big)\big(z^{-1},t\big),\\
\Psi_2^R(z,t) &= 2\pi z^{-1}\exp\big(-t^L\chi\big(z^{-1}\big)\big)F_\mu\big(z^{-1}\big)\phi_1^R\big(z^{-1},t\big),\\
\big(\Psi_2^R\big)^*(z,t) &= \exp\big(\overline{t^L}\chi\big(z^{-1}\big)\big)\phi_2^R(z,t).
\end{aligned}\tag{43}$$

These wave functions are also eigen-functions of the Lax matrices (38) L_i, R_i, *for* $i=1,2$,

$$\begin{aligned}
L_i\Psi_i^L &= z\Psi_i^L, & \Psi_i^R R_i &= z\Psi_i^R,\\
L_i^\dagger\big(\Psi_i^L\big)^* &= z\big(\Psi_i^L\big)^*, & \big(\Psi_i^R\big)^* R_i^\dagger &= z\big(\Psi_i^R\big)^*.
\end{aligned}$$

3.1.2. CMV matrices and matrix Toeplitz lattice

For the CMV ordering of the Laurent basis, the Lax equations acquire a dynamical non-linear system form that is the matrix version, in the CMV context, of the Toeplitz lattice developed in [3]. In [27] Mattia Cafasso presented a non-Abelian extension of the TL which corresponds to our total flows. The partial flows presented here are, to our knowledge, new in the literature.

Proposition 28. *The Lax equations result in the following non-linear dynamical system for the matrix Verblunsky coefficients* $H=L,R$:

- *Partial flows*

$$\begin{aligned}
\frac{\partial}{\partial t_{1,a}^L}\alpha_{1,k}^R &= -\big(h_{k-1}^L\big)^{-1}\alpha_{1,k-1}^L E_{a,a}h_k^R,\\
\frac{\partial}{\partial t_{1,a}^L}\big[\alpha_{2,k}^L\big]^\dagger &= \big(h_{k-1}^R\big)^{-1}E_{a,a}\big[\alpha_{2,k+1}^R\big]^\dagger h_k^L,\\
\frac{\partial}{\partial t_{1,a}^R}\big[\alpha_{2,k}^R\big]^\dagger &= h_k^R\big[\alpha_{2,k+1}^L\big]^\dagger E_{a,a}\big(h_{k-1}^L\big)^{-1},\\
\frac{\partial}{\partial t_{1,a}^R}\alpha_{1,k}^L &= -\big(h_k^L\big)E_{a,a}\alpha_{1,k-1}^R\big(h_{k-1}^R\big)^{-1},\\
\frac{\partial}{\partial t_{2,a}^L}\alpha_{1,k}^R &= \big(h_{k-1}^L\big)^{-1}E_{a,a}\alpha_{1,k+1}^L h_k^R,\\
\frac{\partial}{\partial t_{2,a}^L}\big[\alpha_{2,k}^L\big]^\dagger &= -\big(h_{k-1}^R\big)^{-1}\big[\alpha_{2,k-1}^R\big]^\dagger E_{a,a}h_k^L,
\end{aligned}$$

$$\frac{\partial}{\partial t^R_{2,a}}\big[\alpha^R_{2,k}\big]^\dagger = -h^R_k E_{a,a}\big[\alpha^L_{2,k-1}\big]^\dagger\big(h^L_{k-1}\big)^{-1},$$

$$\frac{\partial}{\partial t^R_{2,a}}\alpha^L_{1,k} = h^L_k \alpha^R_{1,k+1} E_{a,a}\big(h^R_{k-1}\big)^{-1},$$

$$\frac{\partial}{\partial t^L_{1,a}}h^L_k = -\alpha^L_{1,k} E_{a,a}\big(\alpha^R_{2,k+1}\big)^\dagger h^L_k,$$

$$\frac{\partial}{\partial t^R_{1,a}}h^R_k = -h^R_k\big(\alpha^L_{2,k-1}\big)^\dagger E_{a,a}\alpha^R_{1,k},$$

$$\frac{\partial}{\partial t^L_{2,a}}h^R_k = -\big(\alpha^R_{2,k}\big)^\dagger E_{a,a}\alpha^L_{1,k+1} h^R_k,$$

$$\frac{\partial}{\partial t^R_{2,a}}h^L_k = -h^L_k \alpha^R_{1,k+1} E_{a,a}\big(\alpha^L_{2,k}\big)^\dagger.$$

- *Total flows [27]*

$$\partial_{H,1}\big[\alpha^R_{2,k}\big]^\dagger = \big[\alpha^R_{2,k+1}\big]^\dagger\big(\mathbb{I} - \alpha_{1,k}\big[\alpha^R_{2,k}\big]^\dagger\big),$$

$$\partial_{H,1}\alpha^L_{1,k} = -\big(\mathbb{I} - \alpha^L_{1,k}\big[\alpha^R_{2,k}\big]^\dagger\big)\alpha^L_{1,k-1},$$

$$\partial_{H,1}\big[\alpha^L_{2,k}\big]^\dagger = \big(\mathbb{I} - \big[\alpha^L_{2,k}\big]^\dagger\alpha^R_{1,k}\big)\big[\alpha^L_{2,k+1}\big]^\dagger,$$

$$\partial_{H,1}\alpha^R_{1,k} = -\alpha^R_{1,k-1}\big(\mathbb{I} - \big[\alpha^L_{2,k}\big]^\dagger\alpha^R_{1,k}\big),$$

$$\partial_{H,2}\big[\alpha^R_{2,k}\big]^\dagger = -\big(\mathbb{I} - \big[\alpha^R_{2,k}\big]^\dagger\alpha^L_{1,k}\big)\big[\alpha^R_{2,k-1}\big]^\dagger,$$

$$\partial_{H,2}\alpha^L_{1,k} = \alpha^L_{1,k+1}\big(\mathbb{I} - \big[\alpha^R_{2,k}\big]^\dagger\alpha^L_{1,k}\big),$$

$$\partial_{H,2}\big[\alpha^L_{2,k}\big]^\dagger = -\big[\alpha^L_{2,k-1}\big]^\dagger\big(\mathbb{I} - \alpha^R_{1,k}\big[\alpha^L_{2,k}\big]^\dagger\big),$$

$$\partial_{H,2}\alpha^R_{1,k} = \big(\mathbb{I} - \alpha^R_{1,k}\big[\alpha^L_{2,k}\big]^\dagger\big)\alpha^R_{1,k+1},$$

$$\partial_{H,1}h^L_k = -\alpha^L_{1,k}\big(\alpha^R_{2,k+1}\big)^\dagger h^L_k,$$

$$\partial_{H,1}h^R_k = -h^R_k\big(\alpha^L_{2,k-1}\big)^\dagger\alpha^R_{1,k},$$

$$\partial_{H,2}h^R_k = -\big(\alpha^R_{2,k}\big)^\dagger\alpha^L_{1,k+1}h^R_k,$$

$$\partial_{H,2}h^L_k = -h^L_k\alpha^R_{1,k+1}\big(\alpha^L_{2,k}\big)^\dagger.$$

Proof. To obtain the partial flows, it is enough to use the Lax equations for $j, p = 1, 2$ and operate. In order to obtain the total flows, we go back to the partial flows, and sum in a. From the Lax equations, we know that in this total case we no longer need to distinguish between R, L. This procedure leads to the result that is finally rewritten using the relations in Proposition 17. □

3.1.3. Bilinear equations

Bilinear equations are an alternative way of expressing an integrable hierarchy developed by the Japanese school, see [40–42]. We are going to show that these MOLPUC also fulfill a particular type of bilinear equations. These results are the matrix extensions of the scalar situation described in [13]. Let us start by considering the wave semi-infinite matrices $W_i^H(t)$ 36 associated to the moment matrix g^H, $H = L, R$. Since the last one is time independent, the reader can easily check that

Proposition 29.

(1) *The wave matrices associated to different times satisfy*

$$\begin{aligned} W_1^L(t)\big(W_1^L(t')\big)^{-1} &= W_2^L(t)\big(W_2^L(t')\big)^{-1}, \\ \big(W_1^R(t)\big)^{-1}W_1^R(t') &= \big(W_2^R(t)\big)^{-1}W_2^R(t'). \end{aligned} \tag{44}$$

(2) *The vectors χ, χ^* fulfill*

$$\operatorname{Res}_{z=0}\big[\chi(z)\big(\chi^*(\bar z)\big)^\dagger\big] = \operatorname{Res}_{z=0}\big[\chi^*(z)\big(\chi(\bar z)\big)^\dagger\big] = \mathbb{I}.$$

(3) *One has that the product of two matrices can be expressed as*

$$\begin{aligned} UV &= \operatorname{Res}_{z=0}\big[U\chi(z)\big(V^\dagger\chi^*(\bar z)\big)^\dagger\big] = \operatorname{Res}_{z=0}\big[U\chi^*(z)\big(V^\dagger\chi(\bar z)\big)^\dagger\big] \\ &= \operatorname{Res}_{z=0}\big[\big(\chi^*(\bar z)^\top U^\dagger\big)^\dagger\chi(z)^\top V\big] = \operatorname{Res}_{z=0}\big[\big(\chi(\bar z)^\top U^\dagger\big)^\dagger\chi^*(z)^\top V\big] \end{aligned}$$

From where we derive

Theorem 3. *For two different set of times $t, \tilde t$ the wave functions satisfy*

$$\operatorname{Res}_{z=0}\big[\Psi_1^L(z,t)\big[\big(\Psi_1^L\big)^*(\bar z,\tilde t)\big]^\dagger\big] = \operatorname{Res}_{z=0}\big[\Psi_2^L(z,t)\big[\big(\Psi_2^L\big)^*(\bar z,\tilde t)\big]^\dagger\big], \tag{45}$$

$$\operatorname{Res}_{z=0}\big[\big[\big(\Psi_1^R\big)^*(\bar z,t)\big]^\dagger\Psi_1^R(z,\tilde t)\big] = \operatorname{Res}_{z=0}\big[\big[\big(\Psi_2^R\big)^*(\bar z,t)\big]^\dagger\Psi_2^R(z,\tilde t)\big]. \tag{46}$$

From the identities in (43) the previous theorem can be rewritten in terms of CMV polynomials as

$$\begin{aligned} &\operatorname{Res}_{z=0}\big[\big(\varphi_1^L\big)^{(l)}(z,t)\big(\exp\big(\big(t^L-\tilde t^L\big)\chi(z)\big)z^{-1}F_\mu(z)\big)\big[\big(\varphi_2^L\big)^{(m)}\big(\bar z^{-1},\tilde t\big)\big]^\dagger\big] \\ &\quad = -\operatorname{Res}_{z=\infty}\big[\big(\varphi_1^L\big)^{(l)}(z,t)\big(z^{-1}F_\mu(z)\exp\big(\chi(z)^\top\big(\tilde t^R-t^R\big)\big)\big)\big[\big(\varphi_2^L\big)^{(m)}\big(\bar z^{-1},\tilde t\big)\big]^\dagger\big], \\ &\operatorname{Res}_{z=0}\big[\big[\big(\varphi_2^R\big)^{(l)}\big(\bar z^{-1},t\big)\big]^\dagger\big(z^{-1}F_\mu(z)\exp\big(\chi(z)^\top\big(t_R'-t_R\big)\big)\big)\big(\varphi_1^R\big)^{(m)}(z,\tilde t)\big] \\ &\quad = -\operatorname{Res}_{z=\infty}\big[\big[\big(\varphi_2^R\big)^{(l)}\big(\bar z^{-1},t\big)\big]^\dagger\big(\exp\big(\big(t_L-t_L'\big)\chi(z)\big)z^{-1}F_\mu(z)\big)\big(\varphi_1^R\big)^{(m)}(z,\tilde t)\big]. \end{aligned}$$

Here we have used that $\operatorname{Res}_{z=\infty} F(z) = -\operatorname{Res}_{z=0} z^{-2}F(z^{-1})$. Alternatively, we can write all the previous expressions using integrals instead of using residues. To do this, let us denote by γ_0 and γ_∞ two positively oriented circles around $z=0$ and $z=\infty$, respectively, included in the annulus of convergence of the Fourier series of the matrix measure, such that they do not include different simple poles that $z=0,\infty$, respectively. Then,

$$\oint_{\gamma_0} \Psi_1^L(z,t)\big[\big(\Psi_1^L\big)^*(\bar z,\tilde t)\big]^\dagger \mathrm{d}z = \oint_{\gamma_0} \Psi_2^L(z,t)\big[\big(\Psi_2^L\big)^*(\bar z,\tilde t)\big]^\dagger \mathrm{d}z, \tag{47}$$

$$\oint_{\gamma_0} \big[\big(\Psi_1^R\big)^*(\bar z,t)\big]^\dagger \Psi_1^L(z,\tilde t)\mathrm{d}z = \oint_{\gamma_0} \big[\big(\Psi_2^R\big)^*(\bar z,t)\big]^\dagger \Psi_2^R(z,\tilde t)\mathrm{d}z \tag{48}$$

or, in terms of matrix Laurent orthogonal polynomials and Fourier series of the matrix measure:

Proposition 30. *The evolved MOLPUC satisfy*

$$\begin{aligned}
&\oint_{\gamma_0} \big(\varphi_1^L\big)^{(l)}(z,t)\big(\exp\big(\big(t^L-\tilde t^L\big)\chi(z)\big)z^{-1}F_\mu(z)\big)\big[\big(\varphi_2^L\big)^{(m)}\big(\bar z^{-1},\tilde t\big)\big]^\dagger \mathrm{d}z\\
&\quad = \oint_{\gamma_\infty} \big(\varphi_1^L\big)^{(l)}(z,t)\big(z^{-1}F_\mu(z)\exp\big(\chi(z)^\top\big(\tilde t^R-t^R\big)\big)\big)\big[\big(\varphi_2^L\big)^{(m)}\big(\bar z^{-1},\tilde t\big)\big]^\dagger \mathrm{d}z,\\
&\oint_{\gamma_0} \big[\big(\varphi_2^R\big)^{(l)}\big(\bar z^{-1},t\big)\big]^\dagger\big(z^{-1}F_\mu(z)\exp\big(\chi(z)^\top\big(\tilde t^R-t^R\big)\big)\big)\big(\varphi_1^R\big)^{(m)}(z,\tilde t)\mathrm{d}z\\
&\quad = \oint_{\gamma_\infty} \big[\big(\varphi_2^R\big)^{(l)}\big(\bar z^{-1},t\big)\big]^\dagger\big(\exp\big(\big(t^L-\tilde t^L\big)\chi(z)\big)z^{-1}F_\mu(z)\big)\big(\varphi_1^R\big)^{(m)}(z,\tilde t)\mathrm{d}z.
\end{aligned}$$

3.2. 2D Toda discrete flows

Given a couple of sequences of diagonal matrices

$$d=\{d_+,d_-\},\quad d_\pm=\{d_{\pm,0}=0,d_{\pm,1},d_{\pm,2},\dots\},\ d_{\pm,j}\in\operatorname{diag}_m,$$

and a pair of non-negative integers $n=\{n_+,n_-\}\in\mathbb{Z}_+^2$, we consider the next semi-infinite block matrices

$$\begin{aligned}
\Delta_d^L(n) &= \big(\mathbb{I}-d_{-,0}\Upsilon^{-1}\big)\cdots\big(\mathbb{I}-d_{-,n_-}\Upsilon^{-1}\big)(\mathbb{I}-d_{+,0}\Upsilon)\cdots(\mathbb{I}-d_{+,n_+}\Upsilon),\\
\Delta_d^R(n) &= (\mathbb{I}-d_{-,0}\Upsilon)\cdots(\mathbb{I}-d_{-,n_-}\Upsilon)\big(\mathbb{I}-d_{+,0}\Upsilon^{-1}\big)\cdots\big(\mathbb{I}-d_{+,n_+}\Upsilon^{-1}\big).
\end{aligned}$$

Observe that the order of the factors does not alter the product as each of them commutes with the others.

Definition 16. Given two couples of sequences of diagonal matrices, say $d^H = \{d_+^H, d_-^H\}$, $H = L, R$, we introduce the discrete flows for the right and left moment matrices

$$\begin{aligned} &g^H(n^L, n^R) = \Delta^H_{d^L}(n^L) g^H \Delta^H_{d^R}(n^R), \qquad n^H = \{n_+^H, n_-^H\} \in \mathbb{Z}_+^2, \\ &g^H(0,0) = g^H, \quad H = L, R. \end{aligned}$$

The property $\eta g^L(n^L, n^R) = g^R(n^L, n^R)\eta$ is easily checked and it follows that we have an associated measure of which these are the corresponding left and right moment matrices given by

$$\begin{aligned} \mathrm{d}\mu(n^L, n^R) = {}& \left[\prod_{i=0}^{n_-^L} (\mathbb{I} - d^L_{-,i} z^{-1}) \prod_{j=0}^{n_+^L} (\mathbb{I} - d^L_{+,j} z)\right] \\ & \times \mathrm{d}\mu \left[\prod_{i=0}^{n_-^R} (\mathbb{I} - d^R_{-,i} z^{-1}) \prod_{j=0}^{n_+^R} (\mathbb{I} - d^R_{+,j} z)\right]. \end{aligned}$$

The measure is Hermitian if the following conditions are fulfilled

$$[d^R_{\mp,j}]^\dagger = d^L_{\pm,j} = d_{\pm,j}, \qquad n^L_\pm = n^R_\mp = n_\pm,$$

being the evolved measure

$$\begin{aligned} \mathrm{d}\mu(n_+, n_-) = {}& \left[\prod_{i=0}^{n_-} (\mathbb{I} - d_{-,i} z^{-1}) \prod_{j=0}^{n_+} (\mathbb{I} - d_{+,j} z)\right] \\ & \times \mathrm{d}\mu \left[\prod_{i=0}^{n_-} (\mathbb{I} - d_{-,i} z^{-1}) \prod_{j=0}^{n_+} (\mathbb{I} - d_{+,j} z)\right]^\dagger. \end{aligned}$$

Positive definiteness for the Hermitian situation can be ensured if we request $d_i := d_{+,i} = [d_{-,i}]^\dagger$ and $n := n_+ = n_-$ so that

$$\mathrm{d}\mu(n) = \left[\sum_{a=1}^{m} \left(\prod_{j=0}^{n} |1 - d_{j,a} z|^2\right) E_{a,a}\right] \mathrm{d}\mu \left[\sum_{a=1}^{m} \left(\prod_{j=0}^{n} |1 - d_{j,a} z|^2\right) E_{a,a}\right].$$

As in the continuous case, we introduce

Definition 17. The wave matrices, depending on discrete variables $n^L, n^R \in \mathbb{Z}_+^2$, are defined as

$$\begin{aligned} W_1^L(n^L, n^R) &= S_1(n^L, n^R) \Delta^L_{d^L}(n^L), & W_2^L(n^L, n^R) &= S_2(n^L, n^R) \big(\Delta^L_{d^R}(n^R)\big)^{-1}, \\ W_1^R(n^L, n^R) &= \Delta^R_{d^R}(n^R) Z_1(n^L, n^R), & W_2^R(n^L, n^R) &= \big(\Delta^R_{d^L}(n^L)\big)^{-1} Z_2(n^L, n^R). \end{aligned}$$

Hence

$$g^L = \big[W_1^L(n^L,n^R)\big]^{-1} W_2^L(n^L,n^R), \qquad g^R = W_2^R(n^L,n^R)\big[W_1^R(n^L,n^R)\big]^{-1}.$$

We also need to introduce the following objects

Definition 18.

(1) Given a diagonal matrix $d \in \operatorname{diag}_m$, we define the semi-infinite block matrices

$$\delta_\pm^{H,H'}(d) = \begin{cases} S_1(\mathbb{I} - d\Upsilon^{\pm 1})S_1^{-1}, & H = L,\ H' = L, \\ S_2(\mathbb{I} - d\Upsilon^{\pm 1})S_2^{-1}, & H = R,\ H' = L, \\ Z_2^{-1}(\mathbb{I} - d\Upsilon^{\mp 1})Z_2, & H = L,\ H' = R, \\ Z_1^{-1}(\mathbb{I} - d\Upsilon^{\mp 1})Z_1, & H = R,\ H' = R. \end{cases}$$

(2) The shifts are

$$T_+^L : \begin{pmatrix} (n_+^L, n_-^L) \longrightarrow (n_+^L + 1, n_-^R) \\ n^R \longrightarrow n^R \end{pmatrix}, \qquad T_-^L : \begin{pmatrix} (n_+^L, n_-^R) \longrightarrow (n_+^L, n_-^L + 1) \\ n^R \longrightarrow n^R \end{pmatrix},$$
$$T_+^R : \begin{pmatrix} n_L \longrightarrow n_L \\ (n_+^R, n_-^R) \longrightarrow (n_+^R + 1, n_-^R) \end{pmatrix}, \qquad T_-^R : \begin{pmatrix} n_L \longrightarrow n_L \\ (n_+^R, n_-^R) \longrightarrow (n_+^R, n_-^R + 1) \end{pmatrix}.$$

For any diagonal matrix $d = \sum_{a=1}^m d_a E_{a,a} \in \operatorname{diag}_m$, $d_a \in \mathbb{C}$, we introduce the semi-infinite matrices

$$d^{H,H'} = \sum_{a=1}^m d_a P_a^{H,H'},$$

where $P_a^{H,H'}$ was defined in (37); observe that when $d = c\mathbb{I}_m$, $c \in \mathbb{C}$, we have $d^{H,H'} = c\mathbb{I}$.

Notice that the $\delta_\pm^{H,H'}$ are just particular combinations of the block Jacobi matrices J^H

$$\delta_\pm^{H,H'}(d) = \mathbb{I} - d^{H,H'}\big(J^{H'}\big)^{\pm 1}.$$

Proposition 31. *If $g^H(n^L,n^R)$, $(T_\pm^H g^{H'})(n^L,n^R)$, $H,H' = L,R$, admit a block LU factorization, then the δ matrices introduced in Definition 18(1) also admit a block LU factorization.*

Proof. We have

$$T_\pm^L g^L = \big(\mathbb{I} - d_{\pm,n_\pm^L+1}^L \Upsilon^{\pm 1}\big) g^L$$
$$\Longrightarrow \quad S_1\big(T_\pm^L S_1\big)^{-1}\big(T_\pm^L S_2\big)S_2^{-1} = S_1\big(\mathbb{I} - d_{\pm,n_\pm^L+1}^L \Upsilon^{\pm 1}\big)S_1^{-1} = \delta_\pm^{L,L}\big(d_{\pm,n_\pm^L+1}^L\big),$$
$$T_\pm^L g^R = \big(\mathbb{I} - d_{\pm,n_\pm^L+1}^L \Upsilon^{\mp 1}\big) g^R$$

$$\implies \quad Z_2^{-1}\big(T^L_\pm Z_2\big)\big(T^L_\pm Z_1\big)^{-1}Z_1 = Z_2^{-1}\big(\mathbb{I}-d^L_{\pm,n^L_\pm+1}\varUpsilon^{\mp 1}\big)Z_2 = \delta^{L,R}_\pm\big(d^L_{\pm,n^L_\pm+1}\big),$$

$$T^R_\pm g^L = g^L\big(\mathbb{I}-d^R_{\pm,n^R_\pm+1}\varUpsilon^{\pm 1}\big)$$

$$\implies \quad S_1\big(T^R_\pm S_1\big)^{-1}\big(T^R_\pm S_2\big)S_2^{-1} = S_2\big(\mathbb{I}-d^R_{\pm,n^{R,\pm}+1}\varUpsilon^{\pm 1}\big)S_2^{-1} = \delta^{R,L}_\pm\big(d^R_{\pm,n^R_\pm+1}\big),$$

$$T^R_\pm g^R = g^R\big(\mathbb{I}-d^R_{\pm,n^R_\pm+1}\varUpsilon^{\mp 1}\big)$$

$$\implies \quad Z_2^{-1}\big(T^R_\pm Z_2\big)\big(T^R_\pm Z_1\big)^{-1}Z_1 = Z_1^{-1}\big(\mathbb{I}-d^R_{\pm,n^R_\pm+1}\varUpsilon^{\mp 1}\big)Z_1 = \delta^{R,R}_\pm\big(d^R_{\pm,n^R_\pm+1}\big).$$

Therefore, for $H = L, R$,

$$\begin{aligned} \delta^{H,L}_\pm &= \big(\delta^{H,L}_\pm\big)_-^{-1}\cdot\big(\delta^{H,L}_\pm\big)_+, & \big(\delta^{H,L}_\pm\big)_- &= \big(T^H_\pm S_1\big)S_1^{-1}\in\mathscr{L}, \\ \big(\delta^{H,L}_\pm\big)_+ &= \big(T^H_\pm S_2\big)S_2^{-1}\in\mathscr{U}, & \delta^{H,R}_\pm &= \big(\delta^{H,R}_\pm\big)_-\cdot\big(\delta^{H,R}_\pm\big)_+^{-1}, \\ \big(\delta^{H,R}_\pm\big)_- &= Z_2^{-1}\big(T^H_\pm Z_2\big)\in\mathscr{L}, & \big(\delta^{H,R}_\pm\big)_+ &= Z_1^{-1}\big(T^H_\pm Z_1\big)\in\mathscr{U}. \end{aligned} \qquad \square$$

Definition 19. We define

$$\omega^{H,H'}_\pm := \begin{cases} (\delta^{L,L}_\pm)_+ = (T^L_\pm S_2)S_2^{-1}, & H=L,\ H'=L, \\ (\delta^{R,L}_\pm)_- = (T^R_\pm S_1)S_1^{-1}, & H=R,\ H'=L, \\ (\delta^{L,R}_\pm)_- = Z_1^{-1}(T^L_\pm Z_1), & H=L,\ H'=R, \\ (\delta^{R,R}_\pm)_+ = Z_2^{-1}(T^R_\pm Z_2), & H=R,\ H'=R. \end{cases}$$

We are ready to derive discrete integrability.

Theorem 4.

- *The discrete linear systems*

$$T^H_\pm W^L_i = \omega^{H,L}_\pm W^L_i, \qquad T^H_\pm W^R_i = W^R_i\omega^{H,R}_\pm, \quad i=1,2,\ H=L,R.$$

- *Discrete Lax equations hold*

$$T^H_\pm J^L = \omega^{H,L}_\pm J^L\big(\omega^{H,L}_\pm\big)^{-1}, \qquad T^H_\pm J^R = \big(\omega^{H,R}_\pm\big)^{-1}J^R\omega^{H,R}_\pm, \quad H=L,R.$$

- *Intertwining matrix*

$$T^H_\pm C_{[p]} = \big(\omega^{H,R}_\pm\big)^{-1}C_{[p]}\big(\omega^{H,L}_\pm\big)^{-1}, \quad H=L,R.$$

- *Zakharov–Shabat equations*

$$\big(T_a^H \omega_b^{H',L}\big)\omega_a^{H,L} = \big(T_b^{H'} \omega_a^{H,L}\big)\omega_b^{H',L}, \qquad \omega_a^{H,R}\big(T_a^H \omega_b^{H',R}\big) = \omega_b^{H',R}\big(T_b^{H'} \omega_a^{H,R}\big),$$
$$a, b = \pm, \ H, H' = L, R.$$

- *Continuous-discrete equations*

$$\frac{\partial \omega_\pm^{H',L}}{\partial t_{j,a}^H} + \omega_\pm^{H',L} B_{j,a}^{H,L} = \big(T_\pm^{H'} B_{j,a}^{H,L}\big)\omega_\pm^{H',L},$$
$$\frac{\partial \omega_\pm^{H',R}}{\partial t_{j,a}^H} + B_{j,a}^{H,R} \omega^{H',R}\pm = \omega_\pm^{H',R}\big(T_\pm^{H'} B_{j,a}^{H,R}\big),$$

with $H, H' = L, R$, $a = 1, \dots, m$ *and* $j = 0, 1, \dots$.

From these results, one may derive discrete matrix equations for the Verblunsky coefficients.

It also follows that these flows are extensions of Darboux transformations, see [13] for the scalar case. Each of these discrete shifts is generalization of the typical Darboux transformation corresponding to the flip of the upper and lower triangular factors of the operators $\delta_\pm^{H,H'}$. These flips occur in some specific cases as follows. Let us assume that the diagonal matrices $d_{\pm,j}^H$ do not depend on j; then,

$$\delta_\pm^{H,H'} = \begin{cases} (\delta^{H,L})_-^{-1}(\delta_\pm^{H,L})_+, & H' = L, \\ (\delta^{H,R})_-(\delta_\pm^{H,R})_+^{-1}, & H' = L, \end{cases}$$
$$\xrightarrow{T_\pm^H} \quad T_\pm^H \delta_\pm^{H,H'} = \begin{cases} (\delta_\pm^{H,L})_+(\delta^{H,L})_-^{-1}, & H' = L, \\ (\delta_\pm^{H,R})_+^{-1}(\delta^{H,R})_-, & H' = L. \end{cases}$$

It is clear that the shift corresponds to the flip of the factors in the Gaussian factorization of the $\delta_\pm^{H,H'}$ matrices, just as in the Darboux transformations. When the constant sequences $d_{\pm,j}^H = c_\pm^H \mathbb{I}_m$, with $c_\pm^H \in \mathbb{C}$ scalars, we have that $\delta_\pm^{H,H'}$ are pentadiagonal block matrices (main diagonal and the two next diagonals above and below it), and therefore the Gauss–Borel factorizations give upper or lower block tridiagonal matrices, $(\delta_\pm^{H,H'})_+$ and $(\delta_\pm^{H,H'})_-$, respectively. This is quite close to some results in the talk [28].

3.3. Miwa shifts

In our unsuccessful search for a neat τ-function theory in this matrix scenario, we have studied the action of Miwa shifts. Despite we did not find appropriate τ-functions, we found interesting relations among Christoffel–Darboux kernels and the Miwa transformations of the MOLPUC. These relations do in fact lead in the scalar case to the τ-function representation of MOLPUC. Unfortunately, apparently that is not the case in the matrix scenario.

Miwa shifts are coherent time translations that lead to discrete type flows. Given a diagonal matrix $w = \operatorname{diag}(w_1, \dots, w_m) \in \mathbb{C}^{m\times m}$, we introduce four different $\mathfrak{M}_w^{H,\pm}$, $H = L, R$, coherent shifts

$$\begin{aligned} \mathfrak{M}_w^{L,+} &: t_{2k}^L \mapsto t_{2k}^L - \frac{w^k}{k}, & \mathfrak{M}_w^{L,-} &: t_{2k-1}^L \mapsto t_{2k-1}^L - \frac{w^k}{k}, \\ \mathfrak{M}_w^{R,+} &: t_{2k}^R \mapsto t_{2k}^R - \frac{w^k}{k}, & \mathfrak{M}_w^{R,-} &: t_{2k-1}^R \mapsto t_{2k-1}^R - \frac{w^k}{k}. \end{aligned}$$

For each Miwa shift, we only write down those times with a non-trivial transformation. When these shifts act on the deformed matrix measure, we get new matrix measures

$$\mathrm{d}\mathfrak{M}_w^{L,\pm}[\mu] = \big(1 - wz^{\pm 1}\big)\mathrm{d}\mu, \qquad \mathrm{d}\mathfrak{M}_w^{R,\pm}[\mu] = \mathrm{d}\mu\big(1 - wz^{\pm 1}\big), \tag{49}$$

with corresponding left and right moment matrices given by

$$\begin{aligned} \mathfrak{M}_w^{L,\pm}\big[g^L\big] &= \big(\mathbb{I} - w\Upsilon^{\pm 1}\big)g^L, & \mathfrak{M}_w^{L,\pm}\big[g^R\big] &= \big(\mathbb{I} - w\Upsilon^{\mp 1}\big)g^R, \\ \mathfrak{M}_w^{R,\pm}\big[g^L\big] &= g^L\big(\mathbb{I} - w\Upsilon^{\pm 1}\big), & \mathfrak{M}_w^{R,\pm}\big[g^R\big] &= g^R\big(\mathbb{I} - w\Upsilon^{\mp 1}\big). \end{aligned} \tag{50}$$

From these we can deduce the next

Theorem 5. *For every diagonal matrix $w \in \operatorname{diag}_m$, the following relations between Miwa shifted and non-shifted Christoffel–Darboux kernels and MOLPUC hold*

$$\begin{aligned} K^{L,[2l+1]}(z,u) &= \mathfrak{M}_w^{L,+}\big[K^{L,[2l]}\big](z,u)(\mathbb{I} - wu) \\ &\quad + \big(\mathfrak{M}_w^{L,+}\big[\big(\varphi_2^L\big)^{(2l)}\big](z)\big)^\dagger \mathfrak{M}_w^{L,+}\big[h_{2l}^L\big]\big(h_{2l}^L\big)^{-1}\big(\varphi_1^L\big)^{(2l)}(u), \\ K^{R,[2l]}(z,u) &= \mathfrak{M}_w^{L,+}\big[K^{R,[2l-1]}\big](z,u)\big(\mathbb{I} - w\bar{u}^{-1}\big) \\ &\quad + \mathfrak{M}_w^{L,+}\big[\big(\varphi_1^R\big)^{(2l-1)}\big](z)\mathfrak{M}_w^{L,+}\big[h_{2l-1}^L\big]\big(h_{2l-1}^L\big)^{-1}\big(\big(\varphi_2^R\big)^{(2l-1)}(u)\big)^\dagger, \\ K^{L,[2l]}(z,u) &= \mathfrak{M}_w^{L,-}\big[K^{L,[2l-1]}\big](z,u)\big(\mathbb{I} - wu^{-1}\big) \\ &\quad + \big(\mathfrak{M}_w^{L,-}\big[\big(\varphi_2^L\big)^{(2l-1)}\big](z)\big)^\dagger \mathfrak{M}_w^{L,-}\big[h_{2l-1}^R\big]\big(h_{2l-1}^R\big)^{-1}\big(\varphi_1^L\big)^{(2l-1)}(u), \\ K^{R,[2l+1]}(z,u) &= \mathfrak{M}_w^{L,-}\big[K^{R,[2l]}\big](z,u)(\mathbb{I} - w\bar{u}) \\ &\quad + \mathfrak{M}_w^{L,-}\big[\big(\varphi_1^R\big)^{(2l)}\big](z)\mathfrak{M}_w^{L,-}\big[h_{2l}^L\big]\big(h_{2l}^L\big)^{-1}\big(\big(\varphi_2^R\big)^{(2l)}(u)\big)^\dagger, \\ K^{L,[2l]}(z,u) &= \big(\mathbb{I} - w\bar{z}^{-1}\big)\mathfrak{M}_w^{R,+}\big[K^{L,[2l-1]}\big](z,u) \\ &\quad + \big(\big(\varphi_2^L\big)^{(2l-1)}(z)\big)^\dagger \mathfrak{M}_w^{R,+}\big[\big(\varphi_1^L\big)^{(2l-1)}\big](u), \\ K^{R,[2l+1]}(z,u) &= (\mathbb{I} - wz)\mathfrak{M}_w^{R,+}\big[K^{R,[2l]}\big](z,u) \\ &\quad + \big(\varphi_1^R\big)^{(2l)}(z)\big(\mathfrak{M}_w^{R,+}\big[\big(\varphi_2^R\big)^{(2l)}\big](u)\big)^\dagger, \end{aligned}$$

$$
\begin{aligned}
K^{L,[2l+1]}(z,u) &= (\mathbb{I} - w\bar{z})\mathfrak{M}_w^{R,-}\big[K^{L,[2l]}\big](z,u) \\
&\quad + \big(\big(\varphi_2^L\big)^{(2l)}(z)\big)^\dagger \mathfrak{M}_w^{R,-}\big[\big(\varphi_1^L\big)^{(2l)}\big](u), \\
K^{R,[2l]}(z,u) &= \big(\mathbb{I} - wz^{-1}\big)\mathfrak{M}_w^{R,-}\big[K^{R,[2l-1]}\big](z,u) \\
&\quad + \big(\varphi_1^R\big)^{(2l-1)}(z)\big(\mathfrak{M}_w^{R,-}\big[\big(\varphi_2^R\big)^{(2l-1)}\big](u)\big)^\dagger.
\end{aligned}
$$

Proof. We just give the main ideas of the proof not dealing with details. Let us consider (50) at the light of the Gauss–Borel factorizations (7) and (8)

$$
\begin{aligned}
\mathfrak{M}_w^{L,\pm}[S_2]S_2^{-1} &= \mathfrak{M}_w^{L,\pm}[S_1]\big[\mathbb{I} - w\Upsilon^{\pm 1}\big]S_1^{-1}, \\
\big(\mathfrak{M}_w^{L,\pm}[Z_1]\big)^{-1}Z_1 &= \big(\mathfrak{M}_w^{L,\pm}[Z_2]\big)^{-1}\big[\mathbb{I} - w\Upsilon^{\mp 1}\big]Z_2, \\
S_1\big(\mathfrak{M}_w^{R,\pm}[S_1]\big)^{-1} &= S_2\big[\mathbb{I} - w\Upsilon^{\pm 1}\big]\big(\mathfrak{M}_w^{R,\pm}[S_2]\big)^{-1}, \\
Z_2^{-1}\mathfrak{M}_w^{R,\pm}[Z_2] &= Z_1^{-1}\big[\mathbb{I} - w\Upsilon^{\mp 1}\big]\mathfrak{M}_w^{R,\pm}[Z_1].
\end{aligned}
$$

Each of these equalities defines a semi-infinite matrix relating shifted and non-shifted polynomials. At this point it is important to stress that the LHS in the two first equations are upper triangular semi-infinite matrices, while the two last equations have in the RHS upper triangular semi-infinite matrices. Observe also that in the two first equations, because of the RHS only the main, the first and the second block diagonals over the first have non-zero blocks while in the LHS of the two last equations only the main diagonal and the two immediate diagonals below it have non-zero blocks. Then we proceed as in the proof of the Christoffel–Darboux formula in Theorem 2. To get a glance of the technique, let us illustrate it for the first equation. On the one hand, we have for the $2l$-th and $(2l+1)$-th block rows

$$
\begin{aligned}
&\mathfrak{M}_w^{L,+}[S_2]S_2^{-1} \\
&\quad = \mathfrak{M}_w^{L,+}[S_1][\mathbb{I} - w\Upsilon]S_1^{-1} \\
&\quad = \begin{pmatrix}
\ddots & \ddots & & & & & \\
\cdots & 0 & \mathfrak{M}_w^{L,+}[h_{2l}^L](h_{2l}^L)^{-1} & w\alpha_{1,2l+2}^L & -w & 0 & \cdots \\
\cdots & 0 & 0 & \mathfrak{M}_w^{L,+}[h_{2l+1}^R](h_{2l+1}^R)^{-1} & -(\mathfrak{M}_w^{L,+}[\alpha_{2,2l+1}^R])^\dagger w & * & 0 \\
 & & & & \ddots & &
\end{pmatrix}.
\end{aligned}
$$

On the other hand,

$$
\begin{aligned}
\big(\mathfrak{M}_w^{L,+}\big[\phi_2^L\big](z)\big)^\dagger \mathfrak{M}_w^{L,+}[S_2]S_2^{-1} &= \big(\phi_2^L(z)\big)^\dagger, \\
\mathfrak{M}_w^{L,+}[S_1][\mathbb{I} - w\Upsilon]S_1^{-1}\phi_1^L(u) &= \mathfrak{M}_w^{L,+}\big[\phi_1^L\big](u)(\mathbb{I} - wu).
\end{aligned}
$$

Then, by appropriate scalar product pairings, we get the result. $\square$

An appropriated choice of the variables allows us to express the rows or columns of the kernel in terms of the rows or columns of a product of a shifted and a non-shifted polynomial

Corollary 3. *If $w = \operatorname{diag}(w_1, \dots, w_m)$, $w_k \in \mathbb{C}$, we have*

$$
\begin{aligned}
&K^{L,[2l+1]}\big(z, w_k^{-1}\big)E_{k,k}\\
&\quad = \big(\mathfrak{M}_w^{L,+}\big[\big(\varphi_2^L\big)^{(2l)}\big](z)\big)^\dagger \mathfrak{M}_w^{L,+}\big[h_{2l}^L\big]\big(h_{2l}^L\big)^{-1}\big(\varphi_1^L\big)^{(2l)}\big(w_k^{-1}\big)E_{k,k},\\
&K^{R,[2l]}(z, \bar{w}_k)E_{k,k}\\
&\quad = \mathfrak{M}_w^{L,+}\big[\big(\varphi_1^R\big)^{(2l-1)}\big](z)\mathfrak{M}_w^{L,+}\big[h_{2l-1}^L\big]\big(h_{2l-1}^L\big)^{-1}\big(\big(\varphi_2^R\big)^{(2l-1)}(\bar{w}_k)\big)^\dagger E_{k,k},\\
&K^{L,[2l]}(z, w_k)E_{k,k}\\
&\quad = \big(\mathfrak{M}_w^{L,-}\big[\big(\varphi_2^L\big)^{(2l-1)}\big](z)\big)^\dagger \mathfrak{M}_w^{L,-}\big[h_{2l-1}^R\big]\big(h_{2l-1}^R\big)^{-1}\big(\varphi_1^L\big)^{(2l-1)}(w_k)E_{k,k},\\
&K^{R,[2l+1]}\big(z, \bar{w}_k^{-1}\big)E_{k,k}\\
&\quad = \mathfrak{M}_w^{L,-}\big[\big(\varphi_1^R\big)^{(2l)}\big](z)\mathfrak{M}_w^{L,-}\big[h_{2l}^L\big]\big(h_{2l}^L\big)^{-1}\big(\big(\varphi_2^R\big)^{(2l)}\big(\bar{w}_k^{-1}\big)\big)^\dagger E_{k,k},\\
&E_{k,k}K^{L;[2l]}(\bar{w}_k, u) = E_{k,k}\big(\big(\varphi_2^L\big)^{(2l-1)}(\bar{w}_k)\big)^\dagger \mathfrak{M}_w^{R,+}\big[\big(\varphi_1^L\big)^{(2l-1)}\big](u),\\
&E_{k,k}K^{R,[2l+1]}\big(w_k^{-1}, u\big) = E_{k,k}\big(\varphi_1^R\big)^{(2l)}\big(w_k^{-1}\big)\big(\mathfrak{M}_w^{R,+}\big[\big(\varphi_2^R\big)^{(2l)}\big](u)\big)^\dagger,\\
&E_{k,k}K^{L,[2l+1]}\big(\bar{w}_k^{-1}, u\big) = E_{k,k}\big(\big(\varphi_2^L\big)^{(2l)}\big(\bar{w}_k^{-1}\big)\big)^\dagger \mathfrak{M}_w^{R,-}\big[\big(\varphi_1^L\big)^{(2l)}\big](u),\\
&E_{k,k}K^{R,[2l]}(w_k, u) = E_{k,k}\big(\varphi_1^R\big)^{(2l-1)}(w_k)\big(\mathfrak{M}_w^{R,-}\big[\big(\varphi_2^R\big)^{(2l-1)}\big](u)\big)^\dagger.
\end{aligned}
$$

Let us consider what happens when instead of a diagonal matrix w is proportional to the identity matrix. In this case, (49) informs us that left and right handed Miwa shifts coincide. We only have two Miwa shifts $\mathfrak{M}_w^\pm$ where now $w \in \mathbb{C}$

$$
\mathrm{d}\mathfrak{M}_w^\pm[\mu] = \big(1 - wz^{\pm 1}\big)\mathrm{d}\mu. \tag{51}
$$

In this case, Corollary 3 would be written in much a simpler way (closer to the scalar case):

Proposition 32. *The following relations hold*

$$
\begin{aligned}
\big[\big(\varphi_2^L\big)^{(2l-1)}(\bar{w})\big]^\dagger \mathfrak{M}_w^+\big(h_{2l-1}^R\big) &= \big(\varphi_1^R\big)^{(2l)}\big(w^{-1}\big)h_{2l}^R,\\
\mathfrak{M}_w^-\big(h_{2l-1}^R\big)\big(h_{2l-1}^R\big)^{-1}\big(\varphi_1^L\big)^{(2l-1)}(w) &= \big[\big(\varphi_2^R\big)^{(2l)}\big(\bar{w}^{-1}\big)\big]^\dagger,\\
\big(\varphi_1^R\big)^{(2l-1)}(w)\mathfrak{M}_w^-\big(h_{2l-1}^L\big) &= \big[\big(\varphi_2^L\big)^{(2l)}\big(\bar{w}^{-1}\big)\big]^\dagger h_{2l}^L,\\
\mathfrak{M}_w^+\big(h_{2l-1}^L\big)\big(h_{2l-1}^L\big)^{-1}\big[\big(\varphi_2^R\big)^{(2l-1)}(\bar{w})\big]^\dagger &= \big(\varphi_1^L\big)^{(2l)}\big(w^{-1}\big),\\
\big(\varphi_1^R\big)^{(2l)}\big(w^{-1}\big)\mathfrak{M}_w^+\big(h_{2l}^R\big) &= \big[\bar{w}\big(\varphi_2^L\big)^{(2l+1)}(\bar{w})\big]^\dagger h_{2l+1}^R,
\end{aligned}
$$

$$\mathfrak{M}_w^-\big(h_{2l}^R\big)\big(h_{2l}^R\big)^{-1}\big[\big(\varphi_2^R\big)^{(2l)}\big(\bar{w}^{-1}\big)\big]^\dagger = z\big(\varphi_1^L\big)^{(2l+1)}(w),$$
$$\big[\big(\varphi_2^L\big)^{(2l)}\big(\bar{w}^{-1}\big)\big]^\dagger \mathfrak{M}_w^-\big(h_{2l}^L\big) = w\big(\varphi_1^R\big)^{(2l+1)}(z)h_{2l+1}^L,$$
$$\mathfrak{M}_w^+\big(h_{2l}^L\big)\big(h_{2l}^L\big)^{-1}\big(\varphi_1^L\big)^{(2l)}\big(w^{-1}\big) = \big[\bar{w}\big(\varphi_2^R\big)^{(2l+1)}(\bar{w})\big]^\dagger.$$

Proof. See Appendix A □

Now, we can state

Theorem 6. *The CMV matrix Laurent orthogonal polynomials can be expressed as follows*

$$\big(\varphi_1^L\big)^{(2l)}(z) = z^l\big[\mathfrak{M}_{z^{-1}}^+\big(h_{2l-1}^L\big)\big(h_{2l-1}^L\big)^{-1}\big]\cdots\big[\mathfrak{M}_{z^{-1}}^+\big(h_0^L\big)\big(h_0^L\big)^{-1}\big], \tag{52}$$

$$\big(\varphi_1^L\big)^{(2l+1)}(z) = z^{-(l+1)}\big[\mathfrak{M}_z^-\big(h_{2l}^R\big)\big(h_{2l}^R\big)^{-1}\big]\cdots\big[\mathfrak{M}_z^-\big(h_0^R\big)\big(h_0^R\big)^{-1}\big], \tag{53}$$

$$\big[\big(\varphi_2^L\big)^{(2l)}\big(\bar{z}^{-1}\big)\big]^\dagger = z^{-l}\big[\big(h_0^L\big)^{-1}\mathfrak{M}_z^-\big(h_0^L\big)\big]\cdots\big[\big(h_{2l-1}^L\big)\mathfrak{M}_z^-\big(h_{2l-1}^L\big)\big]\big(h_{2l}^L\big)^{-1}, \tag{54}$$

$$\big[\big(\varphi_2^L\big)^{(2l+1)}\big(\bar{z}^{-1}\big)\big]^\dagger = z^{l+1}\big[\big(h_0^R\big)^{-1}\mathfrak{M}_{z^{-1}}^+\big(h_0^R\big)\big]\cdots\big[\big(h_{2l}^R\big)^{-1}\mathfrak{M}_{z^{-1}}^+\big(h_{2l}^R\big)\big]\big(h_{2l+1}^R\big)^{-1}, \tag{55}$$

$$\big(\varphi_1^R\big)^{(2l)}(z) = z^l\big[\big(h_0^R\big)^{-1}\mathfrak{M}_{z^{-1}}^+\big({h_0}^R\big)\big]\cdots\big[\big(h_{2l-1}^R\big)^{-1}\mathfrak{M}_{z^{-1}}^+\big(h_{2l-1}^R\big)\big]\big(h_{2l}^R\big)^{-1}, \tag{56}$$

$$\big(\varphi_1^R\big)^{(2l+1)}(z) = z^{-(l+1)}\big[\big(h_0^L\big)^{-1}\mathfrak{M}_z^-\big(h_0^L\big)\big]\cdots\big[\big(h_{2l}^L\big)^{-1}\big]\mathfrak{M}_z^-\big(h_{2l}^L\big)\big(h_{2l+1}^L\big)^{-1}, \tag{57}$$

$$\big[\big(\varphi_2^R\big)^{(2l)}\big(\bar{z}^{-1}\big)\big]^\dagger = z^{-l}\big[\mathfrak{M}_z^-\big(h_{2l-1}^R\big)\big(h_{2l-1}^R\big)^{-1}\big]\cdots\big[\mathfrak{M}_z^-\big(h_0^R\big)\big(h_0^R\big)^{-1}\big], \tag{58}$$

$$\big[\big(\varphi_2^R\big)^{(2l+1)}\big(\bar{z}^{-1}\big)\big]^\dagger = z^{l+1}\big[\mathfrak{M}_{z^{-1}}^+\big(h_{2l}^L\big)\big(h_{2l}^L\big)^{-1}\big]\cdots\big[\mathfrak{M}_{z^{-1}}^+\big(h_0^L\big)\big(h_0^L\big)^{-1}\big]. \tag{59}$$

Proof. See Appendix A □

This is the furthest we have managed to take our τ description of the MOLPUC search. The reader may have noticed that forgetting about the R and L labels and the noncommutativity of the matrix norms we would be left with a quotient of Miwa shifted and non-shifted norms which in the scalar case coincides with the quotient of the determinants of the truncated Miwa shifted and non-shifted moment matrices.

Appendix A. Proofs

Proof of Proposition 1. Assuming $\det A \neq 0$ for any block matrix $M = \left(\begin{smallmatrix} A & B \\ C & D \end{smallmatrix}\right)$, we can write in terms of Schur complements

$$M = \begin{pmatrix} \mathbb{I} & 0 \\ CA^{-1} & \mathbb{I} \end{pmatrix}\begin{pmatrix} A & 0 \\ 0 & M/A \end{pmatrix}\begin{pmatrix} \mathbb{I} & A^{-1}B \\ 0 & \mathbb{I} \end{pmatrix}.$$

Thus, as g^H is given for a matrix quasi-definite measure

$$(g^H)^{[l+1]} = \left(\begin{array}{c|c} \mathbb{I}_{l\times l} & 0 \\ \hline v_{[l]} & \mathbb{I} \end{array}\right) \left(\begin{array}{c|c} (g^H)^{[l]} & 0 \\ \hline 0 & (g^H)^{[l+1]}/(g^H)^{[l]} \end{array}\right) \left(\begin{array}{c|c} \mathbb{I}_{l\times l} & w^{[l]} \\ \hline 0 & \mathbb{I} \end{array}\right),$$

where $v_{[l]} = (v_0, \dots, v_{l-1})$ and $w^{[l]} = \begin{pmatrix} w^0 \\ w^1 \\ \vdots \\ w^{l-1} \end{pmatrix}$ are two matrix vectors. Applying the same factorization to $(g^H)^{[l]}$, we get

$$(g^H)^{[l+1]} = \left(\begin{array}{c|cc} \mathbb{I}_{(l-1)\times(l-1)} & 0 & 0 \\ \hline r_{[l-1]} & \mathbb{I} & 0 \\ v'_{[l-1]} & * & \mathbb{I} \end{array}\right) \left(\begin{array}{c|cc} (g^H)^{[l-1]} & 0 & 0 \\ \hline 0 & (g^H)^{[l]}/(g^H)^{[l-1]} & 0 \\ 0 & 0 & (g^H)^{[l+1]}/(g^H)^{[l]} \end{array}\right)$$
$$\times \left(\begin{array}{c|cc} \mathbb{I}_{(l-1)\times(l-1)} & s^{[l-1]} & w'^{[l-1]} \\ \hline 0 & \mathbb{I} & * \\ 0 & 0 & \mathbb{I} \end{array}\right).$$

Finally, the iteration of these factorizations leads to

$$\begin{aligned}(g^H)^{[l+1]} &= \begin{pmatrix} \mathbb{I} & 0 & \dots & 0 \\ * & \mathbb{I} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ * & \dots & * & \mathbb{I} \end{pmatrix} \\ &\times \begin{pmatrix} (g^H)^{[1]}/(g^H)^{[0]} & 0 & \dots & 0 \\ 0 & (g^H)^{[2]}/(g^H)^{[1]} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & (g^H)^{[l+1]}/(g^H)^{[l]} \end{pmatrix} \\ &\times \begin{pmatrix} \mathbb{I} & * & \dots & * \\ 0 & \mathbb{I} & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \dots & 0 & \mathbb{I} \end{pmatrix},\end{aligned} \tag{A.1}$$

for $H = L, R$. Since this would have been valid for any l, it would also hold for the direct limit $\varinjlim (g^H)^{[l]}$; i.e., for g^H with $H = L, R$. □

Proof of Lemma 1. Notice is that the third equality of each expression is the second one written in terms of Schur complements. Therefore, just the first and second equalities of each expression need to be proven. The LU factorization leads to

$$S_1^{[\geq l,l]}(g^L)^{[l]} = -S_1^{[\geq l]}(g^L)^{[\geq l,l]},$$

$$S_2^{[l]}\big(\big(g^L\big)^{[l]}\big)^{-1}=S_1^{[l]}$$

from where the result follows immediately. As an illustration, let us derive the first expression; on the one hand,

$$\begin{aligned}\big(\varphi^L\big)_1^{(l)}(z)&=\sum_{j=0}^{l}(S_1)_{l,j}\chi^{(j)}\\&=\chi^{(l)}+\big(S_1^{[\geq l,l]}\chi^{[l]}\big)^{(0)}\\&=\chi^{(l)}-\sum_{m=0}^{l-1}\sum_{j=0}^{l}\big(\big(g^L\big)^{[\geq l,l]}\big)_{0,m}\big(\big(\big(g^L\big)^{[l]}\big)^{-1}\big)_{m,j}\chi^{(j)}\\&=\chi^{(l)}-\big((g^L)_{l,0}\quad(g^L)_{l,1}\quad\cdots\quad(g^L)_{l,l-1}\big)\big(\big(g^L\big)^{[l]}\big)^{-1}\chi^{[l]},\end{aligned}$$

and on the other,

$$\begin{aligned}\big(\varphi^L\big)_1^{(l)}(z)&=\sum_{m=0}^{l}\sum_{j=0}^{l}\big(S_2^{[l+1]}\big)_{l,m}\big(\big(\big(g^L\big)^{[l+1]}\big)^{-1}\big)_{m,j}\chi^{(j)}\\&=\big(S_2^{[l+1]}\big)_{l,l}\big(\big[\big(g^L\big)^{[l+1]}\big]^{-1}\big)_{l,j}\chi^{(j)}\\&=(S_2)_{ll}\,(0\quad 0\quad\ldots\quad 0\quad\mathbb{I})\,\big(\big(g^L\big)^{[l+1]}\big)^{-1}\chi^{[l+1]}.\end{aligned}$$

Proceeding in a similar manner, one gets all the other identities.[3] □

Proof of Theorem 2. We will only prove the first equation as the other three are proven in a similar way. In particular, we first prove the second equality of the first equation. We are interested in evaluating the expression

$$\big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^{\dagger}\big[\big(J^L\big)^{[2k]}\big]\big(\varphi_1^L\big)^{[2k]}\big(z'\big)$$

in two different ways. On the one hand, we could first let J act to the right. Truncating the expression

$$J^L\varphi_1^L(z)=z\varphi_1^L(z),$$

we have

[3] It is interesting to notice that in order to prove the right case expressions, once we have worked out the left ones, there is no need to go over the same calculations again. It is enough to realize that

$$\begin{aligned}&\varphi_1^R(\bar z)=\chi^{\dagger}(z)Z_1\text{ same structure as }\big[\varphi_2^L(z)\big]^{\dagger}=\chi^{\dagger}(z)S_2^{-1},\\&\varphi_1^L(z)=S_1\chi(z)\text{ same structure as }\big[\varphi_2^R(\bar z)\big]^{\dagger}=Z_2^{-1}\chi(z).\end{aligned}$$

$$\big[\big(J^L\big)^{[2k]}\big]\big(\varphi_1^L\big)^{[2k]}\big(z'\big) =$$
$$\begin{pmatrix} z'(\varphi_1^L)^{(0)}(z') \\ z'(\varphi_1^L)^{(1)}(z') \\ \vdots \\ -\alpha_{1,2k-1}^L h_{2k-2}^R [h_{2k-3}^R]^{-1}(\varphi_1^L)^{(2k-3)}(z') - \alpha_{1,2k-1}^L [\alpha_{2,2k-2}^R]^\dagger (\varphi_1^L)^{(2k-2)}(z') - \alpha_{1,2k}^L (\varphi_1^L)^{(2k-1)}(z') \\ h_{2k-1}^R [h_{2k-3}^R]^{-1}(\varphi_1^L)^{(2k-3)}(z') + h_{2k-1}^R [h_{2k-2}^R]^{-1}[\alpha_{2,2k-2}^R]^\dagger (\varphi_1^L)^{(2k-2)}(z') - [\alpha_{2,2k-1}^R]^\dagger \alpha_{1,2k}^L (\varphi_1^L)^{(2k-1)}(z') \end{pmatrix}.$$

But

$$\begin{aligned} z'\big(\varphi_1^L\big)^{(2k-2)}\big(z'\big) &= -\alpha_{1,2k-1}^L h_{2k-2}^R \big[h_{2k-3}^R\big]^{-1}\big(\varphi_1^L\big)^{(2k-3)}\big(z'\big) \\ &\quad - \alpha_{1,2k-1}^L \big[\alpha_{2,2k-2}^R\big]^\dagger \big(\varphi_1^L\big)^{(2k-2)}\big(z'\big) \\ &\quad - \alpha_{1,2k}^L \big(\varphi_1^L\big)^{(2k-1)}\big(z'\big) + \big(\varphi_1^L\big)^{(2k)}\big(z'\big) \\ z'\big(\varphi_1^L\big)^{(2k-1)}\big(z'\big) &= h_{2k-1}^R \big[h_{2k-3}^R\big]^{-1}\big(\varphi_1^L\big)^{(2k-3)}\big(z'\big) \\ &\quad + h_{2k-1}^R \big[h_{2k-2}^R\big]^{-1}\big[\alpha_{2,2k-2}^R\big]^\dagger \big(\varphi_1^L\big)^{(2k-2)}\big(z'\big) \\ &\quad - \big[\alpha_{2,2k-1}^R\big]^\dagger \alpha_{1,2k}^L \big(\varphi_1^L\big)^{(2k-1)}\big(z'\big) + \big[\alpha_{2,2k-1}^R\big]^\dagger \big(\varphi_1^L\big)^{(2k)}\big(z'\big), \end{aligned}$$

so that we obtain

$$\begin{aligned} &\big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^\dagger \big[J_L^{[2k]}\big]\big(\varphi_1^L\big)^{[2k]}\big(z'\big) \\ &\quad = \big(\big[\big(\varphi_2^L\big)^{(0)}(z)\big]^\dagger, \big[\big(\varphi_2^L\big)^{(1)}(z)\big]^\dagger, \dots, \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger\big) \\ &\qquad \cdot \begin{pmatrix} z'(\varphi_1^L)^{(0)}(z') \\ z'(\varphi_1^L)^{(1)}(z') \\ \vdots \\ z'(\varphi_1^L)^{(2k-2)}(z') - (\varphi_1^L)^{(2k)}(z') \\ z'(\varphi_1^L)^{(2k-1)}(z') - [\alpha_{2,2k-1}^R]^\dagger (\varphi_1^L)^{(2k)}(z') \end{pmatrix}. \end{aligned}$$

On the other hand, we could let J^L act to the left and remember that

$$\big[\big(\varphi_2^L\big)(z)\big]^\dagger J^L = \bar{z}^{-1}\big[\big(\varphi_2^L\big)(z)\big]^\dagger.$$

So, truncating the expression as we did when J^L acted to the right, we are left with

$$\begin{aligned} &\big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^\dagger \big(J^L\big)^{[2k]} \\ &\quad = \big(\bar{z}^{-1}\big[\big(\varphi_2^L\big)^{(0)}(z)\big]^\dagger, \dots, \bar{z}^{-1}\big[\big(\varphi_2^L\big)^{(2k-2)}(z)\big]^\dagger, \\ &\qquad \big[\big(\varphi_2^L\big)^{(2k-2)}(z)\big]^\dagger\big(-\alpha_{1,2k}^L\big) + \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger\big(-\big[\alpha_{2,2k-1}^R\big]^\dagger \alpha_{1,2k}^L\big)\big). \end{aligned}$$

But we also have

$$\begin{aligned}\bar{z}^{-1}\big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger &= \big[\big(\varphi_2^L\big)^{(2k-2)}(z)\big]^\dagger\big(-\alpha_{1,2k}^L\big) + \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger\big(-\big[\alpha_{2,2k-1}^R\big]^\dagger\alpha_{1,2k}^L\big)\\ &\quad + \big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger\big[-\alpha_{1,2k+1}^L h_{2k}^R\big(h_{2k-1}^R\big)^{-1}\big]\\ &\quad + \big[\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger\big[h_{2k+1}^R\big(h_{2k-1}^R\big)^{-1}\big].\end{aligned}$$

So, inserting it into the equation we are interested in, we have

$$\begin{aligned}&\big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^\dagger\big[\big(J^L\big)^{[2k]}\big]\big(\varphi_1^L\big)^{[2k]}\big(z'\big)\\ &= \big(\bar{z}^{-1}\big[\big(\varphi_2^L\big)^{(0)}(z)\big]^\dagger,\ldots,\bar{z}^{-1}\big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger\big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger\alpha_{1,2k+1}^L h_{2k}^R\big(h_{2k-1}^R\big)^{-1}\\ &\quad - \big[\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger h_{2k+1}^R\big(h_{2k-1}^R\big)^{-1}\big)\\ &\quad\times\begin{pmatrix}(\varphi_1^L)^{(0)}(z')\\ (\varphi_1^L)^{(1)}(z')\\ \vdots\\ (\varphi_1^L)^{(2k-1)}(z')\end{pmatrix}.\end{aligned}$$

Hence, we are left with the result we wanted to prove

$$\begin{aligned}&\bar{z}^{-1}\big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^\dagger\cdot\big(\varphi_1^L\big)^{[2k]}\big(z'\big) + \big[\big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger\alpha_{1,2k+1}^L h_{2k}^R\big(h_{2k-1}^R\big)^{-1}\\ &\qquad - \big[\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger h_{2k+1}^R\big(h_{2k-1}^R\big)^{-1}\big]\big(\varphi_1^L\big)^{(2k-1)}\big(z'\big)\\ &\quad = \big[\big(\varphi_2^L\big)^{[2k]}(z)\big]^\dagger\cdot z'\big(\varphi_1^L\big)^{[2k]}\big(z'\big) - \big[\big[\big(\varphi_2^L\big)^{(2k-2)}(z)\big]^\dagger\\ &\qquad + \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger\big[\alpha_{2,2k-1}^R\big]^\dagger\big]\big(\varphi_1^L\big)^{(2k)}\big(z'\big).\end{aligned}$$

Finally, the first equality in the first equation follows from the just proven result and Proposition 20. As was said at the beginning of this proof, the rest of the relations are proven in the exact same way. □

Proof of Proposition 26. First of all, we have

$$\begin{aligned}\frac{\partial}{\partial t_{j,a}^L}W_0\big(t^L\big) = \big[E_{a,a}\chi(\Upsilon)^{(j)}\big]W_0\big(t^L\big) \quad&\Longrightarrow\quad \partial_{L,j}W_0\big(t^L\big) = \chi(\Upsilon)^{(j)}W_0\big(t^L\big),\\ \frac{\partial}{\partial t_{j,a}^R}V_0\big(t^R\big) = V_0\big(t^R\big)\big[\chi\big(\Upsilon^{-1}\big)^{(j)}E_{a,a}\big] \quad&\Longrightarrow\quad \partial_{R,j}V_0\big(t^R\big) = V_0\big(t^R\big)\chi\big(\Upsilon^{-1}\big)^{(j)}.\end{aligned}$$

The previous derivatives make sense and are well defined since the two factors in the results commute. Hence, we have

$$\begin{aligned}\frac{\partial}{\partial t_{j,a}^L}W_1^L(t) &= \bigg[\bigg(\frac{\partial}{\partial t_{j,a}^L}S_1(t)\bigg)S_1(t)^{-1} + S_1(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_1(t)^{-1}\bigg]W_1^L(t),\\ \frac{\partial}{\partial t_{j,a}^R}W_1^L(t) &= \bigg[\bigg(\frac{\partial}{\partial t_{j,a}^R}S_1(t)\bigg)S_1(t)^{-1}\bigg]W_1^L(t),\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial t^R_{j,a}}W^L_2(t) &= \Big[\Big(\frac{\partial}{\partial t^R_{j,a}}S_2(t)\Big)S_2(t)^{-1} - S_2(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_2(t)^{-1}\Big]W^L_2(t),\\
\frac{\partial}{\partial t^L_{j,a}}W^L_2(t) &= \Big[\Big(\frac{\partial}{\partial t^L_{j,a}}S_2(t)\Big)S_2(t)^{-1}\Big]W^L_2(t),\\
\frac{\partial}{\partial t^L_{j,a}}W^R_2(t) &= W^R_2(t)\Big[Z_2(t)^{-1}\Big(\frac{\partial}{\partial t^L_{j,a}}Z_2(t)\Big) - Z_2(t)^{-1}\big[E_{a,a}\chi\big(\Upsilon^{-1}\big)^{(j)}\big]Z_2(t)\Big],\\
\frac{\partial}{\partial t^R_{j,a}}W^R_2(t) &= W^R_2(t)\Big[Z_2(t)^{-1}\Big(\frac{\partial}{\partial t^R_{j,a}}Z_2(t)\Big)\Big],\\
\frac{\partial}{\partial t^R_{j,a}}W^R_1(t) &= W^R_1(t)\Big[Z_1(t)^{-1}\Big(\frac{\partial}{\partial t^R_{j,a}}Z_1(t)\Big) + Z_1(t)^{-1}\big[\chi\big(\Upsilon^{-1}\big)^{(j)}E_{a,a}\big]Z_1(t)\Big],\\
\frac{\partial}{\partial t^L_{j,a}}W^R_1(t) &= W^R_1(t)\Big[Z_1(t)^{-1}\Big(\frac{\partial}{\partial t^L_{j,a}}Z_1(t)\Big)\Big].
\end{aligned}
$$

Now, if we let $\frac{\partial}{\partial t^H_{j,a}}$ act on both sides of the first expression in (42), we obtain

$$
\begin{aligned}
\Big(\frac{\partial}{\partial t^L_{j,a}}S_1(t)\Big)S_1(t)^{-1} + S_1(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_1(t)^{-1} &= \Big(\frac{\partial}{\partial t^L_{j,a}}S_2(t)\Big)S_2(t)^{-1},\\
\Big(\frac{\partial}{\partial t^R_{j,a}}S_2(t)\Big)S_2(t)^{-1} - S_2(t)\big[\chi(\Upsilon)^{(j)}E_{a,a}\big]S_2(t)^{-1} &= \Big(\frac{\partial}{\partial t^R_{j,a}}S_1(t)\Big)S_1(t)^{-1},
\end{aligned}
$$

which implies

$$
\begin{aligned}
\Big(\frac{\partial}{\partial t^L_{j,a}}S_2(t)\Big)S_2(t)^{-1} &= \big(S_1(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_1(t)^{-1}\big)_+,\\
\Big(\frac{\partial}{\partial t^L_{j,a}}S_1(t)\Big)S_1(t)^{-1} &= -\big(S_1(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_1(t)^{-1}\big)_-,\\
\Big(\frac{\partial}{\partial t^R_{j,a}}S_2(t)\Big)S_2(t)^{-1} &= \big(S_2(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_2(t)^{-1}\big)_+,\\
\Big(\frac{\partial}{\partial t^R_{j,a}}S_1(t)\Big)S_1(t)^{-1} &= -\big(S_2(t)\big[E_{a,a}\chi(\Upsilon)^{(j)}\big]S_2(t)^{-1}\big)_-.
\end{aligned}
$$

Similarly, let $\frac{\partial}{\partial t^H_{j,a}}$ act on both sides of the second expression in (42)

$$
\begin{aligned}
Z_2(t)^{-1}\Big(\frac{\partial}{\partial t^L_{j,a}}Z_2(t)\Big) - Z_2(t)^{-1}\big[E_{a,a}\chi\big(\Upsilon^{-1}\big)^{(j)}\big]Z_2(t) &= Z_1^{-1}\Big(\frac{\partial}{\partial t^L_{j,a}}Z_1(t)\Big),\\
Z_1(t)^{-1}\Big(\frac{\partial}{\partial t^R_{j,a}}Z_1(t)\Big) + Z_1(t)^{-1}\big[\chi\big(\Upsilon^{-1}\big)^{(j)}E_{a,a}\big]Z_1(t) &= Z_2^{-1}\Big(\frac{\partial}{\partial t^R_{j,a}}Z_2(t)\Big).
\end{aligned}
$$

This means

$$Z_2(t)^{-1}\left(\frac{\partial}{\partial t^L_{j,a}}Z_2(t)\right)=\left(Z_2(t)^{-1}\left[E_{a,a}\chi\left(\Upsilon^{-1}\right)^{(j)}\right]Z_2(t)\right)_-,$$
$$Z_1(t)^{-1}\left(\frac{\partial}{\partial t^L_{j,a}}Z_1(t)\right)=-\left(Z_2(t)^{-1}\left[E_{a,a}\chi\left(\Upsilon^{-1}\right)^{(j)}\right]Z_2(t)\right)_+,$$
$$Z_2(t)^{-1}\left(\frac{\partial}{\partial t^R_{j,a}}Z_2(t)\right)=\left(Z_1(t)^{-1}\left[E_{a,a}\chi\left(\Upsilon^{-1}\right)^{(j)}\right]Z_1(t)\right)_-,$$
$$Z_1(t)^{-1}\left(\frac{\partial}{\partial t^R_{j,a}}Z_1(t)\right)=-\left(Z_1(t)^{-1}\left[E_{a,a}\chi\left(\Upsilon^{-1}\right)^{(j)}\right]Z_1(t)\right)_+.$$

With all these results, it is easy to prove both the linear systems for the wave functions and the Lax equations. For the flows of the intertwining operators, we use these relations together with (41), the first expression for the right times and the second one for the left times; then just recall (40). Finally, the Zakharov–Shabat equations are just the compatibility conditions of the Lax equations. □

Proof of Proposition 32. First we use (19) and apply a Miwa shift

$$\mathfrak{M}^+_w\left(h^R_{2l-1}\right)=\oint_{\mathbb{T}}\mathfrak{M}^+_w\left(\left(\varphi^L_1\right)^{(2l-1)}\right)(u)\frac{\mathrm{d}\mathfrak{M}^+_w(\mu)(u)}{\mathrm{i}u}u^l.$$

Second, from Theorem 5 ($w_k=w$ for $k=1,\dots,m$), we get

$$\left(\left[\left(\varphi^L_2\right)^{(2l-1)}(\bar w)\right]^\dagger\right)^{-1}K^{L,[2l]}(\bar w,u)=\mathfrak{M}^+_w\left[\left(\varphi^L_1\right)^{(2l-1)}\right](u),$$

so that

$$\begin{aligned}
&\left[\left(\varphi^L_2\right)^{(2l-1)}(\bar w)\right]^\dagger\mathfrak{M}^+_w\left(h^R_{2l-1}\right)\\
&\quad=\oint_{\mathbb{T}}K^{L,[2l]}(\bar w,u)\frac{\mathrm{d}\mathfrak{M}^+_w(\mu)(u)}{\mathrm{i}u}u^l\\
&\quad=\oint_{\mathbb{T}}K^{L,[2l]}(\bar w,u)(1-wu)\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}u^l\quad\text{(by (51))}\\
&\quad=\oint_{\mathbb{T}}\Big(\left(\varphi^R_1\right)^{(2l)}\left(w^{-1}\right)h^R_{2l}\left(h^R_{2l-1}\right)^{-1}\left(\varphi^L_1\right)^{(2l-1)}(u)\\
&\qquad-\left(\varphi^R_1\right)^{(2l-1)}\left(w^{-1}\right)\left(\varphi^L_1\right)^{(2l)}(u)\Big)\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}u^l\quad\text{(by (31))}\\
&\quad=\oint_{\mathbb{T}}\left(\varphi^R_1\right)^{(2l)}\left(w^{-1}\right)h^R_{2l}h^R_{2l-1}\left(\varphi^L_1\right)^{(2l-1)}(u)\frac{\mathrm{d}\mu(u)}{\mathrm{i}u}u^l\quad\text{(by (15))}\\
&\quad=\left(\varphi^R_1\right)^{(2l)}\left(w^{-1}\right)h^R_{2l}\quad\text{(by (19))}.
\end{aligned}$$

This same procedure applies for the proof of the remaining formulae. □

Proof of Theorem 6. Since we can take any value for w_1, w_2, let us consider them our variables and name them $w_1 = w_2 = z$. Now by iteration of the formulae in Proposition 32, we get

$$
\begin{aligned}
\big[\big(\varphi_2^R\big)^{(2l+1)}\big(\bar z^{-1}\big)\big]^\dagger &= \mathfrak{M}^+_{z^{-1}}\big(h^L_{2l}\big)\big(h^L_{2l}\big)^{-1}\cdots\mathfrak{M}^+_{z^{-1}}\big(h^L_1\big)\big(h^L_1\big)^{-1}\big[\big(\varphi_2^R\big)^{(1)}\big(\bar z^{-1}\big)\big]^\dagger z^l,\\
\big(\varphi_1^L\big)^{(2l)}(z) &= \mathfrak{M}^+_{z^{-1}}\big(h^L_{2l-1}\big)\big(h^L_{2l-1}\big)^{-1}\cdots\mathfrak{M}^+_{z^{-1}}\big(h^L_1\big)\big(h^L_1\big)^{-1}\mathfrak{M}^+_{z^{-1}}\big(h^L_0\big)\big(h^L_0\big)^{-1}z^l,\\
\big[\big(\varphi_2^L\big)^{(2l+1)}\big(\bar z^{-1}\big)\big]^\dagger &= z^l\big[\big(\varphi_2^L\big)^{(1)}\big(\bar z^{-1}\big)\big]^\dagger\mathfrak{M}^+_{z^{-1}}\big(h^R_1\big)\cdots\big(h^R_{2l}\big)^{-1}\mathfrak{M}^+_{z^{-1}}\big(h^R_{2l}\big)\big(h^R_{2l+1}\big)^{-1},\\
\big(\varphi_1^R\big)^{(2l)}(z) &= z^l\big[h^R_0\big]^{-1}\mathfrak{M}^z_2\big(h^R_0\big)\big(h^R_1\big)^{-1}\mathfrak{M}^+_{z^{-1}}\big(h^R_1\big)\cdots\big(h^R_{2l-2}\big)^{-1}\mathfrak{M}^z_2\big(h^R_{2l-2}\big)\\
&\quad\times\big(h^R_{2l-1}\big)^{-1}\mathfrak{M}^z_2\big(h^R_{2l-1}\big)\big(h^R_{2l}\big)^{-1},\\
\big(\varphi_1^L\big)^{(2l+1)}(z) &= \mathfrak{M}^-_z\big(h^R_{2l}\big)\big(h^R_{2l}\big)^{-1}\mathfrak{M}^-_z\big(h^R_{2l-1}\big)\big(h^R_{2l-1}\big)^{-1}\cdots\mathfrak{M}^-_z\big(h^R_1\big)\\
&\quad\times\big(h^R_1\big)^{-1}\big(\varphi_1^L\big)^{(1)}(z)z^{-l},\\
\big[\big(\varphi_2^R\big)^{(2l)}\big(\bar z^{-1}\big)\big]^\dagger &= \mathfrak{M}^-_z\big(h^R_{2l-1}\big)\big(h^R_{2l-1}\big)^{-1}\cdots\mathfrak{M}^-_z\big(h^R_0\big)\big(h^R_0\big)^{-1}z^{-l},\\
\big(\varphi_1^R\big)^{(2l+1)}(z) &= z^{-l}\big(\varphi_1^R\big)^{(1)}(z)\mathfrak{M}^-_z\big(h^L_1\big)\cdots\big(h^L_{2l-1}\big)^{-1}\mathfrak{M}^-_z\big(h^L_{2l-1}\big)\\
&\quad\times\big(h^L_{2l}\big)^{-1}\mathfrak{M}^-_z\big(h^L_{2l}\big)\big(h^L_{2l+1}\big)^{-1},\\
\big[\big(\varphi_2^L\big)^{(2l)}\big(\bar z^{-1}\big)\big]^\dagger &= z^{-l}\big(h^L_0\big)^{-1}\mathfrak{M}^-_z\big(h^L_0\big)\cdots\big(h^L_{2l-1}\big)^{-1}\mathfrak{M}^-_z\big(h^L_{2l-1}\big)\big(h^L_{2l}\big)^{-1}.
\end{aligned}
$$

Finally, noticing that

$$
\begin{aligned}
\mathfrak{M}^+_{z^{-1}}\big(h^L_0\big) = \mathfrak{M}^+_{z^{-1}}\big(h^R_0\big) &= \mathbb{I} - \frac{1}{z}\big(g^L\big)_{01} = \mathbb{I} - \frac{1}{z}\big(g^R\big)_{10},\\
\mathfrak{M}^-_z\big(h^L_0\big) = \mathfrak{M}^-_z\big(h^R_0\big) &= \mathbb{I} - z\big(g^L\big)_{10} = \mathbb{I} - z\big(g^R\big)_{01},
\end{aligned}
$$

we get

$$
\begin{aligned}
\big[\big(\varphi_2^R\big)^{(1)}\big(\bar z^{-1}\big)\big]^\dagger &= z\left(\mathbb{I} - \frac{1}{z}\big(g^R\big)_{10}\right) = z\mathfrak{M}^+_{z^{-1}}\big(h^L_0\big),\\
\big[\big(\varphi_2^L\big)^{(1)}\big(\bar z^{-1}\big)\big]^\dagger &= z\left(\mathbb{I} - \frac{1}{z}\big(g^L\big)_{01}\right)\big(h^R_1\big)^{-1} = z\mathfrak{M}^+_{z^{-1}}\big(h^R_0\big)\big(h^R_1\big)^{-1},\\
\big(\varphi_1^L\big)^{(1)}(z) &= \frac{1}{z}\big(\mathbb{I} - z\big(g^L\big)_{10}\big) = \frac{1}{z}\mathfrak{M}^-_z\big(h^R_0\big),\\
\big(\varphi_1^R\big)^{(1)}(z) &= \frac{1}{z}\big(\mathbb{I} - z\big(g^R\big)_{01}\big)\big(h^L_1\big)^{-1} = \frac{1}{z}\mathfrak{M}^-_z\big(h^L_0\big)\big(h^L_1\big)^{-1},
\end{aligned}
$$

and the result is proven. □

Appendix B. Explicit coefficients of J and C

Proposition 33. *The following expressions correspond to the block non-zero elements of* $(J^H)^{\pm 1}$*:*

$$\begin{aligned}
(J^L)_{2k,2k-1} &= -h^L_{2k}\alpha^R_{1,2k+1}(h^R_{2k-1})^{-1}, & (J^R)_{2k,2k-1} &= h^R_{2k}[\alpha^L_{2,2k+1}]^\dagger(h^L_{2k-1})^{-1},\\
(J^L)_{2k,2k} &= -h^L_{2k}\alpha^R_{1,2k+1}[\alpha^L_{2,2k}]^\dagger(h^L_{2k})^{-1}, & (J^R)_{2k,2k} &= -h^R_{2k}[\alpha^L_{2,2k+1}]^\dagger\alpha^R_{1,2k}(h^R_{2k})^{-1},\\
(J^L)_{2k,2k+1} &= -\alpha^L_{1,2k+2}, & (J^R)_{2k,2k+1} &= -[\alpha^R_{2,2k+2}]^\dagger,\\
(J^L)_{2k,2k+2} &= \mathbb{I}, & (J^R)_{2k,2k+2} &= \mathbb{I},\\
(J^L)_{2k+1,2k-1} &= h^R_{2k+1}(h^R_{2k-1})^{-1}, & (J^R)_{2k+1,2k-1} &= h^L_{2k+1}(h^L_{2k-1})^{-1},\\
(J^L)_{2k+1,2k} &= h^R_{2k+1}[\alpha^L_{2,2k}]^\dagger(h^L_{2k})^{-1}, & (J^R)_{2k+1,2k} &= h^L_{2k+1}\alpha^R_{1,2k}(h^R_{2k})^{-1},\\
(J^L)_{2k+1,2k+1} &= -[\alpha^R_{2,2k+1}]^\dagger\alpha^L_{1,2k+2}, & (J^R)_{2k+1,2k+1} &= -\alpha^L_{1,2k+1}[\alpha^R_{2,2k+2}]^\dagger,\\
(J^L)_{2k+1,2k+2} &= [\alpha^R_{2,2k+1}]^\dagger, & (J^R)_{2k+1,2k+2} &= \alpha^L_{1,2k+1},
\end{aligned}$$

$$\begin{aligned}
(J^L)_{0,0} &= -h^L_0\alpha^R_{1,1}(h^L_0)^{-1}, & (J^R)_{0,0} &= -h^R_0[\alpha^L_{2,1}]\dagger(h^R_0)^{-1},\\
(J^L)_{0,1} &= -\alpha^L_{1,2}, & (J^R)_{0,1} &= -[\alpha^R_{2,2}]^\dagger,\\
(J^L)_{0,2} &= \mathbb{I}, & (J^R)_{0,2} &= \mathbb{I},\\
(J^L)_{1,0} &= h^R_1(h^L_0)^{-1}, & (J^R)_{1,0} &= h^L_1(h^R_0)^{-1},\\
(J^L)_{1,1} &= -[\alpha^R_{1,2}]^\dagger\alpha^L_{1,2}, & (J^R)_{1,1} &= -\alpha^L_{1,1}[\alpha^R_{2,2}]^\dagger,\\
(J^L)_{1,2} &= [\alpha^R_{1,2}]^\dagger, & (J^R)_{1,2} &= \alpha^L_{1,2},
\end{aligned}$$

$$\begin{aligned}
&((J^L)^{-1})_{2k-1,2k} = -[\alpha^R_{2,2k+1}]^\dagger, \qquad ((J^R)^{-1})_{2k-1,2k} = -\alpha^R_{1,2k+1},\\
&((J^L)^{-1})_{2k,2k} = -\alpha^L_{1,2k}[\alpha^R_{2,2k+1}]^\dagger, \qquad ((J^R)^{-1})_{2k,2k} = -[\alpha^R_{2,2k}]^\dagger\alpha^L_{1,2k+1},\\
&((J^L)^{-1})_{2k+1,2k} = -h^R_{2k+1}[\alpha^L_{2,2k+2}]^\dagger(h^L_{2k})^{-1},\\
&\qquad ((J^R)^{-1})_{2k+1,2k} = -h^L_{2k+1}\alpha^R_{1,2k+2}(h^R_{2k})^{-1},\\
&((J^L)^{-1})_{2k+2,2k} = (h^L)_{2k+2}(h^L_{2k})^{-1}, \qquad ((J^R)^{-1})_{2k+2,2k} = (h^R)_{2k+2}(h^R_{2k})^{-1},\\
&((J^L)^{-1})_{2k-1,2k+1} = \mathbb{I}, \qquad ((J^R)^{-1})_{2k-1,2k+1} = \mathbb{I},\\
&((J^L)^{-1})_{2k,2k+1} = \alpha^L_{1,2k}, \qquad ((J^R)^{-1})_{2k,2k+1} = [\alpha^R_{2,2k}]^\dagger,
\end{aligned}$$

$$
\begin{aligned}
\big((J^L)^{-1}\big)_{2k+1,2k+1} &= -h^R_{2k+1}\big[\alpha^L_{2,2k+2}\big]^\dagger\alpha^R_{1,2k+1}\big(h^R_{2k+1}\big)^{-1},\\
\big((J^R)^{-1}\big)_{2k+1,2k+1} &= -h^L_{2k+1}\alpha^R_{1,2k+2}\big[\alpha^L_{2,2k+1}\big]^\dagger\big(h^L_{2k+1}\big)^{-1},\\
\big((J^L)^{-1}\big)_{2k+2,2k+1} &= h^L_{2k+2}\alpha^R_{1,2k+1}\big(h^R_{2k+1}\big)^{-1},\\
\big((J^R)^{-1}\big)_{2k+2,2k+1} &= \big(h^R\big)_{2k+2}\big[\alpha^L_{2,2k+1}\big]^\dagger\big(h^L_{2k+1}\big)^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\big((J^L)^{-1}\big)_{0,0} &= -\big[\alpha^R_{2,1}\big]^\dagger, & \big((J^R)^{-1}\big)_{0,0} &= -\alpha^L_{1,1},\\
\big((J^L)^{-1}\big)_{1,0} &= -h^R_1\big[\alpha^L_{2,2}\big]^\dagger\big(h^L_0\big)^{-1}, & \big((J^R)^{-1}\big)_{1,0} &= -h^L_1\alpha^R_{1,2}\big(h^R_0\big)^{-1},\\
\big((J^L)^{-1}\big)_{2,0} &= h^L_2\big(h^L_0\big)^{-1}, & \big((J^R)^{-1}\big)_{2,0} &= h^R_2\big(h^R_0\big)^{-1},\\
\big((J^L)^{-1}\big)_{0,1} &= \mathbb{I}, & \big((J^R)^{-1}\big)_{0,1} &= \mathbb{I},\\
\big((J^L)^{-1}\big)_{1,1} &= -h^R_1\big[\alpha^L_{2,2}\big]^\dagger\alpha^R_{1,1}\big(h^R_1\big)^{-1}, & \big((J^R)^{-1}\big)_{1,1} &= -h^L_1\alpha^R_{1,2}\big[\alpha^L_{2,1}\big]^\dagger\big(h^L_1\big)^{-1},\\
\big((J^L)^{-1}\big)_{2,1} &= h^L_2\alpha^R_{1,1}\big(h^R_1\big)^{-1}, & \big((J^R)^{-1}\big)_{2,1} &= h^R_2\big[\alpha^L_{2,1}\big]^\dagger\big(h^L_1\big)^{-1}.
\end{aligned}
$$

Proposition 34. *The following expressions correspond to the block non-zero elements of $C^{[\pm 1]}_{[0]}$:*

$$
\begin{aligned}
(C_{[0]})_{2k,2k-1} &= h^R_{2k}\big[\big(h^R\big)_{2k-1}\big]^{-1} = \mathbb{I} - \big[\alpha^R_{2,2k}\big]^\dagger\alpha^L_{1,2k},\\
(C_{[0]})_{2k,2k} &= \big[\alpha^R_{2,2k}\big]^\dagger = h^R_{2k}\big[\alpha^L_{2,2k}\big]^\dagger\big[h^L_{2k}\big]^{-1},\\
(C_{[0]})_{2k+1,2k+1} &= -\alpha^L_{1,2k+2} = -h^L_{2k+1}\alpha^R_{1,2k+2}\big[h^R_{2k+1}\big]^{-1},\\
(C_{[0]})_{2k+1,2k+2} &= \mathbb{I} = h^L_{2k+1}\big[\mathbb{I} - \alpha^R_{1,2k+2}\big[\alpha^L_{2,2k+2}\big]^\dagger\big]\big[\big(h^L\big)_{2k+2}\big]^{-1},
\end{aligned}
$$

$$
\begin{aligned}
\big(C^{-1}_{[0]}\big)_{2k,2k-1} &= h^L_{2k}\big[\big(h^L\big)_{2k-11}\big]^{-1} = \mathbb{I} - \alpha^L_{1,2k}\big[\alpha^R_{2,2k}\big]^\dagger,\\
\big(C^{-1}_{[0]}\big)_{2k,2k} &= \alpha^L_{1,2k} = h^L_{2k}\alpha^R_{1,2k}\big[h^R_{2k}\big]^{-1},\\
\big(C^{-1}_{[0]}\big)_{2k+1,2k+1} &= -\big[\alpha^R_{2,2k+2}\big]^\dagger = -h^R_{2k+1}\big[\alpha^L_{2,2k+2}\big]^\dagger\big[h^L_{2k+1}\big]^{-1},\\
\big(C^{-1}_{[0]}\big)_{2k+1,2k+2} &= \mathbb{I} = h^R_{2k+1}\big[\mathbb{I} - \big[\alpha^L_{2,2k+2}\big]^\dagger\alpha^R_{1,2k+2}\big]\big[h^R_{2k+2}\big]^{-1}.
\end{aligned}
$$

Proposition 35. *The following expressions correspond to the block non-zero elements of $C^{[\pm 1]}_{[-1]}$:*

$$
\begin{aligned}
(C_{[-1]})_{2k,2k} &= -\big[\alpha^R_{2,2k+1}\big]^\dagger = -h^R_{2k}\big[\alpha^L_{2,2k+1}\big]^\dagger\big[h^L_{2k}\big]^{-1},\\
(C_{[-1]^{-1}})_{2k,2k} &= -\alpha^L_{1,2k+1} = -h^L_{2k}\alpha^L_{2,2k+1}\big[h^R_{2k}\big]^{-1},\\
(C_{[-1]})_{2k,2k+1} &= \mathbb{I} \quad \big(C^{-1}_{[-1]}\big)_{2k,2k+1} = \mathbb{I},
\end{aligned}
$$

$$
\begin{aligned}
(C_{[-1]})_{2k+1,2k} &= \mathbb{I} - \alpha^L_{1,2k+1}\big[\alpha^R_{2,2k+1}\big]^\dagger = h^L_{2k+1}\big[h^L_{2k}\big]^{-1},\\
\big(C^{-1}_{[-1]}\big)_{2k+1,2k} &= \mathbb{I} - \big[\alpha^R_{2,2k+1}\big]^\dagger\alpha^L_{1,2k+1} = h^R_{2k+1}\big[h^R_{2k}\big]^{-1},\\
(C_{[-1]})_{2k+1,2k+1} &= \alpha^L_{1,2k+1} = h^L_{2k+1}\alpha^R_{1,2k+1}\big[h^R_{2k+1}\big]^{-1},\\
\big(C^{-1}_{[-1]}\big)_{2k+1,2k+1} &= \big[\alpha^R_{2,2k+1}\big]^\dagger = h^R_{2k+1}\big[\alpha^L_{2,2k+1}\big]^\dagger\big[h^L_{2k+1}\big]^{-1}.
\end{aligned}
$$

Appendix C. Complete recursion relations

Here we give a more complete set of recursion relations for the MOLPUC.

Proposition 36. *The five term CMV recursion relations are*

$$
\begin{aligned}
z\big(\varphi^L_1\big)^{(2k)}(z) &= -\alpha^L_{1,2k+1}\big(\mathbb{I} - \big[\alpha^R_{2,2k}\big]^\dagger\alpha^L_{1,2k}\big)\big(\varphi^L_1\big)^{(2k-1)} - \alpha^L_{1,2k+1}\big[\alpha^R_{2,2k}\big]^\dagger\big(\varphi^L_1\big)^{(2k)}\\
&\quad - \alpha^L_{1,2k+2}\big(\varphi^L_1\big)^{(2k+1)} + \big(\varphi^L_1\big)^{(2k+2)},\\
z\big(\varphi^L_1\big)^{(2k+1)}(z) &= \big(\mathbb{I} - \big[\alpha^R_{2,2k+1}\big]^\dagger\alpha^L_{1,2k+1}\big)\big(\mathbb{I} - \big[\alpha^R_{2,2k}\big]^\dagger\alpha^L_{1,2k}\big)\big(\varphi^L_1\big)^{(2k-1)}\\
&\quad + \big(\mathbb{I} - \big[\alpha^R_{2,2k+1}\big]^\dagger\alpha^L_{1,2k+1}\big)\big[\alpha^R_{2,2k}\big]^\dagger\big(\varphi^L_1\big)^{(2k)}\\
&\quad - \big[\alpha^R_{2,2k+1}\big]^\dagger\alpha^L_{1,2k+2}\big(\varphi^L_1\big)^{(2k+1)} + \big[\alpha^R_{2,2k+1}\big]^\dagger\big(\varphi^L_1\big)^{(2k+2)},\\
z\big(\varphi^L_1\big)^{(0)}(z) &= -\alpha^R_{1,1}\big(\varphi^L_1\big)^{(0)} - \alpha^L_{1,2}\big(\varphi^L_1\big)^{(1)} + \big(\varphi^L_1\big)^{(2)},\\
z\big(\varphi^L_1\big)^{(1)}(z) &= \big(\mathbb{I} - \big[\alpha^R_{2,1}\big]^\dagger\alpha^L_{1,1}\big)\big(\varphi^L_1\big)^{(0)} - \big[\alpha^R_{2,1}\big]^\dagger\alpha^L_{1,2}\big(\varphi^L_1\big)^{(1)} + \big[\alpha^R_{2,1}\big]^\dagger\big(\varphi^L_1\big)^{(2)},
\end{aligned}
$$

$$
\begin{aligned}
z^{-1}\big(\varphi^L_1\big)^{(2k)}(z) &= \big(\mathbb{I} - \alpha^L_{1,2k}\big[\alpha^R_{2,2k}\big]^\dagger\big)\big(\mathbb{I} - \alpha^L_{1,2k-1}\big[\alpha^R_{2,2k-1}\big]^\dagger\big)\big(\varphi^L_1\big)^{(2k-2)}\\
&\quad + \big(\mathbb{I} - \alpha^L_{1,2k}\big[\alpha^R_{2,2k}\big]^\dagger\big)\alpha^L_{1,2k-1}\big(\varphi^L_1\big)^{(2k-1)} - \alpha^L_{1,2k}\big[\alpha^R_{2,2k+1}\big]^\dagger\big(\varphi^L_1\big)^{(2k)}\\
&\quad + \big[\alpha^L_{1,2k}\big]^\dagger\big(\varphi^L_1\big)^{(2k+1)},\\
z^{-1}\big(\varphi^L_1\big)^{(2k+1)}(z) &= -\big[\alpha^R_{2,2k+2}\big]^\dagger\big(\mathbb{I} - \alpha^L_{1,2k+1}\big[\alpha^R_{2,2k+1}\big]^\dagger\big)\big(\varphi^L_1\big)^{(2k)}\\
&\quad - \big[\alpha^R_{2,2k+2}\big]^\dagger\alpha^L_{1,2k+1}\big(\varphi^L_1\big)^{(2k+1)} - \big[\alpha^R_{2,2k+3}\big]^\dagger\big(\varphi^L_1\big)^{(2k+2)}\\
&\quad + \big(\varphi^L_1\big)^{(2k+3)},\\
z^{-1}\big(\varphi^L_1\big)^{(0)}(z) &= -\big[\alpha^R_{2,1}\big]^\dagger\big(\varphi^L_1\big)^{(0)} + \big(\varphi^L_1\big)^{(1)},
\end{aligned}
$$

$$
\begin{aligned}
\big[z\big(\varphi^L_2\big)^{(2k)}(z)\big]^\dagger &= -\big[\big(\varphi^L_2\big)^{(2k-1)}(z)\big]^\dagger\big[\alpha^R_{2,2k+1}\big]^\dagger - \big[\big(\varphi^L_2\big)^{(2k)}\big]^\dagger\alpha^L_{1,2k}\big[\alpha^R_{2,2k+1}\big]^\dagger\\
&\quad - \big[\big(\varphi^L_2\big)^{(2k+1)}(z)\big]^\dagger\big[\alpha^R_{2,2k+2}\big]^\dagger\big(\mathbb{I} - \alpha^L_{1,2k+1}\big[\alpha^R_{2,2k+1}\big]^\dagger\big)\\
&\quad + \big[\big(\varphi^L_2\big)^{(2k+2)}(z)\big]^\dagger\big(\mathbb{I} - \alpha^L_{1,2k+2}\big[\alpha^R_{2,2k+2}\big]^\dagger\big)\big(\mathbb{I} - \alpha^L_{1,2k+1}\big[\alpha^R_{2,2k+1}\big]^\dagger\big),
\end{aligned}
$$

$$\begin{aligned}
\big[z\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger ={}& \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger + \big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger \alpha_{1,2k}^L \\
&- \big[\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger \big[\alpha_{2,2k+2}^R\big]^\dagger \alpha_{1,2k+1}^L \\
&+ \big[\big(\varphi_2^L\big)^{(2k+2)}(z)\big]^\dagger \big(\mathbb{I} - \alpha_{1,2k+2}^L \big[\alpha_{2,2k+2}^R\big]^\dagger\big)\alpha_{1,2k+1}^L,
\end{aligned}$$

$$\begin{aligned}
\big[z^{-1}\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger ={}& \big[\big(\varphi_2^L\big)^{(2k-2)}(z)\big]^\dagger - \big[\big(\varphi_2^L\big)^{(2k-1)}(z)\big]^\dagger \big[\alpha_{2,2k-1}^R\big]^\dagger \\
&- \big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger \alpha_{1,2k+1}^L \big[\alpha_{2,2k}^R\big]^\dagger \\
&+ \big[\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^\dagger \alpha_{1,2k+1}^L\big)\big[\alpha_{2,2k}^R\big]^\dagger, \\
\big[z^{-1}\big(\varphi_2^L\big)^{(2k+1)}(z)\big]^\dagger ={}& -\big[\big(\varphi_2^L\big)^{(2k)}(z)\big]^\dagger \alpha_{1,2k+2}^L - \big[\big(\varphi_2^L\big)^{(2k+1)}\big]^\dagger \big[\alpha_{2,2k+1}^R\big]^\dagger \alpha_{1,2k+2}^L \\
&- \big[\big(\varphi_2^L\big)^{(2k+2)}(z)\big]^\dagger \alpha_{1,2k+3}^L \big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^\dagger \alpha_{1,2k+2}^L\big) \\
&+ \big[\big(\varphi_2^L\big)^{(2k+3)}(z)\big]^\dagger \big(\mathbb{I} - \big[\alpha_{2,2k+3}^R\big]^\dagger \alpha_{1,2k+3}^L\big) \\
&\times \big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^\dagger \alpha_{1,2k+2}^L\big),
\end{aligned}$$

$$\begin{aligned}
z\big(\varphi_1^R\big)^{(2k)} ={}& -\big(\varphi_1^R\big)^{(2k-1)}\alpha_{1,2k+1}^L - \big(\varphi_1^R\big)^{(2k)}\big[\alpha_{2,2k}^R\big]^\dagger \alpha_{1,2k+1}^L \\
&- \big(\varphi_1^R\big)^{(2k+1)}\alpha_{1,2k+2}^L \big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^\dagger \alpha_{1,2k+1}^L\big) \\
&+ \big(\varphi_1^R\big)^{(2k+2)}\big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^\dagger \alpha_{1,2k+2}^L\big)\big(\mathbb{I} - \big[\alpha_{2,2k+1}^R\big]^\dagger \alpha_{1,2k+1}^L\big), \\
z\big(\varphi_1^R\big)^{(2k+1)} ={}& \big(\varphi_1^R\big)^{(2k-1)} + \big(\varphi_1^R\big)^{(2k)}\big[\alpha_{2,2k}^R\big]^\dagger - \big(\varphi_1^R\big)^{(2k+1)}\alpha_{1,2k+2}^L \big[\alpha_{2,2k+1}^R\big]^\dagger \\
&+ \big(\varphi_1^R\big)^{(2k+2)}\big(\mathbb{I} - \big[\alpha_{2,2k+2}^R\big]^\dagger \alpha_{1,2k+2}^L\big)\big[\alpha_{2,2k+1}^R\big]^\dagger,
\end{aligned}$$

$$\begin{aligned}
z^{-1}\big(\varphi_1^R\big)^{(2k)} ={}& \big(\varphi_1^R\big)^{(2k-2)} + \big(\varphi_1^R\big)^{(2k-1)}\alpha_{1,2k-1}^L \\
&- \big(\varphi_1^R\big)^{(2k)}\alpha_{1,2k}^L \big[\alpha_{2,2k+1}^R\big]^\dagger + \big(\varphi_1^R\big)^{(2k+1)}\big(\mathbb{I} - \alpha_{1,2k+1}^L \big[\alpha_{2,2k+1}^R\big]^\dagger\big)\alpha_{1,2k}^L, \\
z^{-1}\big(\varphi_1^R\big)^{(2k-1)} ={}& -\big(\varphi_1^R\big)^{(2k-2)}\big[\alpha_{2,2k}^R\big]^\dagger - \big(\varphi_1^R\big)^{(2k-1)}\alpha_{1,2k-1}^L \big[\alpha_{2,2k}^R\big]^\dagger \\
&- \big(\varphi_1^R\big)^{(2k)}\big[\alpha_{1,2k+1}^L\big]^\dagger \big(\mathbb{I} - \alpha_{1,2k}^L \big[\alpha_{2,2k}^R\big]^\dagger\big) \\
&+ \big(\varphi_1^R\big)^{(2k+1)}\big(\mathbb{I} - \alpha_{1,2k+1}^L \big[\alpha_{2,2k+1}^R\big]^\dagger\big)\big(\mathbb{I} - \alpha_{1,2k}^L \big[\alpha_{2,2k}^R\big]^\dagger\big),
\end{aligned}$$

$$\begin{aligned}
\big[z\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger ={}& -\big[\alpha_{2,2k+1}^R\big]^\dagger \big(\mathbb{I} - \alpha_{1,2k}^L \big[\alpha_{2,2k}^R\big]^\dagger\big)\big[\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger \\
&- \big[\alpha_{2,2k+1}^R\big]^\dagger \alpha_{1,2k}^L \big[\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger - \big[\alpha_{2,2k+2}^R\big]^\dagger \big[\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger \\
&+ \big[\big(\varphi_2^R\big)^{(2k+2)}(z)\big]^\dagger,
\end{aligned}$$

$$
\begin{aligned}
\big[z\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger &= \big(\mathbb{I}-\alpha_{1,2k+1}^L\big[\alpha_{2,2k+1}^R\big]^\dagger\big)\big(\mathbb{I}-\alpha_{1,2k}^L\big[\alpha_{2,2k}^R\big]^\dagger\big)\big[\big(\varphi_2^R\big)^{(2k-1)}(z)\big]^\dagger \\
&\quad + \big(\mathbb{I}-\alpha_{1,2k+1}^L\big[\alpha_{2,2k+1}^R\big]^\dagger\big)\alpha_{1,2k}^L\big[\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger \\
&\quad - \alpha_{1,2k+1}^L\big[\alpha_{2,2k+1}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger + \alpha_{1,2k+1}^L\big[\big(\varphi_2^R\big)^{(2k+2)}(z)\big]^\dagger, \\
\big[z\big(\varphi_2^R\big)^{(0)}(z)\big]^\dagger &= -\big[\alpha_{2,1}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(0)}(z)\big]^\dagger - \big[\alpha_{2,2}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(1)}(z)\big]^\dagger + \big[\big(\varphi_2^R\big)^{(2)}(z)\big]^\dagger, \\
\big[z\big(\varphi_2^R\big)^{(1)}(z)\big]^\dagger &= \big(\mathbb{I}-\alpha_{1,1}^L\big[\alpha_{2,1}^R\big]^\dagger\big)\big[\big(\varphi_2^R\big)^{(0)}(z)\big]^\dagger \\
&\quad - \alpha_{1,1}^L\big[\alpha_{2,2}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(1)}(z)\big]^\dagger\alpha_{1,1}^L\big[\big(\varphi_2^R\big)^{(2)}(z)\big]^\dagger,
\end{aligned}
$$

$$
\begin{aligned}
\big[z^{-1}\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger &= \big(\mathbb{I}-\big[\alpha_{2,2k}^R\big]^\dagger\alpha_{1,2k}^L\big)\big(\mathbb{I}-\big[\alpha_{2,2k-1}^R\big]^\dagger\alpha_{1,2k-1}^L\big)\big[\big(\varphi_2^R\big)^{(2k-2)}(z)\big]^\dagger \\
&\quad + \big(\mathbb{I}-\big[\alpha_{2,2k}^R\big]^\dagger\alpha_{1,2k}^L\big)\big[\alpha_{2,2k-1}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(2k-1)}(z)\big]^\dagger \\
&\quad - \big[\alpha_{2,2k}^R\big]^\dagger\alpha_{1,2k+1}^L\big[\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger + \big[\alpha_{2,2k}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger, \\
\big[z^{-1}\big(\varphi_2^R\big)^{(2k-1)}(z)\big]^\dagger &= -\alpha_{1,2k}^L\big(\mathbb{I}-\big[\alpha_{2,2k-1}^R\big]^\dagger\alpha_{1,2k-1}^L\big)\big[\big(\varphi_2^R\big)^{(2k-2)}(z)\big]^\dagger \\
&\quad - \alpha_{1,2k}^L\big[\alpha_{2,2k-1}^R\big]^\dagger\big[\big(\varphi_2^R\big)^{(2k-1)}(z)\big]^\dagger - \alpha_{1,2k+1}^L\big[\big(\varphi_2^R\big)^{(2k)}(z)\big]^\dagger \\
&\quad + \big[\big(\varphi_2^R\big)^{(2k+1)}(z)\big]^\dagger, \\
\big[z^{-1}\big(\varphi_2^R\big)^{(0)}(z)\big]^\dagger &= -\alpha_{1,1}\big[\big(\varphi_2^R\big)^{(0)}(z)\big]^\dagger + \big[\big(\varphi_2^R\big)^{(1)}(z)\big]^\dagger.
\end{aligned}
$$

Appendix D. Projections in modules

For a ring $\mathbb{M}$ and left and right modules V and W over $\mathbb{M}$, respectively, bilinear forms are applications

$$G : V \times W \longrightarrow \mathbb{M}$$

such that

$$
\begin{aligned}
&G(m_1v_1 + m_2v_2, w) = m_1G(v_1, w) + m_2G(v_2, w), \quad \forall m_1, m_2 \in \mathbb{M},\ v, v_1, v_2 \in V, \\
&G(v, w_1m_1 + w_2m_2) = G(v, w_1)m_1 + G(v, w_2)m_2, \quad \forall m_1, m_2 \in \mathbb{M},\ w, w_1, w_2 \in W.
\end{aligned}
$$

In free modules, any such bilinear form can be represented by a unique $l \times r$ matrix denoted also by G, with coefficients in the ring $\mathbb{M}$, as follows

$$
\begin{aligned}
&G : V \times W \longrightarrow \mathbb{M}, \\
&G(v, w) := \begin{pmatrix} v_0 & \dots & v_{l-1} \end{pmatrix} G \begin{pmatrix} w_0 \\ \vdots \\ w_{l-1} \end{pmatrix}.
\end{aligned}
$$

Given free submodules $\tilde{V} \subset V$ and $\tilde{W} \subset W$ of the modules (not necessarily free) V, W and two bases $\{e_0, \dots, e_{\tilde{l}-1}\} \subset \tilde{V}$ and $\{f_0, \dots, f_{\tilde{r}-1}\} \subset \tilde{W}$ of $\tilde{V}$ and $\tilde{W}$, respectively, we denote $G_{i,j} = G(e_i, f_j)$. For the same rank, $\tilde{l} = \tilde{r}$, the matrix $\tilde{G} = (G_{i,j})$ can be assumed to be invertible, $\tilde{G} \in \mathrm{GL}(\tilde{l}, \mathbb{M}) \cong \mathrm{GL}(\tilde{l}m, \mathbb{C})$. In such case, we introduce the G-dual vectors to e_i, f_j defined as

$$e_i^* = \sum_{j=0}^{\tilde{l}-1} f_j (\tilde{G}^{-1})_{j,i}, \qquad f_j^* = \sum_{i=0}^{\tilde{r}-1} (\tilde{G}^{-1})_{j,i} e_i.$$

These vectors have some interesting properties:

(1) If we change basis $\hat{e}_j = \sum_{i=0}^{\tilde{l}-1} a_{j,i} e_i$ and $\hat{f}_j = \sum_{i=0}^{\tilde{l}-1} f_i b_{i,j}$, then

$$\hat{e}_j^* = \sum_{i=0}^{\tilde{l}-1} e_i^* (a^{-1})_{i,j}, \qquad \hat{f}_i^* = \sum_{i=0}^{\tilde{l}-1} (b^{-1})_{i,j} f_j^*,$$

where we have used the matrices $a = (a_{i,j})$ and $b = (b_{i,j})$, $a, b \in \mathrm{GL}(\tilde{l}, \mathbb{M})$.

(2) The sets of dual vectors $\{e_i^*\}_{i=0}^{\tilde{l}-1}$ and $\{f_i^*\}_{i=0}^{\tilde{l}-1}$ are bases with duals given by

$$(e_i^*)^* = e_i, \qquad (f_j^*)^* = f_j.$$

(3) It is easy to see that they satisfy the biorthogonal type identity

$$G(e_i, e_j^*) = G(f_i^*, f_j) = \delta_{i,j}, \quad \forall i, j = 0, \dots, \tilde{l} - 1.$$

Given the bilinear form G, we can construct the associated projections on these

$$p : V \to \tilde{V}, \qquad p(v) := \sum_{i=0}^{\tilde{l}-1} G(v, e_i^*) e_i,$$

$$q : W \to \tilde{W}, \qquad q(w) := \sum_{j=0}^{\tilde{l}-1} f_j G(f_j^*, w).$$

These constructions are relevant when considering the Christoffel–Darboux operators and formulae in the matrix context.

References

[1] M.J. Ablowitz, J.F. Ladik, Nonlinear differential–difference equations, J. Math. Phys. 16 (1975) 598–603.

[2] M.J. Ablowitz, J.F. Ladik, Nonlinear differential–difference equations and Fourier analysis, J. Math. Phys. 17 (1976) 1011–1018.

[3] M. Adler, P. van Moerbecke, Integrals over classical groups, random permutations, Toda and Toeplitz lattices, Comm. Pure Appl. Math. 54 (2001) 153–205.

[4] M. Adler, P. van Moerbeke, Group factorization, moment matrices and Toda lattices, Int. Math. Res. Not. IMRN 12 (1997) 556–572.

[5] M. Adler, P. van Moerbeke, Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems, Comm. Math. Phys. 207 (1999) 589–620.

[6] M. Adler, P. van Moerbeke, The spectrum of coupled random matrices, Ann. of Math. 149 (1999) 921–976.

[7] M. Adler, P. van Moerbeke, Vertex operator solutions to the discrete KP hierarchy, Comm. Math. Phys. 203 (1999) 185–210.

[8] M. Adler, P. van Moerbeke, Darboux transforms on band matrices, weights and associated polynomials, Int. Math. Res. Not. IMRN 18 (2001) 935–984.

[9] M. Adler, P. van Moerbeke, P. Vanhaecke, Moment matrices and multi-component KP, with applications to random matrix theory, Comm. Math. Phys. 286 (2009) 1–38.

[10] M. Alfaro, Una expresión de los polinomios ortogonales sobre la circunferencia unidad, in: Actas III Jornadas Matemáticas Hispano-Lusas, vol. 2, Sevilla, 1974, 1982, pp. 1–8.

[11] C. Álvarez-Fernández, U. Fidalgo, M. Mañas, The multicomponent 2D Toda hierarchy: generalized matrix orthogonal polynomials, multiple orthogonal polynomials and Riemann–Hilbert problems, Inverse Problems 26 (2010) 055009, 17 pp.

[12] C. Álvarez-Fernández, U. Fidalgo, M. Mañas, Multiple orthogonal polynomials of mixed type: Gauss–Borel factorization and the multi-component 2D Toda hierarchy, Adv. Math. 227 (2011) 1451–1525.

[13] C. Álvarez-Fernández, M. Mañas, Orthogonal Laurent polynomials on the unit circle, extended CMV ordering and 2D Toda type integrable hierarchies, Adv. Math. 240 (2013) 132–193.

[14] C. Álvarez-Fernández, M. Mañas, On the Christoffel–Darboux formula for generalized matrix orthogonal polynomials of multigraded Hankel type, arXiv:1311.0563 [math.CA], 2013.

[15] R. Álvarez-Nodarse, J. Arvesú, F. Marcellán, Modifications of quasi-definite linear functionals via addition of delta and derivatives of delta Dirac functions, Indag. Math. 15 (2004) 1–20.

[16] M. Ambroladze, On exceptional sets of asymptotics relations for general orthogonal polynomials, J. Approx. Theory 82 (1995) 257–273.

[17] A.I. Aptekarev, E.M. Nikishin, The scattering problem for a discrete Sturm–Liouville operator, Math. USSR Sb. 49 (1984) 325–355.

[18] G. Araznibarreta, M. Mañas, Another Christoffel–Darboux formula for multiple orthogonal polynomials of mixed type, arXiv:1310.7240 [math.CA], 2013.

[19] D. Barrios, G. López, Ratio asymptotics for orthogonal polynomials on arcs of the unit circle, Constr. Approx. 15 (1999) 1–31.

[20] Yu.M. Berezanskii, Expansions in Eigenfunctions of Self-Adjoint Operators, Transl. Math. Monogr., vol. 17, American Mathematical Society, 1968.

[21] M.J. Bergvelt, A.P.E. ten Kroode, Partitions, vertex operators constructions and multi-component KP equations, Pacific J. Math. 171 (1995) 23–88.

[22] E. Berriochoa, A. Cachafeiro, J. Garcia-Amor, Connection between orthogonal polynomials on the unit circle and bounded interval, J. Comput. Appl. Math. 177 (2005) 205–223.

[23] M. Bertola, M. Gekhtman, Biorthogonal Laurent polynomials, Toeplitz determinants, minimal Toda orbits and isomonodromic tau functions, Constr. Approx. 26 (2007) 383–430.

[24] J. Borrego, M. Castro, A.J. Durán, Orthogonal matrix polynomials satisfying differential equations with recurrence coefficients having non-scalar limits, Integral Transforms Spec. Funct. 23 (2012) 685–700.

[25] A. Bultheel, P. González-Vera, E. Hendriksen, O. Njåstad, Orthogonal Rational Functions, Cambridge Monogr. Appl. Comput. Math., vol. 5, Cambridge University Press, Cambridge, 1999.

[26] A. Cachafeiro, F. Marcellán, C. Pérez, Lebesgue perturbation of a quasi-definite Hermitian functional. The positive definite case, Linear Algebra Appl. 369 (2003) 235–250.

[27] M. Cafasso, Matrix Biorthogonal Polynomials on the unit circle and the non-Abelian Ablowitz–Ladik hierarchy, J. Phys. A: Math. Theor. 42 (2009) 365211.

[28] M.J. Cantero, L. Moral, F. Marcellan, L. Velázquez, CMV matrices and spectral transformations of measures, talk given in the Matrices and Orthogonal Polynomials Minisymposium in ILAS 2013, The 18th Conference of the International Linear Algebra Society, Providence, RI, June 3–7, 2013.

[29] M.J. Cantero, L. Moral, L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, Linear Algebra Appl. 362 (2003) 29–56.

[30] M.J. Cantero, L. Moral, L. Velázquez, Differential properties of matrix orthogonal polynomials, J. Concr. Appl. Math. v3 (2008) 313–334.
[31] G.A. Cassatella-Contra, M. Mañas, Riemann–Hilbert problems, matrix orthogonal polynomials and discrete matrix equations with singularity confinement, Stud. Appl. Math. 128 (2011) 252–274.
[32] G.A. Cassatella-Contra, M. Mañas, P. Tempesta, Singularity confinement for matrix discrete Painlevé equations, arXiv:1311.0557 [math.CA], 2013.
[33] L. Cochran, S.C. Cooper, Orthogonal Laurent polynomials on the real line, in: S.C. Cooper, W.J. Thron (Eds.), Continued Fractions and Orthogonal Functions, in: Lect. Notes Pure Appl. Math. Ser., vol. 154, Marcel Dekker, New York, 1994, p. 47100.
[34] R.D. Costin, Matrix valued polynomials generated by the scalar-type Rodrigues' formulas, J. Approx. Theory 161 (2009) 693–705.
[35] R. Cruz-Barroso, L. Daruis, P. González-Vera, O. Njastadb, Sequences of orthogonal Laurent polynomials, biorthogonality and quadrature formulas on the unit circle, J. Comput. Appl. Math. 200 (2007) 424–440.
[36] R. Cruz-Barroso, S. Delvaux, Orthogonal Laurent polynomials on the unit circle and snake-shaped matrix factorizations, J. Approx. Theory 161 (2009) 65–87.
[37] R. Cruz-Barroso, P. González-Vera, A Christoffel–Darboux formula and a Favard's theorem for Laurent orthogonal polynomials on the unit circle, J. Comput. Appl. Math. 179 (2005) 157–173.
[38] E. Daems, A.B.J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions, J. Approx. Theory 146 (2007) 91–114.
[39] D. Damanik, A. Pushnitski, B. Simon, The Analytic Theory of Matrix Orthogonal Polynomials, Surv. Approx. Theory 4 (2008) 1–85.
[40] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Operator approach to the Kadomtsev–Petviashvili equation. Transformation groups for soliton equations. III, J. Phys. Soc. Jpn. 50 (1981) 3806–3812.
[41] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Transformation groups for soliton equations. Euclidean Lie algebras and reduction of the KP hierarchy, Publ. Res. Inst. Math. Sci. 18 (1982) 1077–1110.
[42] E. Date, M. Jimbo, M. Kashiwara, T. Miwa, Transformation groups for soliton equations, in: M. Jimbo, T. Miwa (Eds.), Nonlinear Integrable Systems-Classical Theory and Quantum Theory, World Scientific, Singapore, 1983, pp. 39–120.
[43] C. Díaz-Mendoza, P. González-Vera, M. Jiménez-Paiz, Strong Stieltjes distributions and orthogonal Laurent polynomials with applications to quadratures and Padé approximation, Math. Comp. 74 (2005) 1843–1870.
[44] A.J. Durán, Matrix inner product having a matrix symmetric second order differential operator, Rocky Mountain J. Math. 27 (1997) 585–600.
[45] A.J. Durán, M.D. de la Iglesia, Second order differential operators having several families of orthogonal matrix polynomials as eigenfunctions, Int. Math. Res. Not. IMRN 2008 (2008), Article ID rnn084, 24 pp.
[46] A.J. Durán, F.J. Grünbaum, Orthogonal matrix polynomials satisfying second order differential equations, Int. Math. Res. Not. IMRN 10 (2004) 461–484.
[47] A.J. Durán, F.J. Grünbaum, Orthogonal matrix polynomials, scalar-type Rodrigues' formulas and Pearson equations, J. Approx. Theory 134 (2005) 267–280.
[48] A.J. Durán, F.J. Grünbaum, Structural formulas for orthogonal matrix polynomials satisfying second order differential equations, I, Constr. Approx. 22 (2005) 255–271.
[49] L. Faybusovich, M. Gekhtman, On Schur flows, J. Phys. A: Math. Gen. 32 (1999) 4671–4680.
[50] L. Faybusovich, M. Gekhtman, Elementary Toda orbits and integrable lattices, J. Math. Phys. 41 (2000) 2905–2921.
[51] L. Faybusovich, M. Gekhtman, Inverse moment problem for elementary co-adjoint orbits, Inverse Problems 17 (2001) 1295–1306.
[52] R. Felipe, F. Ongay, Algebraic aspects of the discrete KP hierarchy, Linear Algebra Appl. 338 (2001) 1–17.
[53] G. Freud, Orthogonal Polynomials, Akadémiai Kiadó, Budapest, 1971, Pergamon Press, Oxford, 1985.
[54] P. García, F. Marcellán, On zeros of regular orthogonal polynomials on the unit circle, Ann. Polon. Math. 58 (1993) 287–298.
[55] W. Gautschi, Orthogonal Polynomials: Computation and Approximation, Oxford University Press, New York, 2004.
[56] J.S. Geronimo, Scattering theory and matrix orthogonal polynomials on the real line, Circuits Systems Signal Process. 1 (1982) 471–495.

[57] Ya.L. Geronimus, Polynomials orthogonal on a circle and their applications, series and approximations, in: Amer. Math. Soc. Transl. Ser. 1, vol. 3, Amer. Math. Soc., Providence, RI, 1962, pp. 1–78.
[58] E. Godoy, F. Marcellán, Orthogonal polynomials on the unit circle: distribution of zeros, J. Comput. Appl. Math. 37 (1991) 195–208.
[59] L. Golinskii, Schur flows and orthogonal polynomials on the unit circle, Sb. Math. 197 (2006) 1145.
[60] L. Golinskii, A. Zlatos, Coefficients of orthogonal polynomials on the unit circle and higher order Szegő theorems, Constr. Approx. 26 (2007) 361–382.
[61] G.H. Golub, C.F. Van Loan, Matrix Computations, third ed., The John Hopkins University Press, Baltimore, CA, 1996.
[62] F. Grünbaum, M.D. de la Iglesia, A. Martínez-Finkelshtein, Properties of matrix orthogonal polynomials via their Riemann–Hilbert characterization, SIGMA 7 (2011) 098, 31 pp.
[63] W.B. Jones, O. Njåstad, Applications of Szegő polynomials to digital signal processing, Rocky Mountain J. Math. 21 (1991) 387–436.
[64] W.B. Jones, O. Njåstad, W.J. Thron, Two-point Padé expansions for a family of analytic functions, J. Comput. Appl. Math. 9 (1983) 105.123.
[65] W.B. Jones, O. Njåstad, W.J. Thron, H. Waadeland, Szegő polynomials applied to frequency analysis, J. Comput. Appl. Math. 46 (1993) 217–228.
[66] W.B. Jones, W.J. Thron, Orthogonal Laurent polynomials and Gaussian quadrature, in: K.E. Gustafson, W.P. Reinhardt (Eds.), Quantum Mechanics in Mathematics Chemistry and Physics, Plenum, New York, 1981, pp. 449–455.
[67] W.B. Jones, W.J. Thron, H. Waadeland, A strong Stieltjes moment problem, Trans. Amer. Math. Soc. 261 (1980) 503–528.
[68] R. Killip, I. Nenciu, CMV: the unitary analogue of Jacobi matrices, Comm. Pure Appl. Math. 60 (2006) 1148–1188.
[69] M.G. Krein, Infinite J-matrices and a matrix moment problem, Dokl. Akad. Nauk SSSR 69 (1949) 125–128.
[70] M.G. Krein, Fundamental aspects of the representation theory of hermitian operators with deficiency index (m, m), in: Amer. Math. Soc. Transl. Ser. 2, vol. 97, Amer. Math. Soc., Providence, Rhode Island, 1971, pp. 75–143.
[71] M. Mañas, L. Martínez Alonso, C. Álvarez-Fernández, The multicomponent 2D Toda hierarchy: discrete flows and string equations, Inverse Problems 25 (2009) 065007, 31 pp.
[72] M. Mañas, L. Martínez Alonso, E. Medina, Dressing methods for geometric nets: I. Conjugate nets, J. Phys. A: Math. Gen. 33 (2000) 2871.
[73] M. Mañas, L. Martínez Alonso, E. Medina, Dressing methods for geometric nets: II. Orthogonal and Egorov nets, J. Phys. A: Math. Gen. 33 (2000) 7181.
[74] H.N. Mhaskar, E.B. Saff, On the distribution of zeros of polynomials orthogonal on the unit circle, J. Approx. Theory 63 (1990) 30–38.
[75] L. Miranian, Matrix valued orthogonal polynomials on the real line: some extensions of the classical theory, J. Phys. A: Math. Gen. 38 (2005) 5731–5749.
[76] L. Miranian, Matrix valued orthogonal polynomials on the unit circle: some extensions of the classical theory, Canad. Math. Bull. 52 (2009) 95–104.
[77] A. Mukaihira, Y. Nakamura, Schur flow for orthogonal polynomials on the unit circle and its integrable discretization, J. Comput. Appl. Math. 139 (2002) 75–94.
[78] M. Mulase, Complete integrability of the Kadomtsev–Petviashvili equation, Adv. Math. 54 (1984) 57–66.
[79] I. Nenciu, Lax pairs for the Ablowitz–Ladik system via orthogonal polynomials on the unit circle, Int. Math. Res. Not. IMRN 11 (2005) 647–686.
[80] P. Nevai, V. Totik, Orthogonal polynomials and their zeros, Acta Sci. Math. (Szeged) 53 (1–2) (1989) 99–104.
[81] O. Njåstad, W.J. Thron, The theory of sequences of orthogonal L-polynomials, in: H. Waadeland, H. Wallin (Eds.), Padé Approximants and Continued Fractions, Det Kongelige Norske Videnskabers Selskabs Skrifter, 1983, pp. 54–91.
[82] K. Pan, Asymptotics for Szegő polynomials associated with Wiener–Levinson filters, J. Comput. Appl. Math. 46 (1993) 387–394.
[83] K. Pan, E.B. Saff, Asymptotics for zeros of Szegő polynomials associated with trigonometric polynomials signals, J. Approx. Theory 71 (1992) 239–251.
[84] M. Sato, Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds, Res. Inst. Math. Sci. Kokyuroku 439 (1981) 30–46.

[85] B. Simon, Orthogonal Polynomials on the Unit Circle, part 1: Classical Theory, AMS Colloq. Ser., American Mathematical Society, Providence, RI, 2005.
[86] B. Simon, Orthogonal Polynomials on the Unit Circle, part 2: Spectral Theory, AMS Colloquium Series, American Mathematical Society, Providence, RI, 2005.
[87] B. Simon, CMV matrices: five years after, J. Comput. Appl. Math. 208 (2007) 120–154.
[88] B. Simon, Zeros of OPUC and long time asymptotics of Schur and related flows, Inverse Probl. Imaging 1 (2007) 189–215.
[89] B. Simon, Szegő's Theorem and Its Descendants, Princeton University Press, Princeton, New Jersey, 2011.
[90] G. Szegő, Orthogonal Polynomials, Amer. Math. Soc. Colloq. Publ., vol. XXIII, American Mathematical Society, Providence, Rhode Island, 1975.
[91] W.J. Thron, L-polynomials orthogonal on the unit circle, in: A. Cuyt (Ed.), Nonlinear Methods and Rational Approximation, Reidel Publishing Company, Dordrecht, 1988, pp. 271–278.
[92] K. Ueno, K. Takasaki, Toda lattice hierarchy, in: Group Representations and Systems of Differential Equations, in: Adv. Stud. Pure Math., vol. 4, 1984, pp. 1–95.
[93] D.S. Watkins, Some perspectives on the eigenvalue problem, SIAM Rev. 35 (1993) 430–471.

5 Generalized Sobolev Orthogonal Polynomials, Matrix Moment Problems and Integrable Systems

GENERALIZED SOBOLEV ORTHOGONAL POLYNOMIALS, MATRIX MOMENT PROBLEMS AND INTEGRABLE SYSTEMS

GERARDO ARIZNABARRETA, MANUEL MAÑAS, AND PIERGIULIO TEMPESTA

Abstract. We introduce a large class of Sobolev bi-orthogonal polynomial sequences arising from a LU-factorizable moment matrix and associated with a suitable measure matrix that characterizes the Sobolev bilinear form. A theory of deformations of Sobolev bilinear forms is also proposed. We consider both polynomial deformations and a class of transformations related to the action of linear operators on the entries of a given bilinear form. Transformation formulae among new and old polynomial sequences are determined.

Finally, integrable hierarchies of evolution equations arising from the factorization of a time deformation of the moment matrix are presented.

MSC2010: 33C45, 37L60, 42C05

Contents

Date: December 15, 2016.

1. Introduction

1.1. **Historical background and motivation.** In the last decades, the study of Sobolev orthogonal polynomials has become a field of increasing interest both in Applied Mathematics and Mathematical Physics. The purpose of this article is to extend the notion of Sobolev orthogonality by introducing a theoretical framework allowing to define a new, large class of Sobolev bi-orthogonal polynomial sequences (SBPS).

In order to situate our contribution in the context of the existing literature, we start by mentioning some of the most relevant results of the theory established till now. We focus here only on some aspects of special interest for our research. For a nice review of modern results, historical background and an updated bibliography, the reader is referred to [24], [25].

Sobolev orthogonal polynomials were introduced in 1962 by Althammer [1]. He proposed the idea of defining a class of polynomials orthogonal with respect to a *deformation* of the Legendre inner product, of the form

$$\langle f, g\rangle_A = \int_{-1}^{1} f(x)g(x)dx + \lambda \int_{-1}^{1} f'(x)g'(x)dx \ . \tag{1}$$

The polynomials arising from this inner product are called nowadays the Sobolev-Legendre polynomials.

Perhaps the most relevant of the early contributions to the theory came in the 70's with the works [29], [30]. Indeed, Schäfke and Wolf proposed the following family of inner products

$$\langle f, g\rangle_{SW} = \sum_{j,k=0}^{\infty} \int_a^b f^{(j)}(x)g^{(k)}(x)v_{j,k}(x)w(x)dx \ , \tag{2}$$

where the weight w and the associated integration interval is intended to be one of the three classical cases of Hermite, Laguerre and Jacobi; also, $v_{j,k}(x)$ are suitable polynomials, symmetric in j, k.

Starting from this polynomial deformation of classical measures, and specializing conveniently the functions $v_{j,k}$, Schäfke and Wolf were able to define eight families of new Sobolev orthogonal polynomials, and extended all previously known results on Sobolev orthogonal polynomials.

Since the last decade of the previous century there was a resurgence of interest in the field of Sobolev orthogonality, starting with the seminal paper [13]. In this work, the notion of *coherent pairs*, a fundamental idea which has triggered many new developments, was introduced. Let $\{d\mu_1, d\mu_2\}$ be a pair of Borel measures on the real line with finite moments. To this pair we associate the inner product $\langle f, g\rangle_{(\mu_1,\mu_2)} = \int_a^b f(x)g(x)d\mu_1 + \lambda \int_a^b f'(x)g'(x)d\mu_2$, with $a, b \in \mathbb{R}$. Essentially, the pair of measures $\{d\mu_1, d\mu_2\}$ is said to be a coherent pair whenever the sequence of polynomials associated with $d\mu_2$ can be related in a specific way with the first derivatives of the polynomials of the sequence associated with $d\mu_1$. In [22] a classification of coherent pairs was given when one of the two involved measures is a classical one (Hermite, Laguerre, Jacobi or Bessel). In [26] it was proven that in order for $\{d\mu_1, d\mu_2\}$ to form a coherent pair, at least one of the two measures has to be classical. This result shows that the classification given in [22] is actually a complete one.

Besides, a huge amount of results concerning many analytic and algebraic aspects of the theory has been obtained in the last twenty years, including the relation with differential operators [15], [9], the asymptotic behaviour and the study of zeros of Sobolev polynomials [18], etc.

1.2. **Main results.** In this paper, we generalize significantly the construction of Schäfke and Wolf by introducing a large class of not necessarily symmetric Sobolev bilinear forms $(*, *)_{\mathscr{W}}$. These bilinear forms are defined by means of a matrix of measures $\mathscr{W}$, representing one of the crucial mathematical structures of the present paper. To each measure matrix $\mathscr{W}$, or equivalently to the corresponding bilinear form, we can naturally associate a moment matrix $G_{\mathscr{W}}$. In our analysis, we shall focus on the class of moment matrices that admit an *LU-factorization*. Indeed, for this class one can construct Sobolev bi-orthogonal polynomial sequences (SBPS). We shall prove that many algebraic techniques related to the LU-factorization, that proved to be very useful in order to obtain algebraic properties of the standard orthogonal polynomial sequences (OPS) can be extended naturally to our Sobolev setting.

A crucial notion proposed in this paper is that of *additive perturbations* of a measure matrix $\mathscr{W}$ in the Sobolev context. Precisely, we shall study under which conditions, by performing an additive matrix perturbation of $\mathscr{W}$, one can still produce families of SBPS. This approach turns out to be particularly fruitful. Indeed, one can describe on the same footing, and generalize widely, important constructions as the coherent pairs and the standard approach of discrete Sobolev bilinear forms. Concerning the first aspect, we wish to point out that not only a standard coherent pair can be studied from the perspective of perturbation theory, but it also can be generalized, in terms of the new notion of *$m \times m$ block coherent pair*. The SBPS arising from both standard and block coherent pairs are studied.

When the entries of the measure matrix $\mathscr{W}$ are allowed to depend on δ distributions, we can encompass in our approach the well-known case of discrete Sobolev orthogonality. Once we split a Sobolev bilinear function into a continuous part, involving those entries of $\mathscr{W}$ having a continuous support, and a discrete one, involving those having a discrete support (δ distributions) [1], we can interpret the discrete part as an *additive discrete perturbation* of its continuous part. This leads to an interesting characterization of the SBPS associated to the original measure matrix in terms of quasi-determinantal formulae, involving only the continuous part of the bilinear function.

A related aspect is the possibility of classifying measure matrices in terms of *equivalence classes*: To each class it belongs a set of measure matrices giving rise to the same moment matrix, and therefore to the same SBPS. Indeed, the correspondence between measure matrices and moment matrices is not one to one. Therefore, different Sobolev bilinear forms may lead to the same SBPS. An interesting case arises when inside the same equivalence class possibly Sobolev and non Sobolev-type measure matrices are present. All this is not surprising, taking into account that the integration by parts procedure (at least in a distributional sense) comes into play, allowing to define elementary operations leaving a measure matrix into the same class.

Due to the relevance of measure matrices in our approach, a natural problem is to develop a deformation theory for these matrices which allows us to relate the corresponding deformed and non deformed SBPS.

Special attention will be devoted to certain classes of transformations well known in the literature on orthogonal polynomials: Christoffel's and Geronimo's transformations. The first ones were introduced in 1858 by Christoffel [7], and amount to a polynomial deformation of a given classical measure. Precisely, the standard Christoffel formulae establish connections among families of orthogonal polynomials, allowing to express a polynomial of a family just in terms of a constant number of polynomials of the other family. We generalize this approach by introducing Christoffel-Sobolev transformations. These involve a matrix polynomial deformation of the Sobolev measure matrix $\mathscr{W}$, which can be implemented by means of a right or left action of the deformation on the matrix $\mathscr{W}$. Once suitable resolvents and their adjoints are defined, then it is possible to connect deformed and non-deformed Sobolev polynomial sequences (and related Christoffel-Darboux kernels). In addition, quasi-determinantal expressions for the deformed polynomial sequences in terms of the original ones are obtained.

The second class of deformations we shall generalize is that of Geronimus, which was introduced in [11] (see also [12]). We propose, in our context, the notion of Geronimus-Sobolev transformation of a measure matrix. This very general transformation amounts to a right or left multiplication of the initial measure matrix $\mathscr{W}$ by the inverse of a matrix polynomial, extended by the addition of a discrete deformation. Once again, one can obtain explicit formulae connecting deformed and non-deformed polynomials (and Christoffel-Darboux kernels) that are expressed

[1] Some authors call type I Sobolev products those involving continuous supports only and type II and III those involving a continuous support while the rest are finite subsets

in compact quasi-determinantal expressions.

The previous cases of polynomial and inverse polynomial-type deformations of the measure matrix are of special interest, but do not exhaust the range of possible transformations we can perform over $\mathscr{W}$. Another novel aspect of the present work is that, indeed, we broaden the family of possible deformations by admitting much more general deformations. They are expressed in terms of *linear differential operators* with polynomial coefficients, this is, operators of the form $\boldsymbol{L} = \sum_k p_k(x)\frac{\mathrm{d}^k}{\mathrm{d}x^k}$, acting on the entries of the original bilinear form. Due to its generality, the theory of these operator deformations appears to be extremely rich (see also [2]). In this paper, we focused on several aspects which look of particular interest. Given a couple of linear differential operators of the form given above, it is possible to define a new class of Sobolev bilinear forms, which under certain technical conditions still possesses an associated moment matrix $G_{\mathscr{W}}$ which is LU-factorizable and consequently, give a SBPS.

We mention that an article which in some sense can be related to section 6 of the present one is Ref. [4]. In that work, the authors consider polynomial perturbations of a generic sesquilinear form. The methods used there are specially suited to polynomial perturbations of a matrix bivariate functional, and therefore include matrix Sobolev bilinear forms. The present paper focuses explicitly on the Sobolev scenario, from a different point of view. The fraction of the results of [4] concerning polynomial deformations of sesquilinear forms, in our opinion cannot be translated into our context in a simple or useful way. For that reason, we have introduced Sections 6.3, 6.4 and 6.5, where polynomial perturbations are treated expressly for the Sobolev (scalar) setting. It must be underlined that the deformations of the bilinear forms that the present paper considers (Section 7) are certainly more general since linear differential operator transformations are allowed instead of just polynomial ones.

The paper is organized as follows. In Section 2, we introduce the main notions of our analysis: Measure matrices, Sobolev generalized bilinear forms, moment matrices and the LU-factorization is studied. In Section 3, we construct the family of Sobolev bi-orthogonal polynomial sequences arising from LU-factorizable moment matrices together with the introduction of their associated second kind functions. Chistoffel-Darboux and Cauchy kernels associated with these sequences are also defined. In Section 4, we propose a theory of additive perturbations of measure matrices, which allows us to treat on the same footing coherent pairs (and a generalization of these) and discrete bilinear forms of Sobolev type. The crucial idea of equivalence classes of measure matrices is introduced and developed in Section 5. This idea proves to be of special interest when classical measures are involved in the bilinear form; some attention is devoted to these measures in order to generalize some known results. A polynomial deformation theory of the measure matrices is proposed in Section 6, which includes the important case of linear spectral or Darboux-Sobolev transformations. Section 7 is devoted to an extension of our theory of deformations of measure matrices to the case of linear differential operators. The study of the relation of the present approach with integrable hierarchies of Toda type is presented in the final Appendix.

2. Algebraic preliminaries

2.1. **A generalized Sobolev bilinear form.** We shall first introduce the main definitions necessary for our approach.

Definition 1. *A measure matrix of order $\mathscr{N}$, with $\mathscr{N} \in \mathbb{N}$ is a matrix $\mathscr{W}$ whose entries $\{\mathrm{d}\mu_{i,j}(x)\}_{i,j}$ are Borel measures and $\mathrm{d}\mu_{i,j} = 0\ \forall i, j > \mathscr{N}$:*

$$\mathscr{W}(x) := \begin{pmatrix} \mathrm{d}\mu_{0,0} & \mathrm{d}\mu_{0,1} & \dots & \mathrm{d}\mu_{0,\mathscr{N}} & 0 & \dots \\ \mathrm{d}\mu_{1,0} & \mathrm{d}\mu_{1,1} & \dots & \mathrm{d}\mu_{1,\mathscr{N}} & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ \mathrm{d}\mu_{\mathscr{N},0} & \mathrm{d}\mu_{\mathscr{N},1} & \dots & \mathrm{d}\mu_{\mathscr{N},\mathscr{N}} & 0 & \dots \\ 0 & 0 & & 0 & 0 & \dots \\ \vdots & \vdots & & \vdots & & \ddots \end{pmatrix} \qquad \mathrm{d}\mu_{i,j} : \Omega_{i,j} \subseteq \mathbb{R} \longrightarrow \mathbb{R}$$

Definition 2. *The bilinear form $(*, *; \mathscr{W}) : \mathbb{R}[x] \times \mathbb{R}[x] \longrightarrow \mathbb{R}$ associated with $\mathscr{W}$ is defined to be*

$$(3) \qquad (x^i, x^j; \mathscr{W}) := \sum_{n,r=0}^{\mathscr{N}} \left\langle \frac{\mathrm{d}^n x^i}{\mathrm{d}x^n}, \frac{\mathrm{d}^r x^j}{\mathrm{d}x^r} \right\rangle_{n,r} \qquad \text{where} \qquad \left\langle \frac{\mathrm{d}^n x^i}{\mathrm{d}x^n}, \frac{\mathrm{d}^r x^j}{\mathrm{d}x^r} \right\rangle_{n,r} := \int_{\Omega_{n,r}} \frac{\mathrm{d}^n x^i}{\mathrm{d}x^n} \frac{\mathrm{d}^r x^j}{\mathrm{d}x^r} \mathrm{d}\mu_{n,r}(x)$$

where we assume the condition $|(x^i, x^j; \mathscr{W})| < \infty\ \forall i, j \in \mathbb{N}$.

It is important to notice that the case $\mathscr{N} \longrightarrow \infty$ is also allowed since for given $i, j \in \mathbb{N}$ the bilinear form $(x^i, x^j; \mathscr{W})$ will always involve a finite number of terms only.

We wish to extend the domain of the bilinear form (3) to a more general function space containing $\mathbb{R}[x]$ as a subspace.

Definition 3. *Let $\Omega := \bigcup_{i,j=0}^{\mathscr{N}} \Omega_{i,j}$. The function space $\mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega)$ is defined as*

$$\mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega) := \left\{ f(x) \in C^{\mathscr{N}}(\Omega) \text{ such that } |(f, f; \mathscr{W})| := \left| \sum_{n,r=0}^{\mathscr{N}} \left\langle \frac{\mathrm{d}^n f}{\mathrm{d}x}, \frac{\mathrm{d}^r f}{\mathrm{d}x} \right\rangle_{n,r} \right| < \infty \right\},$$

where $C^k(\Omega)$ denotes the space of functions possessing k continuous derivatives in Ω.

We wish to endow the space $\mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega)$ with a structure of normed vector space, with norm given by $||f||^2 := (f, f; W)$. Therefore, jointly with the existence of finite moments, we need also to require positive definiteness: $\forall f \neq 0, (f, f; W) > 0$. Hereafter we shall tacitly assume that this condition is satisfied.

Observe that, since every continuous bilinear function is bounded, we have that whenever $f(x), g(x) \in \mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega)$ the pairing $(f, g; \mathscr{W})$ satisfies $|(f, g; \mathscr{W})| \leq C||f||||g||$, and therefore is finite. Consequently, we can introduce the notion of Sobolev bilinear function.

Definition 4. *For every $f(x), g(x) \in \mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega)$ we shall call the non degenerate positive definite bilinear function $(*, *; \mathscr{W}) : \mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega) \times \mathscr{A}_{\mathscr{W}}^{\mathscr{N}}(\Omega) \longrightarrow \mathbb{R}$ defined by*

$$(4) \qquad (f, h; \mathscr{W}) := \sum_{n,r=0}^{\mathscr{N}} \langle f^{(n)}, h^{(r)} \rangle_{n,r} \qquad \text{with} \qquad f^{(n)} := \frac{\mathrm{d}^n f(x)}{\mathrm{d}x^n}$$

the Sobolev bilinear function associated with the measure matrix $\mathscr{W}$.

Several comments are in order.

- Definition 4 includes as a particular case the standard inner product, with no derivatives involved, which corresponds to the choice $\mathscr{N} = 0$, namely $\mathrm{d}\mu_{i,j} = 0\ \forall i, j > 0$.
- Choosing a non symmetric $\mathscr{W}$ leads us to extend naturally the concept of orthogonality to that of bi-orthogonality. Indeed, one could have $(f, h; \mathscr{W}) = 0$ while $(h, f; \mathscr{W}) \neq 0$. This situation also occurs in the study of standard matrix orthogonality with respect to a non symmetric matrix measure (see for example [5]) or when dealing with scalar bivariate linear functionals (see for example [4]) .

- If $\mathscr{W}=\mathscr{W}^\top$ we obtain a positive definite symmetric bilinear form $(f,h;\mathscr{W})=(h,f;\mathscr{W})$ which allows us to define a standard inner product. Observe that the literature on the subject specially focuses on diagonal $\mathscr{W}$, for which obviously $\mathscr{W}=\mathscr{W}^\top$.

Remark 1. *Unlike the point of view adopted in* [4]*, based on the bivariate linear functional setting, in this paper we have preferred to work with an integral representation of our bilinear form. This representation exists as a direct consequence of the Riesz-Markov-Kakutani theorem* [14]*. The reason for this choice is the fact that we wish to develop a theory explicitly related with measure matrices.*

2.2. **The moment matrix.** Our approach to SBPS requires the definition of a suitable moment matrix. Notice that the Hankel-type form of the moment matrix, usual in the non Sobolev context, is expected to be lost or generalized; according to [28], the generalized form can be called Hankel–Sobolev matrices. The associated moment problem will involve more than just one sequence of integers (for a study of a diagonal $\mathscr{W}$ see [6],[23]); of course, a propaedeutic problem will be to establish under which conditions a matrix can play the role of a suitable Sobolev moment matrix. Instead, we prefer to proceed in a somewhat different way: we construct a moment matrix suitable for the Sobolev bilinear function (4). We start by settling some notation.

Given two non negative integers m,n we will denote by $(m)^n$ and $(m)_n$ the rising and lower factorial polynomials respectively, i.e.

$$(m)^n := m(m+1)(m+2)\dots(m+(n-1))$$

$$(m)_n := \begin{cases} m(m-1)(m-2)\dots(m-(n-1)) & n<m\\ 0 & n\geq m\end{cases}$$

$$(m)^0=(m)_0:=1 \qquad\qquad (m)^1=(m)_1:=m$$

Definition 5. *We introduce the vectors*

$$\chi(x):=\begin{pmatrix}1\\x\\x^2\\x^3\\\vdots\\x^k\\\vdots\end{pmatrix}\quad \chi'(x):=\begin{pmatrix}0\\1\\2x\\3x^2\\\vdots\\kx^{k-1}\\\vdots\end{pmatrix}\quad \chi''(x):=\begin{pmatrix}0\\0\\2\\(3)(2)x\\\vdots\\k(k-1)x^{k-2}\\\vdots\end{pmatrix}\quad\dots\quad \chi^{(n)}(x):=\begin{pmatrix}0\\0\\\vdots\\(n)_n\\\vdots\\(k)_nx^{k-n}\\\vdots\end{pmatrix}\quad\dots$$

and the lower semi infinite matrix

$$\boldsymbol{\chi}(x):=\begin{pmatrix}\chi(x) & \chi'(x) & \chi''(x) & \dots & \chi^{(k)}(x) & \dots\end{pmatrix}$$

We also define the auxiliary vector

$$\chi^*(x):=\frac{1}{x}\chi\left(\frac{1}{x}\right).$$

The previous definition allows to deal with polynomials in a simple way. Let $p(x)\in\mathbb{R}[x]$ be a polynomial of degree k, i.e. $p(x)=\sum_l p_lx^l$ with $p_l=0\ \forall l>k$. Let us denote by $\boldsymbol{p}:=(p_0,p_1,p_2,\dots)$. Consequently, we have

$$p(x)=\boldsymbol{p}\chi(x)\ \text{ and }\ p^{(k)}(x)=\boldsymbol{p}\chi^{(k)}(x).$$

For each $m \in \mathbb{N}$ (the directed set of natural numbers), we consider the ring of matrices $\mathbb{M}_m := \mathbb{R}^{m\times m}$, and its direct limit $\mathbb{M}_\infty := \lim_{m\to\infty} \mathbb{M}_m$, i.e. the ring of semi-infinite matrices. We will denote by G_∞ the group of invertible semi-infinite matrices of $\mathbb{M}_\infty$. A subgroup of G_∞ is $\mathscr{L}$, that of lower triangular matrices with the identity matrix along its main diagonal. Diagonal matrices will be denoted by $\mathscr{D} = \{M \in \mathbb{M}_\infty : d_{i,j} = d_i \cdot \delta_{i,j}\}$. We will also use the notation $E_{i,j}$ for indicating the matrix canonical basis, this is $(E_{i,j})_{l,m} = \delta_{i,l}\delta_{j,m}$.

Definition 6. *The Sobolev moment matrix associated to the measure matrix* $\mathscr{W}$ *is*

$$G_{\mathscr{W}} := \left(\chi, \chi^\top; \mathscr{W}\right) = \int_\Omega \boldsymbol{\chi}\ \mathscr{W}\ \boldsymbol{\chi}^\top \qquad (G_{\mathscr{W}})_{n,p} := (x^n, x^p; \mathscr{W}) \tag{5}$$

and its truncations will be denoted as

$$G_{\mathscr{W}}^{[k]} := \begin{pmatrix} (G_{\mathscr{W}})_{0,0} & (G_{\mathscr{W}})_{0,1} & \dots & (G_{\mathscr{W}})_{0,k-1} \\ (G_{\mathscr{W}})_{1,0} & (G_{\mathscr{W}})_{1,1} & \dots & (G_{\mathscr{W}})_{1,k-1} \\ \vdots & \vdots & & \\ (G_{\mathscr{W}})_{k-1,0} & (G_{\mathscr{W}})_{k-1,1} & \dots & (G_{\mathscr{W}})_{k-1,k-1} \end{pmatrix} = \int_\Omega \boldsymbol{\chi}^{[k]}\ \mathscr{W}^{[k]}\ \left(\boldsymbol{\chi}^{[k]}\right)^\top$$

By means of the previous notation, the Sobolev bilinear form of two polynomials $p(x), q(x) \in \mathbb{R}[x]$ can be rewritten as

$$(p, q; \mathscr{W}) = \boldsymbol{p} G_{\mathscr{W}} \boldsymbol{q}^\top.$$

The positive definiteness condition on the bilinear function is equivalent to that of $G_{\mathscr{W}}$, i.e., every principal minor of $G_{\mathscr{W}}$ must be greater than zero $\det[G_{\mathscr{W}}^{[k]}] > 0\ \forall k = 1, 2, \dots$ This condition will be discussed in detail later on.

Now we rewrite the moment matrix in a slightly different way, that will be more suitable for our purposes. To this aim, we introduce the derivation matrix $D \in \mathbb{M}_\infty$ defined by

$$D := \begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & 0 & \dots \\ 0 & 2 & 0 & 0 & \dots \\ 0 & 0 & 3 & 0 & \dots \\ 0 & 0 & 0 & 4 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

and its powers D^k, whose action is $D\chi(x) = \chi'(x)$, $D^k\chi(x) = \chi^{(k)}(x)$. We also introduce the shift operator

$$\Lambda := \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ 0 & 0 & 1 & 0 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots & \end{pmatrix}$$

whose action on χ is $\Lambda\chi(x) = x\chi(x)$, and on a polynomial $p(x)$ is $xp(x) = \boldsymbol{p}\Lambda\chi(x)$.
The shift and derivation matrices satisfy for any natural number n

$$\Lambda D^n - D^n\Lambda := [\Lambda, D^n] = nD^{n-1}$$

Definition 7. *We introduce the operator*

$$\boldsymbol{D} := \begin{pmatrix} \mathbb{I} & D & D^2 & \dots & D^k & \dots \end{pmatrix}.$$

It is immediate to verify that

$$\boldsymbol{D}\chi(x) = \boldsymbol{\chi}(x).$$

The following result is a direct consequence of the previous discussion.

Let us denote by $g_{k,r}$ the standard moment matrix associated to the measure $\mathrm{d}\mu_{k,r}$ (notice that $g_{i,j} = 0_{\infty\times\infty}$ is a null matrix when $\mathrm{d}\mu_{i,j} = 0$ and this is the case $\forall i,j > \mathcal{N}$).

Proposition 1. *The moment matrix admits the following representation*

$$G_{\mathscr{W}} = \boldsymbol{D}\begin{pmatrix} g_{0,0} & g_{0,1} & g_{0,2} & g_{0,3} & \cdots \\ g_{1,0} & g_{1,1} & g_{1,2} & g_{1,3} & \cdots \\ g_{2,0} & g_{2,1} & g_{2,2} & g_{2,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}\boldsymbol{D}^\top = \sum_{l,r=0}^{\mathcal{N}} D^l g_{l,r}(D^r)^\top \ , \tag{6}$$

with its truncations

$$G_{\mathscr{W}}^{[k]} = \boldsymbol{D}^{[k]}\begin{pmatrix} g_{0,0} & g_{0,1} & \cdots & g_{0,k-1} \\ g_{1,0} & g_{1,1} & \cdots & g_{1,k-1} \\ \vdots & \vdots & \ddots & \vdots \\ g_{k-1,0} & g_{k-1,1} & \cdots & g_{k-1,k-1} \end{pmatrix}\left(\boldsymbol{D}^{[k]}\right)^\top = \sum_{l,r=0}^{k-1} D^l g_{l,r}(D^r)^\top \ . \tag{7}$$

Proof. Using the previous definitions, for the expression 6 we can write

$$G_{\mathscr{W}} = (\chi,\chi^\top;\mathscr{W}) = \int_\Omega \sum_k\sum_r \chi^{(k)}(x)\mathrm{d}\mu_{k,r}\left(\chi^{(r)}(x)\right)^\top = \int_\Omega \sum_k\sum_r D^k\chi(x)\mathrm{d}\mu_{k,r}\left(D^r\chi(x)\right)^\top = \sum_k\sum_r D^k g_{k,r}\left(D^r\right)^\top$$

while relation 7 follows from the shape of the D^l. Since they are lower

$$\left(\sum_{l,r=0} D^l g_{l,r}(D^r)^\top\right)^{[k]} = \sum_{l,r=0}\left(D^l\right)^{[k]}\left(g_{l,r}\right)^{[k]}\left((D^r)^\top\right)^{[k]} \ ,$$

also observe that $\left(D^l\right)^{[k]} = 0\ \forall l \geq k$. □

This expression is a generalization of the case of a diagonal $\mathscr{W}$, already studied in [6] [23].

3. Sobolev bi-orthogonal polynomial sequences

3.1. **Main definitions and LU factorization.** To introduce the Sobolev bilinear function we have required a positive definiteness condition, which amounts to having every principal minor of $G_{\mathscr{W}}$ greater than zero. This requirement (quasi-definiteness would also be a valid choice) is necessary in order to use the LU factorization techniques of the moment matrix.

In the subsequent considerations, we shall assume that this condition for the minors of the moment matrix holds. Although in this paper we give some requirements on the set $\{\mathrm{d}\mu_{ij}\}_{i,j}$ that would assure definiteness of the associated moment matrix, a thorough analysis of this problem remains open.

In [6], a diagonal measure matrix $\mathscr{W}$ (i.e. $\mathrm{d}\mu_{i,j} = 0\ \forall i \neq j$) was considered. Choosing every $\mathrm{d}\mu_{i,i} := \mathrm{d}\mu_i$ as a positive definite measure makes the resulting Sobolev bilinear form a positive symmetric definite one and therefore a proper inner product.

This result can be easily interpreted in our framework. Observe that according to (6) the moment matrix for the diagonal case is

$$G_{\mathscr{W}} = g_0 + Dg_1D^\top + D^2g_2(D^2)^\top + D^3g_3(D^3)^\top + \dots.$$

If we introduce the matrix $N := diag\{1,2,3,\dots\}$, with the aid of (7), the truncation $G_{\mathscr{W}}^{[k]}$ reads

$$(G_{\mathscr{W}})^{[k]} = (g_0)^{[k]} + \left(\begin{array}{c|c} 0_{1\times 1} & 0 \\ \hline 0 & (Ng_1N)^{[k-1]} \end{array}\right) + \left(\begin{array}{c|c} 0_{2\times 2} & 0 \\ \hline 0 & (N^2g_2N^2)^{[k-2]} \end{array}\right) + \dots + \left(\begin{array}{c|c} 0_{k-1\times k-1} & 0 \\ \hline 0 & (N^{k-1}g_{k-1}N^{k-1})^{[1]} \end{array}\right)$$

The condition that $\mathrm{d}\mu_i$ be positive definite amounts to say that, given any vector $\boldsymbol{v} = (\boldsymbol{v}_0, \boldsymbol{v}_1, \dots, \boldsymbol{v}_{l-1})$, the associated quadratic form $\boldsymbol{v}(g_i)^{[l]}\boldsymbol{v}^\top$ satisfies $\boldsymbol{v}(g_i)^{[l]}\boldsymbol{v}^\top > 0\ \forall \boldsymbol{v}, l$. Therefore, in the computation of $\boldsymbol{v}(G_{\mathscr{W}})^{[k]}\boldsymbol{v}^\top$ only the sum of positive terms is involved; as a result $\boldsymbol{v}(G_{\mathscr{W}})^{[k]}\boldsymbol{v}^\top > 0\ \forall \boldsymbol{v}, k$, ensuring that $G_{\mathscr{W}}$ is positive definite and in turn LU factorizable.

We shall discuss now in the Sobolev context the main algebraic techniques of the present theory: the LU factorization approach for the moment matrix and the existence of bi-orthogonal sequences of polynomials.

Definition 8. *We shall say that the moment matrix $G_{\mathscr{W}}$ admits a LU factorization iff* $\det\left(G_{\mathscr{W}}^{[k]}\right) \neq 0\ \forall k = 1,2,\dots$*; in such a case there exist two matrices $S_1, S_2 \in \mathscr{L}$ such that*

$$G_{\mathscr{W}} := S_1^{-1} H \left(S_2^{-1}\right)^\top \ , \tag{8}$$

where $H := \delta_{r,k} h_k \in \mathscr{D}$.

Definition 9. *The monic SBPS associated with the LU-factorized moment matrix $G_{\mathscr{W}}$* (8) *are defined to be*

$$P_1(x) := S_1\chi(x) := \begin{pmatrix} P_{1,0}(x) \\ P_{1,1}(x) \\ \vdots \\ P_{1,k}(x) \\ \vdots \end{pmatrix}, \qquad P_2(x) := S_2\chi(x) := \begin{pmatrix} P_{2,0}(x) \\ P_{2,1}(x) \\ \vdots \\ P_{2,k}(x) \\ \vdots \end{pmatrix}. \tag{9}$$

As a well known consequence of the previous definitions expressing our polynomials in terms of the LU factorization matrices, we can write the following compact relations.

Proposition 2. *The SBPS can be expressed by means of the following quasi-determinantal formulae*

$$P_{1,k}(x) = \Theta_* \left[\begin{array}{cccc|c} & & & & 1 \\ & G_{\mathscr{W}}^{[k]} & & & x \\ & & & & \vdots \\ & & & & x^{k-1} \\ \hline (G_{\mathscr{W}})_{k,0} & (G_{\mathscr{W}})_{k,1} & \dots & (G_{\mathscr{W}})_{k,k-1} & x^k \end{array}\right], \tag{10}$$

$$P_{2,k}(x) = \Theta_* \left[\begin{array}{cccc|c} & & & & 1 \\ & \left(G_{\mathscr{W}}^\top\right)^{[k]} & & & x \\ & & & & \vdots \\ & & & & x^{k-1} \\ \hline (G_{\mathscr{W}}^\top)_{k,0} & (G_{\mathscr{W}}^\top)_{k,1} & \dots & (G_{\mathscr{W}}^\top)_{k,k-1} & x^k \end{array}\right]. \tag{11}$$

Notice that the definition ensures that $deg[P_{\alpha,k}] = k\ \ \alpha = 1,2\ \ \forall k = 0,1,\dots$ while the condition on the minors of $G_{\mathscr{W}}$ guarantees that the definition always makes sense.

Here we have used the notation $\Theta_*[M]$ to denote the *last quasi-determinant* or *Schur complement* of the matrix in brackets. More precisely, we recall that given $M = \left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \in \mathbb{M}_{(n+m)}$ with $A \in \mathbb{M}_n$, $\det(A) \neq 0$ and

$D \in \mathbb{M}_m$ its last quasi-determinant or Schur complement with respect to A is given by

$$\Theta_* \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := SC(M) := M/A := D - CA^{-1}B$$

It is worth observing that the block Gauss factorization of M involves the last quasi-determinant

$$M = \begin{pmatrix} \mathbb{I}_n & 0 \\ CA^{-1} & \mathbb{I}_m \end{pmatrix} \begin{pmatrix} A & 0 \\ 0 & \Theta_*[M] \end{pmatrix} \begin{pmatrix} \mathbb{I}_n & AB^{-1} \\ 0 & \mathbb{I}_m \end{pmatrix}.$$

From the previous relation one can immediately deduce that

$$\det \left(\Theta_* \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \right) = \frac{\det(M)}{\det(A)}.$$

Therefore, whenever $m = 1$, D reduces to a scalar d , and the quasi-determinants are a ratio of standard determinants

$$\Theta_* \left[\begin{array}{c|c} A & B \\ \hline C & d \end{array}\right] = \frac{\det(M)}{\det(A)} .$$

This indeed is the situation we will deal with. However, we prefer to use quasi-determinants since the relations we obtain will be ready for further generalizations of the theory (matrix Sobolev, multivariate Sobolev), where the expressions in terms of determinants would no longer hold. For further details on the theory of quasi-determinants, see [27].

The following proposition clarifies the notion of bi-orthogonality for SBPS.

Proposition 3. *The monic SBPS P_1 and P_2 are Sobolev-bi-orthogonal, that is, they satisfy the relation*

$$(P_{1,r}, P_{2,k}; \mathscr{W}) := h_r \delta_{r,k}$$

with the further properties

$$(P_{1,l}, x^r; \mathscr{W}) := \delta_{l,r} h_r \quad \forall r \leq l \qquad \Longrightarrow \qquad \sum_{k=0}^{l} \sum_{j=0}^{r} \left\langle P_{1,l}^{(k)}, \frac{\mathrm{d}^j x^r}{\mathrm{d}x^j} \right\rangle_{k,j} = \begin{cases} 0 & \forall r < l \\ h_l & r = l \end{cases}$$

$$(x^r, P_{2,l},; \mathscr{W}) := h_r \delta_{r,l} \quad \forall r \leq l \qquad \Longrightarrow \qquad \sum_{k=0}^{r} \sum_{j=0}^{l} \left\langle \frac{\mathrm{d}^j x^r}{\mathrm{d}x^j}, P_{2,l}^{(k)} \right\rangle_{j,k} = \begin{cases} 0 & \forall r < l \\ 1 & r = l \end{cases}$$

Proof. The previous relations are a direct consequence of the LU factorization of the moment matrix $G_{\mathscr{W}}$. □

Definition 10. *Let $f(x) = \frac{1}{y-x}$ belong to the subspace $\mathscr{A}_{\mathscr{W}}^{\mathcal{N}}(\Omega)$. Then we introduce the second kind functions*

$$C_{1,l}(y) := \int_\Omega \sum_{k=0}^{l} \sum_{j=0}^{\mathcal{N}} P_{1,l}^{(k)}(x) \mathrm{d}\mu_{k,j} \left[\frac{\partial^j}{\partial x^j} \left(\frac{1}{y-x} \right) \right] = \left(P_{1,l}(x), \frac{1}{y-x}; \mathscr{W}(x) \right) , \qquad y \notin \Omega ,$$

$$C_{2,l}(y) := \int_\Omega \sum_{k=0}^{\mathcal{N}} \sum_{j=0}^{l} \left[\frac{\partial^j}{\partial x^j} \left(\frac{1}{y-x} \right) \right] \mathrm{d}\mu_{j,k} P_{2,l}^{(k)}(x) = \left(\frac{1}{y-x}, P_{2,l}(x); \mathscr{W}(x) \right) , \qquad y \notin \Omega .$$

Proposition 4. *The associated Sobolev second kind functions $C_\alpha(y)$ admit the following representation in terms of the LU factorization matrices for all y such that $|y| > max\{|x|, x \in \Omega\}$*

$$C_1(y) = H(S_2^{-1})^\top \chi^*(y) := \begin{pmatrix} C_{1,0}(y) \\ C_{1,1}(y) \\ \vdots \\ C_{1,k}(y) \\ \vdots \end{pmatrix}, \qquad C_2(y) = H(S_1^{-1})^\top \chi^*(y) := \begin{pmatrix} C_{2,0}(y) \\ C_{2,1}(y) \\ \vdots \\ C_{2,k}(y) \\ \vdots \end{pmatrix}.$$

Proof. In order to prove any of the two expressions it is enough to observe that whenever $\forall |x| < |y|$

$$\chi(x)^\top \cdot \chi(y)^* = \frac{1}{y} \sum_{n=0}^{\infty} \left(\frac{x}{y}\right)^n = \frac{1}{y-x}.$$

Also, since the given expressions in the proposition can be rewritten as

$$C_1(y) = S_1 G_{\mathscr{W}} \chi^*(y), \qquad (C_2(y))^\top = (\chi^*(y))^\top G_{\mathscr{W}} S_2^\top,$$

we deduce that, for example for C_1

$$C_1(y) = \quad S_1 G_{\mathscr{W}} \chi^*(y) = S_1 \int_\Omega \boldsymbol{\chi} \ \mathscr{W} \ \boldsymbol{\chi}^\top \cdot \chi^*(y) = \int_\Omega \begin{pmatrix} P_1(x) & P_1'(x) & \dots & P_1^{(k)}(x) & \dots \end{pmatrix} \mathscr{W} \begin{pmatrix} \frac{1}{y-x} \\ \frac{\partial}{\partial x}\left(\frac{1}{y-x}\right) \\ \vdots \\ \frac{\partial^j}{\partial x^j}\left(\frac{1}{y-x}\right) \\ \vdots \end{pmatrix}$$

and similarly for $C_2(y)$. □

A natural question is to establish the relation between the SBPS (and associated second kind functions) that arise from a given measure matrix $\mathscr{W}$ and the ones associated with its transposed $\mathscr{W}^\top$. A simple answer is provided by the following

Proposition 5. *Let $P_{\mathscr{W},\alpha}$ and $C_{\mathscr{W},\alpha}$ with $\alpha = 1,2$ denote the SBPS and second kind functions that arise from the measure matrix $\mathscr{W}$ and $P_{\mathscr{W}^\top,\alpha}$ and $C_{\mathscr{W}^\top,\alpha}$ the ones corresponding to $\mathscr{W}^\top$. Then we have*

$$P_{\mathscr{W},1} = P_{\mathscr{W}^\top,2} \qquad P_{\mathscr{W},2} = P_{\mathscr{W}^\top,1}$$
$$C_{\mathscr{W},1} = C_{\mathscr{W}^\top,2} \qquad C_{\mathscr{W},2} = C_{\mathscr{W}^\top,1}$$

Proof. It is straightforward to see that $G_{\mathscr{W}^\top} = G_{\mathscr{W}}^\top$. The assumption of the *LU* factorization property for the moment matrix implies the proposition. □

The previous proposition implies that if $\mathscr{W} = \mathscr{W}^\top$ then $P_{\mathscr{W},1} = P_{\mathscr{W},2}$ and $C_{\mathscr{W},1} = C_{\mathscr{W},2}$,(usually studied case) as expected since in such a case the *LU* factorization is indeed a Cholesky factorization.

3.2. **Christoffel-Darboux Kernels.** The Christoffel-Darboux and Cauchy kernels will play a crucial role in the following considerations. We present here their formal definition in our context.

Definition 11. *We introduce the Christoffel–Darboux kernel, the Cauchy kernel, and the first and second kind mixed Christoffel–Darboux kernels, given by*

- *Christoffel–Darboux kernel*

$$K^{[l]}(x,y) := \sum_{k=0}^{l-1} P_{2k}(x) h_k^{-1} P_{1k}(y) = [P_2(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [P_1(y)]^{[l]} = \left(\chi(x)^{[l]}\right)^\top \left(G^{[l]}\right)^{-1} \chi(y)^{[l]} \ ,$$

- *Cauchy CD kernel*

$$Q^{[l]}(x,y) := \sum_{k=0}^{l-1} C_{2k}(x) h_k^{-1} C_{1k}(y) = [C_2(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [C_1(y)]^{[l]} = \left(\chi^*(x)^{[l]}\right)^\top \left(G^{[l]}\right) \chi^*(y) \ ,$$

- *Mixed 1st CD kernel*

$$\mathcal{K}_1^{[l]}(x,y) := \sum_{k=0}^{l-1} C_{2k}(x) h_k^{-1} P_{1k}(y) = [C_2(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [P_1(y)]^{[l]} = (\chi(x)^*)^\top \left(\frac{\mathbb{I}_{l\times l}}{(S_1^{-1})^{[\geq l,j]}(S_1^{-1})^{[l]}}\right) \chi^{[l]}(y) \ ,$$

- *Mixed 2nd CD kernel*

$$\mathcal{K}_2^{[l]}(x,y) := \sum_{k=0}^{l-1} P_{2k}(x) h_k^{-1} C_{1k}(y) = [P_2(x)^\top]^{[l]} \left(H^{-1}\right)^{[l]} [C_1(y)]^{[l]} = \left(\chi(x)^{[l]}\right)^\top \left(\ \mathbb{I}_{l\times l} \ \middle| \ (S_2^\top)^{[l]}([S_2^\top]^{-1})^{[j,\geq l]} \ \right) \chi^*(y) \ .$$

Remark 2. *In the previous definition, the expressions of the standard Bezoutian kernels are not present. They would involve only two consecutive orthogonal polynomials (or second kind functions) instead of all polynomials up to the degree of the kernel. The lack of this expression is not surprising, since the Bezoutian kernels would correspond to having a three term recurrence relation for the orthogonal polynomials (and second kind functions), that in principle is missing. Despite that, all of the expected properties of the CD kernel still hold. This is, the CD Kernel still has the reproducing property,*

$$\left(K^{[l]}(x,z), K^{[l]}(z,y)\right)_{\mathscr{W}} = \left(\chi^{[l]}(x)\right)^\top \left(G_{\mathscr{W}}^{[l]}\right)^{-1} \left[\int_\Omega \boldsymbol{\chi}^{[l]}(z)\mathscr{W}(z)\left(\boldsymbol{\chi}^{[l]}(z)\right)^\top\right] \left(G_{\mathscr{W}}^{[l]}\right)^{-1} \chi^{[l]}(y) = K^{[l]}(x,y)$$

and acts as a projector onto the basis of the SBPS. Therefore, given any function $f(x) \in \mathscr{A}_{\mathscr{W}}^{\mathcal{N}}(\Omega)$, one has

$$\Pi_1^{[l]}[f(y)] = \left(f(x), K^{[l]}(x,y)\right)_{\mathscr{W}} = \sum_{k}^{l-1} \left[(f, P_{2,k})_{\mathscr{W}}\, h_k^{-1}\right] P_{1,k}(y)$$

$$\Pi_2^{[l]}[f(x)] = \left(K^{[l]}(x,y), f(y)\right)_{\mathscr{W}} = \sum_{k}^{l-1} P_{2,k}(x) \left[h_k^{-1}\left(P_{1,k}, f\right)_{\mathscr{W}}\right]$$

where we call $\Pi_\alpha^{[l]}[f(x)]$ the best approximation of f (in $(*,*)_{\mathscr{W}}$) in the basis $\{P_{\alpha,l}\}_{k=0}^{(l-1)}$ for $\alpha = \{1,2\}$. Notice also that when $\mathscr{W}$ is symmetric, only one of the two mixed kernels is needed (no distinction between subindices $1,2$ exists).

4. Additive perturbations of the measure matrix

In this section, we are interested in the following problem: Given the pairing (G,g), where G is a moment matrix whose associated SBPS is known, and g is another matrix, find the SBPS associated to the new moment matrix $\breve{G} = G + g$.

The same problem, although from a different point of view, was also studied in [4]. The results proposed in the present work, when they are equivalent, possess alternative proofs. At the same time, they are suited for the Sobolev context.

Generally speaking, the solution to this problem leads to interesting cases when we require that g has some special features. Two nontrivial examples are indeed the cases of *coherent pairs* and of *discrete Sobolev bilinear functions.*

Since their appearance [13], coherent pairs have been largely investigated in the literature. In this work, we will limit ourselves to show how coherent pairs fit within our framework. Instead, the discrete Sobolev bilinear forms will be of considerable relevance in our subsequent discussion; therefore, we will pay special attention to them.

As a starting point of our analysis, suppose that our moment matrix can be written as $\breve{G} = G + g$. Since we assume that G has an associated SBPS, then it must be LU-factorizable; at the same time, the requirement that the SBPS associated to $\breve{G}$ exists implies that the latter matrix should be LU-factorizable too. Therefore, we deduce the relation

$$\breve{S}_1^{-1}\breve{H}\left(\breve{S}_2^{-1}\right)^\top = S_1^{-1}H\left(S_2^{-1}\right)^\top + g. \tag{12}$$

This motivates the following

Definition 12. *We introduce the matrices*

$$A := S_1 g S_2^\top \qquad M_1 := \breve{S}_1 S_1^{-1} \qquad M_2 := \breve{S}_2 S_2^{-1}$$

Proposition 6. *The matrices M_1, M_2 are the connection matrices between old and new polynomials*

$$M_1 P_1(x) = \breve{P}_1(x) \qquad M_2 P_2(x) = \breve{P}_2(x)$$

and provide an LU factorization of the matrix $H + A$:

$$M_1^{-1}H\left(M_2^{-1}\right)^\top = H + A$$

Proof. The result follows from the requirement that both $\breve{G}, G$ admit an LU factorization and from the observation that, by definition, both M_1, M_2 are lower uni-triangular. □

This last proposition allows to derive directly the following consequence.

Proposition 7. *The basis change from the old SBPS to the new one is given*

$$\breve{P}_{1,k}(x) = \Theta_*\left[\begin{array}{cccc|c} \multicolumn{4}{c|}{\multirow{4}{*}{$(H+A)^{[k]}$}} & P_{1,0}(x) \\ &&&& P_{1,1}(x) \\ &&&& \vdots \\ &&&& P_{1,k-1}(x) \\ \hline (A)_{k,0} & (A)_{k,1} & \dots & (A)_{k,k-1} & P_{1,k}(x) \end{array}\right],$$

$$\breve{P}_{2,k}(x) = \Theta_*\left[\begin{array}{cccc|c} \multicolumn{4}{c|}{\multirow{4}{*}{$(H+A)^{[k]}$}} & A_{0,k}(x) \\ &&&& A_{1,k}(x) \\ &&&& \vdots \\ &&&& A_{k,k-1}(x) \\ \hline P_{2,0} & P_{2,1} & \dots & P_{2,k-1} & P_{2,k}(x) \end{array}\right],$$

$$\breve{h}_k = \Theta_*\left[\begin{array}{cccc|c} \multicolumn{4}{c|}{\multirow{4}{*}{$(H+A)^{[k]}$}} & (A)_{0,k} \\ &&&& (A)_{1,k} \\ &&&& \vdots \\ &&&& (A)_{k-1,k} \\ \hline (A)_{k,0} & (A)_{k,1} & \dots & (A)_{k,k-1} & (H+A)_{k,k} \end{array}\right]$$

We shall use this result and focus now on three cases. Firstly we will deal with the situation where G and g are the moment matrices associated to a pair of related classical measures. Secondly we will consider the case when $g=\lambda D g_2 D^\top$ and $G=g_1$ where g_1,g_2 are the moment matrices associated to a couple of measures that form a coherent pair. Finally we will study the case where g is associated to a discrete Sobolev bilinear function.

4.1. **A first relation with classical OPS.** It is a well known fact that classical orthogonal polynomials can be regarded as a very specific case of SOPS. As we are about to see, a consequence of this is that the previous relations become almost trivial when choosing the right measures.

If we denote the classical measures by u_γ, where γ refers to the parameters that define them, they are

- Hermite $u(x)=e^{-x^2}$, $x\in\mathbb{R}$; $(\gamma=\{\varnothing\})$.
- Laguerre $u_\alpha(x)=x^\alpha e^{-x}$, $\alpha>-1$, $x\in\mathbb{R}_+$; $(\gamma=\{\alpha\})$.
- Jacobi $u_{\alpha,\beta}(x)=(1-x)^\alpha(1+x)^\beta$, $\alpha,\beta>-1$, $x\in(-1,1)$; $(\gamma=\{\alpha,\beta\})$.

We will use $P_\gamma(x)=S_\gamma\chi(x)$ to denote the monic orthogonal polynomials $\{P_{\gamma,n}\}_n$ associated to each of them in terms of the LU factorization matrices S_γ of the corresponding moment matrix g_γ.
There are many ways to characterize classical measures; the one that is suited for our purposes is to express them in terms of a Pearson differential equation:

$$p_2(x)\frac{\mathrm{d}u_\gamma}{\mathrm{d}x}=p_{1,\gamma}(x)u_\gamma \qquad p_2^k(x)u_\gamma=u_{\gamma+k} \qquad \text{where } deg[p_2]\leq 2 \text{ and } deg[p_{1,\gamma}]=1.$$

- Hermite $p_1=-2x$, $p_2=1$.
- Laguerre $p_{1,\alpha}=(\alpha-x)$, $p_2=x$.
- Jacobi $p_{1,\alpha,\beta}=-[(\alpha-\beta)+(\alpha+\beta)x]$, $p_2=1-x^2$.

This equation is relevant in the discussion of many properties of the associated OPS. In particular, it implies that $P_{(\gamma+1),n}(x)=\frac{P'_{\gamma,n+1}(x)}{n+1}$, which in matrix form gives the crucial relation $D=S_\gamma D S_{\gamma+1}^{-1}$.

As a simple example, consider the following Sobolev inner product

$$(f,h)=\int f(x)h(x)u_\gamma(x)\mathrm{d}x+\lambda\int f'(x)h'(x)u_{\gamma+1}(x)\mathrm{d}x \qquad \lambda>0,$$

that we wish to interpret as an additive perturbation $\breve{G}=G+g$ with the identifications $G=g_\gamma$ and $g=\lambda D g_{\gamma+1}D^\top$. The crucial relation $D=S_\gamma D S_{\gamma+1}^{-1}$ implies for A the particularly simple form

$$A=\lambda S_\gamma D S_{\gamma+1}^{-1}H_{\gamma+1}\left(S_\gamma D S_{\gamma+1}^{-1}\right)^\top=\lambda D H_{\gamma+1}D^\top=\lambda\begin{pmatrix}0 &&&&&\\ & 1^2h_{\gamma+1,0} &&&&\\ && 2^2h_{\gamma+1,1} &&&\\ &&& \ddots &&\\ &&&& k^2h_{\gamma+1,k-1} &\\ &&&&& \ddots\end{pmatrix},$$

which makes the quasi-determinantal expressions in Proposition 7 almost trivial.

Corollary 1. *The SBPS $\breve{P}_k$ and norms $\breve{h}_k$ for the following inner product*

$$(f,h)=\int f(x)h(x)u_\gamma(x)\mathrm{d}x+\lambda\int f'(x)h'(x)u_{\gamma+1}(x)\mathrm{d}x \qquad \lambda>0$$

are given by

$$\breve{P}_k(x)=P_{\gamma,k}(x) \qquad \breve{h}_k=h_{\gamma,k}+\lambda k^2 h_{\gamma+1,k-1}$$

For future reference we shall also discuss here a couple of additional properties of classical OPS. They will be useful below, in relation with the study of equivalence classes of measure matrices.
i) It is almost straightforward to see that

$$p_2^k(x)u_\gamma\mid_{\partial\Omega}=u_{\gamma+k}\mid_{\partial\Omega}=0 \qquad \forall k\geq 1\ .$$

In the previous relation, one can allow for even smaller values of k, depending on the value of γ. However, to avoid the worst case $\gamma=-1$, taking $k\geq 1$ will be sufficient in all situations.
ii) A less trivial property is expressed by the following

Proposition 8. *The measure $u_{\gamma+k}$ satisfies the relations*

$$\begin{aligned}\frac{\mathrm{d}^r}{\mathrm{d}x^r}\left(p_2^k u_\gamma\right)&=\varphi_{k,r}(x)u_\gamma && 0\leq r\leq k\\ \varphi_{k,r}u_\gamma\mid_{\partial\Omega}&=0 && 0\leq r\leq (k-1)\end{aligned}$$

where $\varphi_{k,r}(x)$ is a suitable polynomial.

Proof. Let Q be a $\mathcal{C}^k$ function; and taking into account the Pearson equation it is easy to see that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left[Qp_2^k u_\gamma\right]=Q'p_2^k u_\gamma+Qkp_2'p_2^{k-1}u_\gamma+Qp_2^k\frac{p_{1,\gamma}}{p_2}u_\gamma=\mathcal{O}_k\left[Q\right]p_2^{k-1}u_\gamma\ ,$$

where $\mathcal{O}_k:=p_2\frac{\mathrm{d}}{\mathrm{d}x}+[kp_2'+p_{1,\gamma}]$ is a first order linear differential operator. Differentiating the previous relation we have

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\left[Qp_2^k u_\gamma\right]=\frac{\mathrm{d}}{\mathrm{d}x}\left[\mathcal{O}_k[Q]p_2^{k-1}u_\gamma\right]\ .=\mathcal{O}_{k-1}\circ\mathcal{O}_k[Q]p_2^{k-2}u_\gamma$$

Therefore, differentiating r times one gets

$$\frac{\mathrm{d}^r}{\mathrm{d}x^r}\left[Qp_2^k u_\gamma\right]=\mathcal{O}_{k-(r-1)}\circ\mathcal{O}_{k-(r-2)}\cdots\circ\mathcal{O}_k[Q]p_2^{k-r}u_\gamma:=\mathcal{O}_k^{k-(r-1)}[Q]p_2^{k-r}u_\gamma$$

Notice that we have defined the operator $\mathcal{O}_k^j[f]$, $\forall j\leq k$, but in order to make our notation a bit more compact let us add to this definition the case $\mathcal{O}_k^{k+1}[f]:=f$ as the identity operator, this way for $0\leq r\leq k$

$$\varphi_{k,r}:=\mathcal{O}_k^{k-(r-1)}[1]p_2^{k-r}$$

(note that according to the definition of $\mathcal{O}_k^{k+1}[f]:=f$ we would have $\varphi_{k,0}=p_2^k$) and now from *i*) the proposition is proven.

□

4.2. **Coherent Pairs.** We are interested in obtaining the SBPS associated to the inner product

$$(f,h)_c:=\int f(x)h(x)\mathrm{d}\mu_1(x)+\lambda\int f'(x)h'(x)\mathrm{d}\mu_2(x) \qquad \lambda>0, \tag{13}$$

where $\mathrm{d}\mu_1(x)$ and $\mathrm{d}\mu_2(x)$ form a *coherent pair of measures.* This inner product, in terms of moment matrices reads

$$\breve{G}=g_1+\lambda Dg_2D^\top$$

and therefore can be studied from the additive perturbation approach. Let us introduce some notation for the moment matrices, their factorization and corresponding OPS. For each of the two involved measures we will denote:

$$\begin{aligned}\mathrm{d}\mu_1(x)&\longrightarrow g_1=S^{-1}H\left(S^{-1}\right)^T\longrightarrow P(x)=S\chi(x)\\ \mathrm{d}\mu_2(x)&\longrightarrow g_2=Z^{-1}K\left(Z^{-1}\right)^T\longrightarrow Q(x)=Z\chi(x)\ .\end{aligned}$$

One of the possible characterizations of a coherent pair is given in terms of a relation between the OPS associated to each of the measures. Precisely, it is said that $\mathrm{d}\mu_1(x)$ and $\mathrm{d}\mu_2(x)$ form a coherent pair if there exist some non zero constants $\{r_k\}_{k=1}^{\infty}$ such that

$$Q_k(x) = \frac{1}{k+1}P'_{k+1}(x) - \frac{r_k}{k}P'_k(x) \qquad \forall k = 1, 2, \dots$$

It is worth pointing out that the coefficient that goes with $P'_{k+1}(x)$ is chosen according to the fact that we wish to generate monic orthogonal polynomials, while both the sign and coefficient that go with $r_k P'_k(x)$ are selected for convenience. To interpret this construction as an additive perturbation and using the notation presented in definition 12 we have

$$A = \lambda \left(SDZ^{-1}\right) K \left(SDZ^{-1}\right)^{\top}$$

We introduce the lower matrix R^{-1} according to the formulae

$$SDZ^{-1} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ 1 & 0 & 0 & \dots \\ * & 2 & 0 & \dots \\ * & * & 3 & 0 \\ \vdots & \vdots & & \ddots \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & \dots \\ (R^{-1})_{0,0} & 0 & 0 & \dots \\ (R^{-1})_{1,0} & (R^{-1})_{1,1} & 0 & \dots \\ (R^{-1})_{2,0} & (R^{-1})_{2,1} & (R^{-1})_{2,2} & \ddots \\ \vdots & \vdots & \vdots & \end{pmatrix}, \quad R^{-1} := \begin{pmatrix} 1 & 0 & 0 & \dots \\ (R^{-1})_{1,0} & 2 & 0 & \dots \\ (R^{-1})_{2,0} & (R^{-1})_{2,1} & 3 & \dots \\ \vdots & \vdots & & \ddots \end{pmatrix}$$

This last definition is motivated by the fact that it allows to write the truncations of A as

$$A^{[k]} = \left(\begin{array}{c|c} 0 & \mathbf{0} \\ \hline \mathbf{0}^{\top} & \lambda\left(R^{-1}K\left(R^{-1}\right)^{T}\right)^{[k-1]} \end{array}\right) = \left(\begin{array}{c|c} 0 & \mathbf{0} \\ \hline \mathbf{0}^{\top} & \lambda\left(R^{[k-1]}\right)^{-1} K^{[k-1]} \left(\left(R^{[k-1]}\right)^{-1}\right)^{T} \end{array}\right)$$

We have used here $\mathbf{0}$ for a row of zeroes. The second equality holds due to the lower triangular shape of R. By means of Proposition 7, we deduce the following expression for the SBPS:

$$\check{P}_k = P_k(x) - \lambda \left(\left(R^{-1}K\left(R^{-1}\right)^{T}\right)^{[k]}_{k-1,0} \quad \dots \quad \left(R^{-1}K\left(R^{-1}\right)^{T}\right)^{[k]}_{k-1,k-2} \right) \left[\left(R^{-1}K\left(R^{-1}\right)^{T}\right)^{[k-1]} + \tilde{H}^{[k-1]} \right]^{-1} \begin{pmatrix} P_1(x) \\ P_2(x) \\ \vdots \\ P_{k-1}(x) \end{pmatrix}$$

Here $\tilde{H}^{[k-1]} := diag\{H_1, H_2, \dots, H_{k-2}\}$. This is a general result that would be valid for any inner product of the form (13). In order to simplify it, we will use the fact that we are working with coherent pairs in order to find a simple expression for $\left(R^{-1}KR^{-T}\right)^{[k]}$. To this aim, remember that

$$SDZ^{-1}Q(x) = P'(x) \qquad \Rightarrow \qquad R^{-1}\begin{pmatrix} Q_0 \\ Q_1 \\ Q_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} P'_1 \\ P'_2 \\ P'_3 \\ \vdots \end{pmatrix} \qquad \Rightarrow \qquad \begin{pmatrix} Q_0 \\ Q_1 \\ Q_2 \\ \vdots \end{pmatrix} = R\begin{pmatrix} P'_1 \\ P'_2 \\ P'_3 \\ \vdots \end{pmatrix}$$

At the same time, due to coherence property, we know that R has a particularly simple lower bi-diagonal shape

$$R = \begin{pmatrix} 1 & & & \\ -\frac{r_1}{1} & \frac{1}{2} & & \\ & -\frac{r_2}{2} & \frac{1}{3} & \\ & & -\frac{r_3}{3} & \ddots \\ & & & \ddots \end{pmatrix}$$

It is now easy to see that after introducing the matrices

$$r := \begin{pmatrix} 0 & & & & \\ r_1 & 0 & & & \\ & r_2 & 0 & & \\ & & r_3 & \ddots & \\ & & & \ddots & \end{pmatrix} \qquad\qquad N := \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & \ddots \end{pmatrix}$$

one obtains

$$RN = (\mathbb{I} - r) \quad \Longrightarrow \quad R^{-1} = N\left(\mathbb{I} - r\right)^{-1} = N\left(\mathbb{I} + r + r^2 + \dots\right) \quad \Longrightarrow \left(R^{[k]}\right)^{-1} = N^{[k]}\left(\mathbb{I}^{[k]} + r^{[k]} + \dots + (r^{k-1})^{[k]}\right)$$

Therefore

$$\lambda\left(R^{-1}K\left(R^{-1}\right)^T\right)^{[k]} = \lambda N^{[k]}\left(\mathbb{I}^{[k]} + r^{[k]} + (r^2)^{[k]} + \dots + (r^{k-1})^{[k]}\right)K^{[k]}\left(\mathbb{I}^{[k]} + r^{[k]} + (r^2)^{[k]} + \dots + (r^{k-1})^{[k]}\right)^{\top} N^{[k]}$$

which finally implies that the $\breve{P}_k$ depend only on the first $k-1$ parameters $\{r_1, r_2, \dots, r_{k-1}\}$ that characterized the coherence and the norms of the original polynomials. For instance, consider

$$\lambda\left(R^{-1}K\left(R^{-1}\right)^T\right)^{[3]} = \lambda\begin{pmatrix} K_0 & 2r_1K_0 & 3r_2r_1K_0 \\ 2r_1K_0 & 2^2(r_1^2K_0 + K_1) & 2\cdot 3(r_1^2r_2K_0 + r_2K_1) \\ 3r_2r_1K_0 & 2\cdot 3(r_1^2r_2K_0 + r_2K_1) & 3^2(r_1^2r_2^2K_0 + r_2^2K_1 + K_2) \end{pmatrix}$$

which yields

$$\begin{aligned} \breve{P}_0 &= P_0 \\ \breve{P}_1 &= P_1 \\ \breve{P}_2 &= P_2 - \lambda(2r_1K_0)[\lambda K_0 + H_1]^{-1}P_1 \\ \breve{P}_3 &= P_3 - \lambda\begin{pmatrix} 3r_2r_1K_0 & 2\cdot 3(r_1^2r_2K_0 + r_2K_1)\end{pmatrix}\begin{pmatrix} K_0 + H_1 & 2r_1K_0 \\ 2r_1K_0 & 2^2(r_1^2K_0 + K_1) + H_2 \end{pmatrix}^{-1}\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} \end{aligned}$$

Observe that the previous nice expressions for the Sobolev polynomials are just a consequence of the lower bi–diagonal structure of R (which came from the characterization of the coherent pair $\{\mathrm{d}\mu_1, \mathrm{d}\mu_2\}$ in terms of their associated OPS).

A possible generalization of the notion of coherent pairs can be obtained by considering a bi–$m \times m$ block diagonal R and proceeding in the same way. This suggests the following

Definition 13. *We shall say that $\{\mathrm{d}\mu_1, \mathrm{d}\mu_2\}$ form a $m \times m$ block coherent pair if their associated OPS are related as follows*

$$\begin{pmatrix} Q_0 \\ Q_1 \\ \vdots \\ Q_{m-1} \end{pmatrix} = (R_m)_{[0][0]}\begin{pmatrix} P'_1 \\ P'_2 \\ \vdots \\ P'_m \end{pmatrix}, \quad \begin{pmatrix} Q_{km} \\ Q_{km+1} \\ \vdots \\ Q_{km+m-1} \end{pmatrix} = (R_m)_{[k][k-1]}\begin{pmatrix} P'_{(k-1)m+1} \\ P'_{(k-1)m+2} \\ \vdots \\ P'_{(k-1)m+m} \end{pmatrix} + (R_m)_{[k][k]}\begin{pmatrix} P'_{km+1} \\ P'_{km+2} \\ \vdots \\ P'_{km+m} \end{pmatrix} \quad \forall k \geq 1$$

where $(R_m)_{[k][k-1]}, (R_m)_{[k][k]} \in \mathbb{M}_m$ and

$$(R_m)_{[k][k]} = \begin{pmatrix} \frac{1}{km+1} & & & \\ * & \frac{1}{km+2} & & \\ \vdots & \vdots & \ddots & \\ * & * & \dots & \frac{1}{(k+1)m} \end{pmatrix}.$$

Note that the case $m=1$ reproduces just the standard concept of coherent pairs that we treated before. The case $m=2$ contains as a particular case the symmetrically coherent pairs since

$$\begin{pmatrix} Q_{2k} \\ Q_{2k+1} \end{pmatrix} = \begin{pmatrix} * & 0 \\ 0 & * \end{pmatrix} \begin{pmatrix} P'_{2k-1} \\ P'_{2k} \end{pmatrix} + \begin{pmatrix} \frac{1}{2k+1} & 0 \\ 0 & \frac{1}{2k+1} \end{pmatrix} \begin{pmatrix} P'_{2k+1} \\ P'_{2k+2} \end{pmatrix}$$

The way to proceed for a general m would follow the same steps as the case $m=1$. Firstly define

$$N_m := \begin{pmatrix} (R_m)^{-1}_{[0][0]} & & \\ & (R_m)^{-1}_{[1][1]} & \\ & & \ddots \end{pmatrix}, \qquad r_m := \begin{pmatrix} 0 & & & \\ (r_m)_{[1][0]} & 0 & & \\ & (r_m)_{[2][1]} & 0 & \\ & & & \ddots \end{pmatrix},$$

$$(r_m)_{[k][k-1]} = -\,(R_m)_{[k][k-1]}\,(R_m)^{-1}_{[k-1][k-1]} \,.$$

Therefore $R_m N_m = \mathbb{I} - r_m$ and taking the inverse of its truncations one obtains

$$\left(R_m^{[km]}\right)^{-1} = N_m^{[km]}\left(\mathbb{I}^{[km]} + r_m^{[km]} + (r_m^2)^{[km]} + \cdots + (r_m^{k-1})^{[km]}\right) ,$$

which would allow us to write the associated SOPS only in terms of the entries of the matrices that characterized the $m\times m$ block coherent pair.

An open problem is to construct examples of $m\times m$ block coherent pairs. An illustrative example is offered by the previously mentioned symmetrically coherent pair case, which has $m=2$. Let us take $k=2$. In this case we have

$$\begin{aligned}
\left(R_2^{[2\cdot 2]}\right)^{-1} &= \begin{pmatrix} 1 & & & \\ & 2 & & \\ & & 3 & \\ & & & 4 \end{pmatrix} \left[\mathbb{I}_{4\times 4} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ r_2 & 0 & 0 & 0 \\ 0 & r_3 & 0 & 0 \end{pmatrix}\right] \\
\Longrightarrow \lambda\left(R^{-1}K\left(R^{-1}\right)^T\right)^{[4]} &= \lambda \begin{pmatrix} K_0 & 0 & 3K_0r_2 & 0 \\ 0 & 4K_1 & 0 & 8K_1r_3 \\ 3K_0r_2 & 0 & 9(K_2+K_0r_2^2) & 0 \\ 0 & 8K_1r_3 & 0 & 16(K_3+K_1r_3^2) \end{pmatrix}
\end{aligned}$$

whence we deduce

$$\breve{P}_0 = P_0 \qquad \breve{P}_1 = P_1 \qquad \breve{P}_2 = P_2$$

$$\breve{P}_3 = P_3 - \lambda\begin{pmatrix} 3K_0r_2 & 0 \end{pmatrix}\left[\begin{pmatrix} K_0 & 0 \\ 0 & 4K_1 \end{pmatrix} + \begin{pmatrix} H_1 & 0 \\ 0 & H_2 \end{pmatrix}\right]^{-1}\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = P_3 - \lambda\beta_1 P_1$$

$$\breve{P}_4 = P_4 - \lambda\begin{pmatrix} 0 & 8K_1r_3 & 0 \end{pmatrix}\left[\begin{pmatrix} K_0 & 0 & 3K_0r_2 \\ 0 & 4K_1 & 0 \\ 3K_0r_2 & 0 & 9(K_2+k_0r_2^2) \end{pmatrix} + \begin{pmatrix} H_1 & 0 & 0 \\ 0 & H_2 & 0 \\ 0 & & H_3 \end{pmatrix}\right]^{-1}\begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} = P_4 - \lambda\beta_2 P_2$$

where the β's are given in terms of K, H, r. A thorough treatment of this approach is beyond the scope of this work and will be studied elsewhere.

4.3. **Discrete Sobolev bilinear forms.** The above definition of the Sobolev bilinear function has been proposed in full generality, i.e. without any reference to the explicit expressions of the entries $\mathrm{d}\mu_{i,j}$ in $\mathscr{W}$. A particularly interesting case is obtained when the entries are allowed to be Dirac's δ distributions.

In this perspective, we shall call the part of the Sobolev bilinear function involving a continuous support the *continuous part* of the bilinear function, and that involving a discrete support its *discrete part*. Thus, once we split a Sobolev bilinear function into its continuous and discrete parts, we can consider the former as an additive perturbation of the latter. According to this philosophy, given a set of nodes and their multiplicities $\{x_i, n_i, m_i\}_{i=1}^{s}$ let us study the following Sobolev bilinear function

$$(f,h)_{\breve{\mathscr{W}}} := (f,h)_{\mathscr{W}} + \sum_{i=1}^{s}\sum_{k=0}^{n_i-1}\sum_{j=0}^{m_i-1} \xi^{(i)}_{k,j} h^{(k)}(x_i) f^{(j)}(x_i) \qquad \Longrightarrow \qquad \breve{G} = G + g$$

Notice that the function space on which this Sobolev bilinear form is defined will be $\mathscr{A}^{\breve{N}}_{\breve{\mathscr{W}}}(\breve{\Omega}) \subseteq \mathscr{A}^{N}_{\mathscr{W}}(\Omega)$ where $\breve{\Omega} = \Omega \bigcup_i x_i$ and $\breve{N} = max\big\{N, \{(n_i-1)\}_i, \{(m_i-1)\}_i\big\}$. In order to see how the matrix A looks like in this case, we propose the following

Definition 14. *Given a function $f \in \mathscr{A}^{\breve{N}}_{\breve{\mathscr{W}}}(\breve{\Omega})$, we introduce the vectors*

$$N[f(x)] := \Big(f(x_1), f'(x_1), \dots, f^{(n_1-1)}, f(x_2), f'(x_2), \dots, f^{(n_2-1)}, \dots, f(x_s), f'(x_s), \dots, f^{(n_s-1)}\Big)$$
$$M[f(x)] := \Big(f(x_1), f'(x_1), \dots, f^{(m_1-1)}, f(x_2), f'(x_2), \dots, f^{(m_2-1)}, \dots, f(x_s), f'(x_s), \dots, f^{(m_s-1)}\Big)$$

and the following matrix $\Xi \in \sum_i n_i \times \sum_i m_i$

$$\Xi := \begin{pmatrix} \xi^{(1)} & & & \\ & \xi^{(2)} & & \\ & & \ddots & \\ & & & \xi^{(2)} \end{pmatrix} \qquad\qquad \xi^{(i)} := \begin{pmatrix} \xi^{(i)}_{0,0} & \xi^{(i)}_{0,1} & \cdots & \xi^{(i)}_{0,m_i-1} \\ \xi^{(i)}_{1,0} & & & \\ \vdots & & & \\ \xi^{(i)}_{n_i-1} & & & \xi^{(i)}_{n_i-1,m_i-1} \end{pmatrix}$$

Proposition 9. *Given an additive perturbation of a discrete Sobolev type form, the matrix A can be written in terms of the old polynomials as*

$$A^{[k]} = N[P_1^{[k]}]\,(\Xi)\,M[P_2^{[k]}]^{\top} .$$

Proof. The proposition follows easily from the relations

$$g = N[\chi]\,(\Xi)\,M[\chi]^{\top} \qquad A^{[k]} = S_1^{[k]} g^{[k]} \left(S_2^{[k]}\right)^{\top} \qquad S_1^{[k]} N[\chi] = N[P_1^{[k]}] \qquad S_2^{[k]} M[\chi] = N[P_2^{[k]}]$$

□

It is useful to define the following $\sum_i n_i \times \sum_i m_i$ matrix, suitable for the discrete Sobolev problem at hand, whose entries are the derivatives of the CD Kernel evaluated at the points $\{x_i\}$ up to $\{(n_i-1), (m_i-1)\}$ times.

Definition 15. *We introduce the CD matrix*

$$\mathbb{K}^{[k]} := \left(M[P_2^{[k]}]\right)^\top \left(H^{[k]}\right)^{-1}\left(N[P_1^{[k]}]\right) = \begin{pmatrix} \mathbb{K}^{[k]}_{[1][1]} & \mathbb{K}^{[k]}_{[1][2]} & \dots & \mathbb{K}^{[k]}_{[1][s]} \\ \mathbb{K}^{[k]}_{[2][1]} & \mathbb{K}^{[k]}_{[2][2]} & \dots & \mathbb{K}^{[k]}_{[2][s]} \\ \\ \mathbb{K}^{[k]}_{[s][1]} & \mathbb{K}^{[k]}_{[s][2]} & \dots & \mathbb{K}^{[k]}_{[s][s]} \end{pmatrix}$$

where

$$\mathbb{K}^{[k]}_{[i][j]} := \begin{pmatrix} \left(K^{[k]}(x_i,x_j)\right)^{(0,0)} & \left(K^{[k]}(x_i,x_j)\right)^{(0,1)} & \dots & \left(K^{[k]}(x_i,x_j)\right)^{(0,n_j-1)} \\ \left(K^{[k]}(x_i,x_j)\right)^{(1,0)} & \left(K^{[k]}(x_i,x_j)\right)^{(1,1)} & \dots & \left(K^{[k]}(x_i,x_j)\right)^{(1,n_j-1)} \\ \\ \left(K^{[k]}(x_i,x_j)\right)^{(m_i-1,0)} & \left(K^{[k]}(x_i,x_j)\right)^{(m_i-1,1)} & \dots & \left(K^{[k]}(x_i,x_j)\right)^{(m_i-1,n_j-1)} \end{pmatrix}.$$

Here we have used the notation $\left(K^{[k]}(x_i,x_j)\right)^{(t,d)} := \frac{\partial^{t+d}K^{[k]}(x,y)}{\partial x^t \partial y^d}|_{(x,y)=(x_i,x_j)}$

The previous definitions and analysis allow us to state the main result of this section.

Proposition 10. *The discrete part of a Sobolev bilinear function is as an additive perturbation of its continuous counterpart. Also, the SBPS associated with the Discrete+Continuous part can be represented in terms of the following quasi-determinantal formulas involving only the continuous part of the Sobolev bilinear function.*

$$\check{P}_{1,k}(x) = \left(\begin{array}{c|c} \mathbb{I}+\mathbb{K}^{[k]}\Xi & M[K^{[k]}(\cdot,x)]^\top \\ \hline N[P_{1,k}]\Xi & P_{1,k}(x)\end{array}\right), \qquad \check{P}_{2,k}(x) = \left(\begin{array}{c|c} \mathbb{I}+\Xi\mathbb{K}^{[k]} & \Xi M[P_{2,k}]^\top \\ \hline N[K^{[k]}(x,\cdot)] & P_{2,k}(x)\end{array}\right). \tag{14}$$

Here the expression $M[K^{[k]}(\cdot,x)]$ $(N[K^{[k]}(x,\cdot)])$ *stands for the action of the operator* M *(respectively* N*), on the first (second) variable of* K*. Alternatively the previous formulas can be rewritten in terms of the original polynomials as follows*

$$\check{P}_{1,k}(x) = \left(\begin{array}{c|c} -N[P_{1,k}]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}\left(M\left[(P_2^{[k])^\top}\right]\right)^\top\left(H^{[k]}\right)^{-1} & 1\end{array}\right)\left(\begin{array}{c} P_1^{[k]}(x) \\ \hline P_{1,k}(x)\end{array}\right), \tag{15}$$

$$\check{P}_{2,k}(x) = \left(\begin{array}{c|c} \left(P_2^{[k]}(x)\right)^\top & P_{2,k}(x)\end{array}\right)\left(\begin{array}{c} -\left(H^{[k]}\right)^{-1}N[P_1^{[k]}]\left(\mathbb{I}+\Xi\mathbb{K}^{[k]}\right)^{-1}\Xi M[P_{2,k}]^\top \\ \hline 1\end{array}\right). \tag{16}$$

Proof. Let us write the expression of the inverse of the matrix $(H+A)^{[k]}$. By using Definition 15, one can check the equalities

$$\begin{aligned}
&\left[(H+A)^{[k]}\right]^{-1} = (H^{[k]})^{-1}\left[\left(\mathbb{I}+AH^{-1}\right)^{[k]}\right]^{-1} = (H^{[k]})^{-1}\left(\mathbb{I}+N[P_1^{[k]}]\Xi M[P_2^{[k]}]^\top (H^{[k]})^{-1}\right)^{-1} \\
&= (H^{[k]})^{-1}\left(\mathbb{I}-N[P_1^{[k]}]\Xi M[P_2^{[k]}]^\top(H^{[k]})^{-1} + N[P_1^{[k]}]\Xi M[P_2^{[k]}]^\top(H^{[k]})^{-1}N[P_1^{[k]}]\Xi M[P_2^{[k]}]^\top(H^{[k]})^{-1} - \dots\right) \\
&= (H^{[k]})^{-1} - (H^{[k]})^{-1}N[P_1^{[k]}]\Xi\left(\mathbb{I}-\mathbb{K}^{[k]}\Xi+(\mathbb{K}^{[k]}\Xi)^2-\dots\right)M[P_2^{[k]}]^\top(H^{[k]})^{-1}.
\end{aligned}$$

Consequently, we get the following expression, assuming that the formal series converges

$$\left[(H+A)^{[k]}\right]^{-1} = (H^{[k]})^{-1} - (H^{[k]})^{-1}N[P_1^{[k]}]\Xi\left(\mathbb{I}+\mathbb{K}^{[k]}\Xi\right)^{-1}M[P_2^{[k]}]^\top(H^{[k]})^{-1}. \tag{17}$$

To prove the second statement, observe that

$$\begin{pmatrix} A_{k,0} & A_{k,1} & \dots & A_{k,k-1} \end{pmatrix} = N[P_{1,k}]\Xi M[P_2^{[k]}]^\top \qquad \begin{pmatrix} A_{0,k} \\ A_{1,k} \\ \vdots \\ A_{k-1,k} \end{pmatrix} = N[P_1^{[k]}]\Xi M[P_{2,k}]^\top$$

Once we substitute these expressions in the quasi-determinantal formulae given in proposition 7, we obtain the relations (14). The expressions (15) and (16) follow from those in (14) by just expanding the quasi-determinants and the CD kernels. □

Remark 3. *Whenever the convergence of the series* (17) *is not fulfilled, no orthogonal polynomial sequences arises. This implies that the LU-factorization assumption for the moment matrix was not satisfied in the specific example considered*

Let us define the following polynomial, which will be useful in dealing with the additive discrete part of a bilinear Sobolev function.

Definition 16. *We define the auxiliary polynomial*

$$W(x) := \prod_{i=1}^{s}(x - x_i)^{max\{n_i, m_i\}} \tag{18}$$

The auxiliary polynomial (18) is the keystone for the following result in concordance with [10] and slightly generalizing [21].

Proposition 11. *Given a non-Sobolev inner product $\langle *, * \rangle$, consider the bilinear form*

$$(f, h)_{\mathscr{W}} := \langle f, h\rangle + \sum_{i=1}^{s}\sum_{k=0}^{n_i-1}\sum_{j=0}^{m_i-1} \xi_{k,j}^{(i)} h^{(k)}(x_i) f^{(j)}(x_i)$$

obtained by adding a discrete Sobolev part to the original standard inner product. Then, the SBPS associated with the new bilinear function $(*, *)_{\mathscr{W}}$ satisfies a $\big(2\,[\deg W(x)]+1\big)$-term recurrence relation, which in matrix form reads

$$R_\alpha \check{P}_\alpha(x) = W(x)\check{P}_\alpha \qquad \alpha = 1, 2\ .$$

Here R_α are $\big(2\,[\deg W(x)] + 1\big)$ banded matrices, related to each other, $R_1 = \check{H} R_2^\top \check{H}^{-1}$, and can be written as

$$R_\alpha = M_\alpha W(J) M_\alpha^{-1}\ .$$

This expression involves the connection matrices $M_\alpha P = \check{P}_\alpha$, whose rows, according to (15), (16) *read*

$$\begin{pmatrix} (M_1)_{k,0} & (M_1)_{k,1} & \dots & (M_1)_{k,k-1} & \big| & (M_1)_{k,k} \end{pmatrix} = \begin{pmatrix} -N[P_k]\Xi\left(\mathbb{I} + \mathbb{K}^{[k]}\Xi\right)^{-1}\left(M\left[(P^{[k]})^\top\right]\right)^\top \left(H^{[k]}\right)^{-1} & \Big| & 1 \end{pmatrix}$$

$$\begin{pmatrix} (M_2)_{k,0} & (M_2)_{k,1} & \dots & (M_2)_{k,k-1} & \big| & (M_2)_{k,k} \end{pmatrix} = \begin{pmatrix} -\left(H^{[k]}\right)^{-1} N[P^{[k]}]\left(\mathbb{I} + \Xi\mathbb{K}^{[k]}\right)^{-1}\Xi M[P_2]^\top \\ \hline 1 \end{pmatrix}^\top$$

and the Jacobi matrix $J := S\Lambda S^{-1}$ of the non perturbed initial inner product $\langle *, * \rangle$ (responsible for their three term recurrence relation $JP(x) = xP(x)$).

Proof. It is straightforward to see that

$$(Wf, h)_{\mathscr{W}} = \langle Wf, h\rangle = \langle f, Wh\rangle = (f, Wh)_{\mathscr{W}}\ .$$

Thus, the moment matrix satisfies

$$W(\Lambda)\breve{G} = \breve{G}W(\Lambda^\top) \ .$$

Taking into account the LU factorization of $\breve{G}$ and the definitions for the connection matrices the proposition follows. □

5. Equivalence classes of measure matrices

A natural question arising from the theory previously developed is the following. Consider two measure matrices $\mathscr{W}_1(\Omega) \neq \mathscr{W}_2(\Omega)$ over the same Ω. Assume that the equality $G_{\mathscr{W}_1} = G_{\mathscr{W}_2}$ holds, or equivalently $(p, q; \mathscr{W}_1) = (p, q; \mathscr{W}_2)$ $\forall p, q \in \mathbb{R}[x]$. Notice that, despite sharing the same moment matrix, and hence the same SBPS, in principle $\mathscr{A}^{\mathcal{N}_1}_{\mathscr{W}_1}(\Omega) \neq \mathscr{A}^{\mathcal{N}_2}_{\mathscr{W}_2}(\Omega)$. At the same time, $\mathbb{R}[x] \in \mathscr{A}^{\mathcal{N}_1}_{\mathscr{W}_1}(\Omega) \cap \mathscr{A}^{\mathcal{N}_2}_{\mathscr{W}_2}(\Omega)$ and for every f, g in this intersection, the equality $(f, g; \mathscr{W}_1) = (f, g; \mathscr{W}_2)$ will hold. These considerations suggest to introduce the notion of equivalence class of measure.

Definition 17. *We shall say that two measure matrices $\mathscr{W}_a$ and $\mathscr{W}_b$ are equivalent, and we write $\mathscr{W}_a \sim \mathscr{W}_b$, if $(p, q; \mathscr{W}_a) = (p, q; \mathscr{W}_b)$ for every $p, q \in \mathbb{R}[x]$. We shall denote by $[\mathscr{W}_a] = \{\mathscr{W}_b \setminus \mathscr{W}_b \sim \mathscr{W}_a\}$ the equivalence class of measure matrices equivalent to a given matrix $\mathscr{W}_a$. Two matrices belonging to the same equivalence class will be said similar.*

In other words, equivalent measure matrices share the same moment matrix. We will use the symbol $G_{[\mathscr{W}_a]}$ to denote the common moment matrix of a given equivalent class.

In this section we will address the equivalence problem, by showing how elements of the same matrix class are related. To this aim, we have to study preliminarily how a measure matrix changes under integration by parts manipulations. Let us focus on the (i, j)-th entry of a given measure matrix and take it to be an absolutely continuous measure, this is, $\mathrm{d}\mu_{i,j}(x) = \omega_{i,j}(x)\mathrm{d}x$. We adopt the notation $I\omega_{i,j} := \mu_{i,j}$ for the anti-derivative or primitive of the absolutely continuous measure $\mathrm{d}\mu_{i,j}$.

Two possibilities arise.

- If $\omega_{i,j} \in \mathcal{C}^1(\Omega_{i,j})$,

$$\int_{\Omega_{i,j}} \chi^{(i)}\omega_{i,j}\left(\chi^{(j)}\right)^\top \mathrm{d}x = \begin{cases} \int_{\Omega_{i,j}} \chi^{(i-1)}\left[\delta\omega_{i,j} - \frac{\mathrm{d}\omega_{i,j}}{\mathrm{d}x}\right]\left(\chi^{(j)}\right)^\top \mathrm{d}x - \int_{\Omega_{i,j}} \chi^{(i-1)}\omega_{i,j}\left(\chi^{(j+1)}\right)^\top \mathrm{d}x \\ \int_{\Omega_{i,j}} \chi^{(i)}\left[\delta\omega_{i,j} - \frac{\mathrm{d}\omega_{i,j}}{\mathrm{d}x}\right]\left(\chi^{(j-1)}\right)^\top \mathrm{d}x - \int_{\Omega_{i,j}} \chi^{(i+1)}\omega_{i,j}\left(\chi^{(j-1)}\right)^\top \mathrm{d}x \end{cases}$$

- For the primitive $\mu_{i,j}$,

$$\int_{\Omega_{i,j}} \chi^{(i)}\mathrm{d}\mu_{i,j}\left(\chi^{(j)}\right)^\top = \int_{\Omega_{i,j}} \chi^{(i)}\delta I\omega_{i,j}\left(\chi^{(j)}\right)^\top \mathrm{d}x - \int_{\Omega_{i,j}} \chi^{(i+1)} I\omega_{i,j}\left(\chi^{(j)}\right)^\top \mathrm{d}x - \int_{\Omega_{i,j}} \chi^{(i)} I\omega_{i,j}\left(\chi^{(j+1)}\right)^\top \mathrm{d}x$$

where we have introduced the operator "δ" that turns the continuous measure into a discrete one on the boundary of its support

$$\int_{\Omega_{i,j}} \delta\omega_{i,j}(x)f(x)\mathrm{d}x := \left(\omega_{i,j}(x)f(x)\right)\Big|_{\partial\Omega_{i,j}}$$

Therefore, we have found the relations among similar measure matrices that arise throughout integrations by parts manipulations.

Proposition 12. *The following elementary transformations characterize an equivalent class of measure matrices:*

(19)
$$\begin{pmatrix} \mathrm{d}\mu_{i-1,j-1} & \mathrm{d}\mu_{i-1,j} & \mathrm{d}\mu_{i-1,j+1} \\ \mathrm{d}\mu_{i,j-1} & \boldsymbol{\omega}_{i,j}\mathrm{d}x & \mathrm{d}\mu_{i,j+1} \\ \mathrm{d}\mu_{i+1,j-1} & \mathrm{d}\mu_{i+1,j} & \mathrm{d}\mu_{i+1,j+1} \end{pmatrix} \sim \begin{cases} \begin{pmatrix} \mathrm{d}\mu_{i-1,j-1} & \left(\mathrm{d}\mu_{i-1,j} - \left[\frac{\mathrm{d}\boldsymbol{\omega}_{i,j}}{\mathrm{d}x}\right]\mathrm{d}x\right) & (\mathrm{d}\mu_{i-1,j+1} - \boldsymbol{\omega}_{i,j}\mathrm{d}x) \\ \mathrm{d}\mu_{i,j-1} & 0 & \mathrm{d}\mu_{i,j+1} \\ \mathrm{d}\mu_{i+1,j-1} & \mathrm{d}\mu_{i+1,j} & \mathrm{d}\mu_{i+1,j+1} \end{pmatrix} + \begin{pmatrix} 0 & \delta\boldsymbol{\omega}_{i,j}\mathrm{d}x & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} \mathrm{d}\mu_{i-1,j-1} & \mathrm{d}\mu_{i-1,j} & \mathrm{d}\mu_{i-1,j+1} \\ \left(\mathrm{d}\mu_{i,j-1} - \left[\frac{\mathrm{d}\boldsymbol{\omega}_{i,j}}{\mathrm{d}x}\right]\mathrm{d}x\right) & 0 & \mathrm{d}\mu_{i,j+1} \\ (\mathrm{d}\mu_{i+1,j-1} - \boldsymbol{\omega}_{i,j}\mathrm{d}x) & \mathrm{d}\mu_{i+1,j} & \mathrm{d}\mu_{i+1,j+1} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ \delta\boldsymbol{\omega}_{i,j}\mathrm{d}x & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \begin{pmatrix} \mathrm{d}\mu_{i-1,j-1} & \mathrm{d}\mu_{i-1,j} & \mathrm{d}\mu_{i-1,j+1} \\ \mathrm{d}\mu_{i,j-1} & 0 & (\mathrm{d}\mu_{i,j+1} - I\boldsymbol{\omega}_{i,j}\mathrm{d}x) \\ \mathrm{d}\mu_{i+1,j-1} & (\mathrm{d}\mu_{i+1,j} - I\boldsymbol{\omega}_{i,j}\mathrm{d}x) & \mathrm{d}\mu_{i+1,j+1} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 \\ 0 & \delta I\boldsymbol{\omega}_{i,j}\mathrm{d}x & 0 \\ 0 & 0 & 0 \end{pmatrix} \end{cases}$$

Iterations of these transformations are obviously allowed. Notice the split between the continuous and discrete parts. The previous transformations can be performed over every entry i, j in the measure matrix as long as $\omega_{i,j}$ can be derived or integrated. This leads to a huge amount of equivalent matrices in $[\mathscr{W}_a]$.

$$\begin{pmatrix} * & * & * & * & * & * & * & * \\ * & * & * & * & * & * & * & \\ * & * & * & * & * & * & & \\ * & * & * & * & * & & & \\ * & * & * & \bigstar & * & * & * & * \\ * & * & * & * & * & * & * & * \\ * & * & & * & * & * & * & * \\ * & & & * & * & * & * & * \\ & & & * & * & * & * & * \\ & & & * & * & * & * & * \end{pmatrix}$$

Upper (lower) anti-diagonal terms come from taking derivatives (integrals) of the entry in the star location.

Each iteration produces a discrete term. Once gathered together in a matrix, these terms will define an additive discrete perturbation of the measure matrix.

According to the previous discussion, if we can obtain the SBPS associated to the continuous part, the SBPS associated to the whole bilinear form can also be obtained with the aid of the CD kernels of the continuous part. However, under certain conditions imposed on the $\omega_{i,j}$ one can get rid of the discrete part.

Definition 18. *Let us denote by $\tilde{\omega}_k$ any weight with finite moments on Ω having the following property*

$$\delta\tilde{\omega}_k^{(t)} = 0 \qquad t = 0, 1, 2, \ldots, (k-1) \tag{20}$$

According to proposition 8, the classical measure $u_{\gamma+k}$ is a particular example of $\tilde{\omega}_k$.

Proposition 13. *Let $\mathscr{W}$ be a $(\mathcal{N}+1)\times(\mathcal{N}+1)$ measure matrix such that $\mathrm{d}\mu_{i,j} = \omega_{i,j}\mathrm{d}x$ and each $\omega_{i,j}$ is a function of class $\mathcal{C}^{|i-j|}$.*

- *If $\mathscr{W} = \mathscr{W}^\top$ then $\mathscr{W}$ is similar to the sum of a diagonal measure matrix and a discrete measure matrix.*

- *If* $\mathscr{W} = \mathscr{W}^\top$ *and additionally each entry* $\omega_{i,j}$ *can has the property of* $\tilde{\omega}_{|i-j|}$ $\forall i, j$, *as in* (20) *then* $\mathscr{W}$ *is similar to a diagonal measure matrix.*

Proof. We shall formulate an inductive procedure to prove the first statement of the proposition. Given any $(\mathcal{N}+1)\times(\mathcal{N}+1)$ symmetric measure matrix $\mathscr{W}$

$$\mathscr{W} = \begin{pmatrix} \omega_{0,0} & \omega_{1,0} & \omega_{2,0} & \dots & \omega_{\mathcal{N}-1,0} & \omega_{\mathcal{N},0} \\ \omega_{1,0} & \omega_{1,1} & \omega_{2,1} & \dots & \omega_{\mathcal{N}-1,1} & \omega_{\mathcal{N},1} \\ \omega_{2,0} & \omega_{2,1} & \omega_{2,2} & \dots & & \\ & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & & \\ \omega_{\mathcal{N}-1,0} & \omega_{\mathcal{N}-1,1} & \dots & & \omega_{\mathcal{N}-1,\mathcal{N}-1} & \omega_{\mathcal{N},\mathcal{N}-1} \\ \omega_{\mathcal{N},0} & \omega_{\mathcal{N},1} & \dots & & \omega_{\mathcal{N},\mathcal{N}-1} & \omega_{\mathcal{N},\mathcal{N}} \end{pmatrix} \mathrm{d}x$$

one can use the first similarity relation stated in (19) for each entry of the last row of $\mathscr{W}$ and the second one for each entry of its last column. In this way, one obtains

$$\mathscr{W} \sim \begin{pmatrix} \omega_{0,0} & \omega_{1,0} & \omega_{2,0} & \dots & \omega_{\mathcal{N}-1,0} - \omega'_{\mathcal{N},0} & 0 \\ \omega_{1,0} & \omega_{1,1} & \omega_{2,1} & \dots & \omega_{\mathcal{N}-1,1} - \omega_{\mathcal{N},0} - \omega'_{\mathcal{N},1} & 0 \\ (\omega_2)_{2,0} & (\omega_1)_{2,1} & \omega_{2,2} & \dots & & \\ & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & & \\ \omega_{\mathcal{N}-1,0} - \omega'_{\mathcal{N},0} & \omega_{\mathcal{N}-1,1} - \omega_{\mathcal{N},0} - \omega'_{\mathcal{N},1} & \dots & & \omega_{\mathcal{N}-1,\mathcal{N}-1} - 2\omega_{\mathcal{N},\mathcal{N}-2} - \omega'_{\mathcal{N},\mathcal{N}-1} & 0 \\ 0 & 0 & \dots & & 0 & \omega_{\mathcal{N},\mathcal{N}} \end{pmatrix} \mathrm{d}x$$

$$+ \begin{pmatrix} 0 & 0 & 0 & \dots & \delta\omega_{\mathcal{N},0} & 0 \\ 0 & 0 & 0 & \dots & \delta\omega_{\mathcal{N},1} & 0 \\ 0 & 0 & 0 & \dots & & \\ & & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & & \\ \delta\omega_{\mathcal{N},0} & \delta\omega_{\mathcal{N},1} & \dots & & \delta\omega_{\mathcal{N},\mathcal{N}-1} & 0 \\ 0 & 0 & \dots & & 0 & 0 \end{pmatrix} \mathrm{d}x$$

This new, equivalent measure matrix is still symmetric. Therefore, the whole procedure can be repeated up to $\mathcal{N}$ times, until the diagonal form is achieved, jointly with the discrete terms that will appear each time.
The second statement of the proposition is just a corollary of the first one since the definition 18 is suited to make the discrete terms disappear. □

Let us consider the example $\mathcal{N} = 3$.

$$\begin{pmatrix}\omega_{0,0} & \omega_{1,0} & \omega_{2,0} & \omega_{3,0}\\ \omega_{1,0} & \omega_{1,1} & \omega_{2,1} & \omega_{3,1}\\ \omega_{2,0} & \omega_{2,1} & \omega_{2,2} & \omega_{3,2}\\ \omega_{3,0} & \omega_{3,1} & \omega_{3,2} & \omega_{3,3}\end{pmatrix}\sim$$

$$\begin{pmatrix}\omega_{0,0}-\omega^{'}_{1,0}+\omega^{''}_{2,0}+\omega^{'''}_{3,0} & 0 & 0 & 0\\ 0 & \omega_{1,1}-\omega^{'}_{2,1}+\omega^{''}_{3,1}-2\omega_{2,0}+3\omega^{'}_{3,0} & 0 & 0\\ 0 & 0 & \omega_{2,2}-\omega^{'}_{2,3}-2\omega_{3,1} & 0\\ 0 & 0 & 0 & \omega_{3,3}\end{pmatrix}$$

$$+\begin{pmatrix}\delta[\omega_{1,0}-\omega'_{2,0}-\omega''_{3,0}] & \delta[\omega_{2,0}-\omega'_{3,0}] & \delta\omega_{3,0} & 0\\ \delta[\omega_{2,0}-\omega'_{3,0}] & \delta[\omega_{2,1}-\omega_{3,0}-\omega'_{3,1}] & \delta\omega_{3,1} & 0\\ \delta\omega_{3,0} & \delta\omega_{3,1} & \delta\omega_{3,2} & 0\\ 0 & 0 & 0 & 0\end{pmatrix}$$

5.1. **Sobolev inner products involving classical measures.** When dealing with classical measures u_γ, the construction of equivalence classes of measure matrices appears to be particularly simple and neat. The reason resides in the possibility of generating equivalence classes without having to deal with any discrete parts (boundary terms). We summarize this properties in the next

Proposition 14. *Let*

$$\mathscr{W}=\begin{pmatrix}\omega_1^0 & 0 & \dots & \\ 0 & 0 & & \\ \vdots & & \ddots & \\ & & & \end{pmatrix}u_{\gamma+0}\mathrm{d}x+\begin{pmatrix}0 & \omega_2^1 & 0 & \dots\\ \omega_2^1 & \omega_1^1 & & \\ 0 & & 0 & \\ \vdots & & & \ddots\end{pmatrix}u_{\gamma+1}\mathrm{d}x+\dots+\begin{pmatrix}0 & \dots & 0 & \omega_{n+1}^n\\ \vdots & \ddots & & \omega_n^n\\ & & & \vdots\\ 0 & & 0 & \vdots\\ \omega_{n+1}^n & \omega_n^n & \dots & \omega_1^n\end{pmatrix}u_{\gamma+n}\mathrm{d}x$$

be a measure matrix such that each $\{\omega_j^r u_{\gamma+r}\}_{j=1}^{r+1}$ is of type $\tilde{\omega}_r$ $\forall r=0,1,\dots,n$. Then if $\mathscr{W}$ determines a SOPS, then there exist linear differential operator $\mathbf{F}$ and constants $\{\alpha_{k,j},\beta_{k,j}\}$ such that

$$(f,h;\mathscr{W})=\langle\mathbf{F}[f]h,u_\gamma\rangle=\langle f\mathbf{F}[h],u_\gamma\rangle \qquad\Longrightarrow\qquad \mathbf{F}[P_{\mathscr{W},k}]=\sum_{j=k}^{r}\alpha_{k,j}P_{\gamma,j}$$

$$(\mathbf{F}[f],h;\mathscr{W})=(f,\mathbf{F}[h];\mathscr{W}) \qquad\Longrightarrow\qquad \mathbf{F}[P_{\mathscr{W},k}]=\sum_{j=k-r}^{k+r}\beta_{k,j}P_{\mathscr{W},j}$$

Proof. Since the selected measure matrix $\mathscr{W}$ satisfies the conditions in proposition 13, using also proposition 8 it is not hard to see that

$$\mathscr{W}\sim\begin{pmatrix}v_0u_\gamma & & & \\ & v_1u_{\gamma+1} & & \\ & & \ddots & \\ & & & v_nu_{\gamma+n}\end{pmatrix} \qquad\Longrightarrow\qquad (f,h;\mathscr{W})=\sum_{r=0}^{n}\langle f^{(r)}h^{(r)},v_ru_{\gamma+r}\rangle$$

where the $\{v_r\}_{r=0}^n$ are functions that depend on the ω and their derivatives and $v_ru_{\gamma+r}$ are of type $\tilde{\omega}_r$ due to the conditions that the proposition imposes on the $\omega_j^ru_{\gamma+r}$. Using proposition 8 (in which the operator $\mathscr{O}_r^j$ was defined) for the r-th term of the sum, the following chain of equalities follow

$$\langle f^{(r)}h^{(r)},v_ru_{\gamma+r}\rangle=(-1)^r\sum_{j=0}^{r}\binom{r}{j}\langle h^{(0)}(f^{(r)})^{(j)},(v_ru_{\gamma+r})^{(r-j)}\rangle=(-1)^r\sum_{j=0}^{r}\binom{r}{j}\langle h^{(0)}\left(\mathscr{O}_r^{j+1}[v_r]p_2^j\right)f^{(r+j)},u_\gamma\rangle$$

this has to be added for each r, after doing so the differential operator $\mathbf{F}$ can be finally defined as

$$\mathbf{F} := \sum_{r=0}^{n}(-1)^r \sum_{j=0}^{r}\binom{r}{j}\left(\mathcal{O}_r^{j+1}[v_r](x)p_2^j(x)\right)\frac{\mathrm{d}^{r+j}}{\mathrm{d}x^{r+j}} .$$

Consequently, $(f,h;\mathscr{W}) = \langle h\mathbf{F}[f], u_\gamma\rangle$. By acting on h instead of f, we also get $(f,h;\mathscr{W}) = \langle \mathbf{F}[h]f, u_\gamma\rangle$. As we already know, this equalities can be translated to relations between the moment matrices

$$(f,h;\mathscr{W}) = \langle \mathbf{F}[f]h, u_\gamma\rangle = \langle f\mathbf{F}[h], u_\gamma\rangle \qquad \Longrightarrow \qquad G_{\mathscr{W}} = Fg_\gamma = g_\gamma F^\top$$

where the matricial representation of $\mathbf{F}$ is, $F := \sum_{r=0}^{n}(-1)^r\sum_{j=0}^{r}\binom{r}{j} D^{r+j}\left(\mathcal{O}_r^{j+1}[v_r](\Lambda)p_2^j(\Lambda)\right)$. Now if $\mathscr{W}$ gives a SOPS then $G_{\mathscr{W}}$ must be LU factorizable, i.e.,

$$G_{\mathscr{W}} = Fg_\gamma \qquad \Longrightarrow \qquad U := S_{\mathscr{W}}FS_\gamma^{-1} = H_{\mathscr{W}}\left(S_\gamma S_{\mathscr{W}}^{-1}\right)^\top H_\gamma^{-1} \qquad \Longrightarrow \qquad UP_\gamma = \mathbf{F}[P_{\mathscr{W}}]$$

The second set of equations imposes an upper triangular form to U, with a finite number r of non vanishing super-diagonal terms only, that will depend on the differential operator.
Multiplying the relation between moment matrices by F and $F^\top$ and LU factorizing once more one obtains

$$FG_{\mathscr{W}} = G_{\mathscr{W}}F^\top \qquad \Longrightarrow \qquad J_F := S_{\mathscr{W}}FS_{\mathscr{W}}^{-1} = H_{\mathscr{W}}J_F^\top H_{\mathscr{W}}^{-1} \qquad \Longrightarrow \qquad J_FP_{\mathscr{W}} = \mathbf{F}[P_{\mathscr{W}}]$$

This time, the second set of relations imposes a $2r+1$ diagonal structure to J_F ($2r$ non vanishing diagonals above and below the main one). □

Some comments are in order.

- The initial condition $\{\omega_j^r u_{\gamma+r}\}_{j=1}^{r+1}$ being of type $\tilde{\omega}_r$ is not so restrictive since $u_{\gamma+r}$ already is of type $\tilde{\omega}_r$. So the ω_j^r just must not spoil this property.
- A particularly simple example is to consider $\omega_j^r = 0\ \forall j \neq 1$ in which case $\omega_1^r = v_r$. Taking now $v_r = p_2^{n-r}\lambda_r$ with $\lambda_r > 0\ \forall r = 0,1,\dots,n$ one is left with the following inner product and corresponding linear differential operator

$$(f,h;\mathscr{W}) = \sum_{r=0}^{n}\lambda_r\langle f^{(r)}h^{(r)}, u_{\gamma+n}\rangle = \langle \mathbf{F}[f]h, u_\gamma\rangle \qquad \mathbf{F} = \sum_{r=0}^{n}(-1)^r\lambda_r\sum_{j=0}^{r}\binom{r}{j}\varphi_{n,r-j}(x)\frac{\mathrm{d}^{r+j}}{\mathrm{d}x^{r+j}}$$

  Although with small differences (the starting inner product's measure and the way the Pearson equation is used), this example is in agreement with the main ideas in [20] and [19].

6. Polynomial deformations of the measure matrix

As we have seen, moment matrices arising from a diagonal $\mathscr{W}$ with positive definite measures (also symmetric $\mathscr{W}$ reducible to a diagonal shape) are examples of Sobolev LU-factorizable moment matrices. In this section, we investigate deformations of a given factorizable case, with the idea of exploring the possibility of new factorizable ones. The deformations of the measure matrix we are interested in can be understood as deformations of the moment matrix, which naturally translate into transformations of the associated bilinear form. These transformations of the bilinear form are expressed in terms of linear differential operators acting on each of its entries, but before studying the general case, we will start with the more simple and usual case of deformations involving polynomials.

On one hand, in the standard case (corresponding to $\omega_{n,r} = 0\ \forall n, r > 0$, which gives $(f,h;\mathscr{W}) = \langle f,h\rangle_{\omega_{0,0}}$) we have the symmetry $\langle xf,h\rangle_{\omega_{0,0}} = \langle f,xh\rangle_{\omega_{0,0}} = \langle f,h\rangle_{x\omega_{0,0}}$. This symmetry is responsible, for instance, for the three term recurrence relation of the OPS, or the Hankel shape of the moment matrix.

On the other hand, given a measure matrix $\mathscr{W}$, in general $(xf,h;\mathscr{W}) \neq (f,xh;\mathscr{W}) \neq (f,h;x\mathscr{W})$. However, we can equivalently say that there exist new measure matrices $\mathscr{W}_2, \mathscr{W}_3$ such that $(xf,h;\mathscr{W}_1) = (f,h;\mathscr{W}_2)$ and

$(f, xh; \mathscr{W}_1) = (f, h; \mathscr{W}_3)$. While multiplication of any of the entries of the standard inner product by a polynomial produces another standard inner product, instead, the same operation in any of the entries of a Sobolev-type bilinear function deforms the initial $\mathscr{W}$ giving a different one, probably spoiling the symmetries of $\mathscr{W}$ if it had any.

Theorem 1. *The operator $\mathcal{X}$ of multiplication by x,*

$$\mathcal{X} := \begin{pmatrix} x & 1 & 0 & 0 & \dots \\ 0 & x & 2 & 0 & \dots \\ 0 & 0 & x & 3 & \dots \\ 0 & 0 & 0 & x & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

once applied to any of the entries of a Sobolev bilinear function, provides the following deformation of the measure matrix [2]

$$\begin{aligned} (xf, h; \mathscr{W}) &= (f, h; \mathcal{X}\mathscr{W}) & \Lambda G_{\mathscr{W}} &= G_{\mathcal{X}\mathscr{W}} \\ (f, xh; \mathscr{W}) &= (f, h; \mathscr{W}(\mathcal{X})^\top) & G_{\mathscr{W}}\Lambda^\top &= G_{\mathscr{W}(\mathcal{X})^\top} \end{aligned}$$

Proof. Using the definition of the moment matrix and taking into account the commutation relations between D^k and Λ, we get

$$\Lambda G_{\mathscr{W}} = \Lambda \boldsymbol{D}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top = \begin{pmatrix}\Lambda\mathbb{I} & \Lambda D & \Lambda D^2 & \dots & \Lambda D^k & \dots\end{pmatrix}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top =$$

$$\begin{pmatrix}\Lambda & D\Lambda + \mathbb{I} & D^2\Lambda + 2D & \dots & D^k\Lambda + kD^{k-1} & \dots\end{pmatrix}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top =$$

$$\boldsymbol{D}\begin{pmatrix} \Lambda & \mathbb{I} & 0 & 0 & \dots \\ 0 & \Lambda & 2\mathbb{I} & 0 & \dots \\ 0 & 0 & \Lambda & 3\mathbb{I} & \dots \\ 0 & 0 & & \Lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top = \boldsymbol{D}\left(\int_\Omega \chi(x)\begin{pmatrix} x & 1 & 0 & 0 & \dots \\ 0 & x & 2 & 0 & \dots \\ 0 & 0 & x & 3 & \dots \\ 0 & 0 & & x & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top.$$

□

We can generalize the previous argument. First, let us compute the powers of $\mathcal{X}$. Then one can observe that $\mathcal{X}^k$ is an upper triangular banded matrix, whose entries for $n = 1, 2, \dots$ are

$$\begin{aligned} (\mathcal{X}^k)_{(n-1),(n-1)+i} &= \binom{k}{i}(n)^i x^{k-i} & 0 &\leq i \leq k \\ (\mathcal{X}^k)_{(n-1),(n-1)+i} &= 0 & i &> k \end{aligned}$$

In addition, due to the bilinearity of the function, we obtain the following

Proposition 15. *Given two real polynomials $P(x)$ and $Q(x)$, the relations*

$$(P(x)f, Q(x)h; \mathscr{W}) = (f, h; P(\mathcal{X})\mathscr{W}[Q(\mathcal{X})]^\top) \qquad P(\Lambda)G_{\mathscr{W}}\left(Q(\Lambda)^\top\right) = G_{P(\mathcal{X})\mathscr{W}(Q(\mathcal{X})^\top)}$$

[2] Being the initial moment matrix $G_{\mathscr{W}}$ a LU-factorizable moment matrix does not imply the new moment matrix $G_{\mathcal{X}\mathscr{W}}$ to be LU-factorizable as well.

hold. If $deg\{P(x)\}=k$, then $P(\mathcal{X})$ is an upper triangular matrix whose entries are

$$(P(\mathcal{X}))_{(n-1),(n-1)+i}=\begin{cases}\frac{(n)^i}{i!}\frac{\mathrm{d}^ip(x)}{\mathrm{d}x^i} & 0\leq i\leq k\\ 0 & i>k\end{cases}\qquad P(\mathcal{X})=\begin{pmatrix}P(x) & P'(x) & P''(x) & P'''(x) & \dots\\ & P(x) & 2P'(x) & 3P''(x) & \dots\\ & & P(x) & 3P'(x) & \dots\\ & & & P(x) & \dots\\ & & & & \ddots\end{pmatrix}.$$

Thus, if $\mathscr{W}$ is a $(\mathcal{N}+1)\times(\mathcal{N}+1)$ measure matrix, then $P(\mathcal{X})\mathscr{W}[Q(\mathcal{X})]^\top$ will still be a $(\mathcal{N}+1)\times(\mathcal{N}+1)$ measure matrix.

The interest of the latter proposition relies on the fact that, although in principle there is no reason why $G_{P(\mathcal{X})\mathscr{W}(Q(\mathcal{X})^\top)}$ should be LU-factorizable if $G_{\mathscr{W}}$ is so, there will be important cases, that we we are about to study, where equations like the one in the right hand side of the proposition will lead to relations between the SBPS associated to the deformed and non deformed measure matrices. Therefore, this proposition will be keystone in order to study a special case where the standard three term recurrence relation holds and to generalize the concept of Darboux transformations [3] to the Sobolev context.

6.1. **A special case where the standard three term recurrence relation holds.** As we have already pointed out, given an arbitrary measure matrix $\mathscr{W}$, in general $(xf,h;\mathscr{W})\neq(f,xh;\mathscr{W})$. However, if we impose some additional symmetry on $\mathscr{W}$, or we specialize it conveniently, we may get the desired equality.

Definition 19. *We introduce the set of matrices*

$$\mathscr{W}_x:=\{\mathscr{W}\setminus\mathcal{X}\mathscr{W}\sim\mathscr{W}\mathcal{X}^\top\}.$$

Theorem 2. *If $\mathscr{W}\in\mathscr{W}_x$ then $G_{\mathscr{W}}$ is Hankel and the associated SOPS satisfy the **standard** three term recurrence relation*

$$xP_n=J_{n,n-1}P_{n-1}+J_{n,n}P_n+P_{n+1}\qquad J_{n,n-1}=\frac{h_n}{h_{n-1}}\qquad J_{n,n}=S_{n,n-1}-S_{n+1,n}$$

Proof. The condition $\mathcal{X}\mathscr{W}\sim\mathscr{W}\mathcal{X}^\top$, due to Theorem 1 is equivalent to $\Lambda G_{\mathscr{W}}=G_{\mathcal{X}\mathscr{W}}=G_{\mathscr{W}\mathcal{X}^\top}=G_{\mathscr{W}}\Lambda^\top$. This symmetry of the moment matrix leads to its Hankel shape and allows to construct the well known tri-diagonal Jacobi matrix ($J:=S\Lambda S^{-1}$) with its entries in terms of the elements of S,h. Note also that if $\mathscr{W}\in\mathscr{W}_x$ then $\mathcal{X}\mathscr{W}\in\mathscr{W}_x$ as well. □

Theorem 3. *$\mathscr{W}_x$ is not an empty set.*

Proof. We give here the following counterexample

$$\mathscr{W}=\begin{pmatrix}\mathrm{d}\mu_0 & \binom{1}{0}\mathrm{d}\mu_1 & \binom{2}{0}\mathrm{d}\mu_2 & \binom{3}{0}\mathrm{d}\mu_3 & \dots & \binom{\mathcal{N}}{0}\mathrm{d}\mu_{\mathcal{N}}\\ \binom{1}{1}\mathrm{d}\mu_1 & \binom{2}{1}\mathrm{d}\mu_2 & \binom{3}{1}\mathrm{d}\mu_3 & & & 0\\ \binom{2}{2}\mathrm{d}\mu_2 & \binom{3}{2}\mathrm{d}\mu_3 & & & & \\ \binom{3}{3}\mathrm{d}\mu_3 & & & & & \\ \vdots & \binom{\mathcal{N}}{\mathcal{N}-1}\mathrm{d}\mu_{\mathcal{N}} & & & & \\ \binom{\mathcal{N}}{\mathcal{N}}\mathrm{d}\mu_{\mathcal{N}} & 0 & & & & \end{pmatrix}\in\mathscr{W}_x \tag{21}$$

which is obtained by imposing $\mathscr{X}\mathscr{W}=\mathscr{W}\mathscr{X}^\top$ and yiels the following Sobolev inner product

$$(f,h;\mathscr{W})=\sum_{n=0}^{\mathscr{N}}\int_{\Omega_n}(fh)^{(n)}\mathrm{d}\mu_n(x)$$

□

For a SOPS associated with a measure matrix in $\mathscr{W}_x$, all of the results of the standard theory of orthogonal polynomial sequences hold: Three term recurrence relation, Christoffel-Darboux formulae, the existence of τ-functions, of associated integrable hierarchies, etc.. All of these properties are indeed a non-trivial consequence of the symmetry $\Lambda G_{\mathscr{W}}=G_{\mathscr{W}}\Lambda^\top$.

A natural question is the relation between the previous result and the classical Favard theorem. Essentially, Favard's theorem assures that given a set of polynomials, satisfying certain initial conditions and a standard three term recurrence relation, there exists a measure μ with respect to which the set of polynomials is actually an OPS. The SOPS associated to a $\mathscr{W}_x$ indeed satisfy the hypotheses of Favard's theorem. Therefore, from both results we deduce that there must exist a measure $\mathrm{d}\mu$ such that $\mathrm{d}\mu E_{00}\sim\mathscr{W}_x$. (Remember that similar measure matrices shared both the moment matrix and the orthogonal polynomial sequence).

Let us consider a particular case of the given counterexample. Let us take $\mathscr{N}=1$ and $\Omega_n:=[x_1,x_2]$ for $n=0,1$ and, using the iterations of Proposition 19 it is not hard to see that (at least in the function spaces where the corresponding integration by parts makes sense)

$$\begin{pmatrix}\mathrm{d}\mu_0 & \mathrm{d}\mu_1\\ \mathrm{d}\mu_1 & 0\end{pmatrix}\sim\begin{pmatrix}\mathrm{d}\mu_0+(\mathrm{d}\mu_1)' & 0\\ 0 & 0\end{pmatrix}+\begin{pmatrix}\delta\mathrm{d}\mu_1 & 0\\ 0 & 0\end{pmatrix}=[\mathrm{d}\mu_0+(\mathrm{d}\mu_1)'+\delta\mathrm{d}\mu_1]\,E_{00}$$

6.2. **Darboux–Sobolev tranformations and quasi-recurrence relations.** In the next three sections we will proceed, with the aid of proposition 15 to deform the measure matrix by means of a (right or left) multiplication by a polynomial in $\mathscr{X}$ or its inverse. Subsequently, we shall study the relation between the new and old SBPS associated to the deformed and non deformed measure matrices, respectively. The reason for the name of these deformations is that whenever $\mathscr{W}=E_{0,0}\omega$ (the "standard" case), then our deformations reduce to the "standard" Darboux transformations or linear spectral transformations. As already noticed in the introduction, this section adapts and completes, for this particular Sobolev scalar case, the more general results given in Ref. [4].

6.3. **Christoffel–Sobolev transformations.** Let us introduce the polynomial $R(x):=\prod_{i=1}^d(x-r_i)^{m_i}$ of degree $\sum_{i=1}^d m_i=M$.

Definition 20. *The right and left Christoffel–Sobolev deformed measure matrices and moment matrices are*

$$\hat{\mathscr{W}_L}:=R(\mathscr{X})\mathscr{W}\qquad\qquad \hat{\mathscr{W}_R}:=\mathscr{W}[R(\mathscr{X})]^\top$$
$$R(\Lambda)G_{\mathscr{W}}=G_{\hat{\mathscr{W}}_L}:=\hat{G}_L\qquad\qquad G_{\mathscr{W}}[R(\Lambda)]^\top=G_{\hat{\mathscr{W}}_R}:=\hat{G}_R$$

The resolvents and adjoint resolvents are defined as

$$(\hat{\omega}_L):=(\hat{S}_{L1})R(\Lambda)S_1^{-1}\qquad\qquad (\hat{\Omega}_L):=S_2(\hat{S}_{L2})^{-1}$$
$$(\hat{\omega}_R):=(\hat{S}_{R2})R(\Lambda)S_2^{-1}\qquad\qquad (\hat{\Omega}_R):=S_1(\hat{S}_{R1})^{-1}$$

Proposition 16. *The resolvents are related to the adjoint resolvents by the formulae*

$$(\hat{\omega}_L)=(\hat{H}_L)(\hat{\Omega}_L)^\top H^{-1}\qquad\qquad (\hat{\omega}_R)=(\hat{H}_R)(\hat{\Omega}_R)^\top H^{-1}$$

and have the following $(M+1)$ diagonal structure

$$\hat{\omega}=\begin{pmatrix}\hat{\omega}_{0,0} & \hat{\omega}_{0,1} & \dots & \hat{\omega}_{0,(M-1)} & \hat{\omega}_{0,M} & 0 & & & \\ 0 & \hat{\omega}_{1,1} & & & \hat{\omega}_{1,M} & \hat{\omega}_{1,(M+1)} & 0 & \dots & \\ 0 & 0 & \ddots & & & & \ddots & & \\ & & & \hat{\omega}_{k,k} & & & \hat{\omega}_{k,k+M-1} & \hat{\omega}_{k,k+M} & 0 \\ & & & & \ddots & & & & \ddots\end{pmatrix},$$

where $\hat{\omega}_{k,k+M}=1$ *and* $\hat{\omega}_{k,k}=\frac{\hat{h}_k}{h_k}$.

Proof. The previous relations follow from a LU-factorization of the expressions defining the Darboux–Sobolev deformed moment matrices. □

Let us establish now some connection formulae relating deformed to non-deformed polynomials. They are based on the notion of resolvent, as clarified by the following

Proposition 17. *Deformed and non deformed polynomials are related by the resolvents*

$$(\hat{\omega}_L)P_1(x)=R(x)(\hat{P}_{L1})(x) \qquad (\hat{\Omega}_L)(\hat{P}_{L2})(x)=P_2(x)$$
$$(\hat{\omega}_R)P_2(x)=R(x)(\hat{P}_{R2})(x) \qquad (\hat{\Omega}_R)(\hat{P}_{R1})(x)=P_1(x)$$

while transformed and non transformed Christoffel–Darboux kernels are related as follows

$$K^{[n+1]}(x,y)=R(y)\hat{K}_L^{[n+1]}(x,y)-$$
$$\begin{pmatrix}(\hat{P}_{L2})_{n+1-M} & \dots & (\hat{P}_{L2})_n\end{pmatrix}\begin{pmatrix}(\hat{h}_L)^{-1}_{n+1-M} & & \\ & \ddots & \\ & & (\hat{h}_L)^{-1}_n\end{pmatrix}\begin{pmatrix}(\hat{\omega}_L)_{n+1-M,n+1} & & 0\\ \vdots & \ddots & \\ (\hat{\omega}_L)_{n,n+1} & \dots & (\hat{\omega}_L)_{n,n+M}\end{pmatrix}\begin{pmatrix}(P_1)_{n+1}(y)\\ \vdots \\ (P_1)_{n+m}(y)\end{pmatrix}$$

$$K^{[n+1]}(y,x)=R(y)\hat{K}_R^{[n+1]}(y,x)-$$
$$\begin{pmatrix}(\hat{P}_{R1})_{n+1-M} & \dots & (\hat{P}_{R1})_n\end{pmatrix}\begin{pmatrix}(\hat{h}_R)^{-1}_{n+1-M} & & \\ & \ddots & \\ & & (\hat{h}_R)^{-1}_n\end{pmatrix}\begin{pmatrix}(\hat{\omega}_R)_{n+1-M,n+1} & & 0\\ \vdots & \ddots & \\ (\hat{\omega}_R)_{n,n+1} & \dots & (\hat{\omega}_R)_{n,n+M}\end{pmatrix}\begin{pmatrix}(P_2)_{n+1}(y)\\ \vdots \\ (P_2)_{n+m}(y)\end{pmatrix}$$

Proof. The first set of relations follow directly by using the definition of the resolvents, and taking into account their action on the SBPS. The second set of relations follow by making explicit the equalities

$$\left[(\hat{P}_{L2})^{\top}(x)(\hat{\Omega}_L)^{\top}\right]H^{-1}P_1(y)=(\hat{P}_{L2})^{\top}(x)(\hat{H}_L)^{-1}\left[(\hat{\omega}_L)P_1(y)\right]$$

for the first one and

$$\left[(\hat{P}_{R1})^{\top}(x)(\hat{\Omega}_R)^{\top}\right]H^{-1}P_2(y)=(\hat{P}_{R1})^{\top}(x)(\hat{H}_R)^{-1}\left[(\hat{\omega}_L)P_2(y)\right]$$

for the second one. □

Let us introduce a vector of "germs" of a function near the points r_i, having multiplicities m_i.

Definition 21. *Given a function* $f(x)$ *and a set* $r:=\{(r_i,m_i)\}_{i=1}^d$ *of points* $r_i\in\mathbb{R}$ *with associated multiplicities* $m_i\in\mathbb{N}$, *we define the vector of germs* $\Pi_r[f]:\mathcal{F}(x)\longrightarrow\mathbb{R}^{\sum m_i}$ *as*

$$\Pi_r[f]:=\left(\frac{f^{(0)}(r_1)}{0!},\frac{f^{(1)}(r_1)}{1!},\dots,\frac{f^{(m_1-1)}(r_1)}{(m_1-1)!};\frac{f^{(0)}(r_2)}{0!},\frac{f^{(1)}(r_2)}{1!},\dots,\frac{f^{(m_2-1)}(r_2)}{(m_2-1)!};\dots;\frac{f^{(0)}(r_d)}{0!},\dots,\frac{f^{(m_d-1)}(r_d)}{(m_d-1)!}\right).$$

Now we can state an useful result.

Proposition 18. *The Christoffel transformed polynomials and their norms are given in terms of the original ones by means of the relations*

$$(\hat{P}_{1L})_n(x) = \frac{1}{R(x)}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_n\\(P_1)_{n+1}\\\vdots\\(P_1)_{n+M-1}\end{bmatrix} & \begin{matrix}(P_1)_n(x)\\(P_1)_{n+1}(x)\\\vdots\\(P_1)_{n+M-1}(x)\end{matrix}\\\hline \Pi_r[(P_1)_{n+M}] & (P_1)_{n+M}(x)\end{array}\right), \quad \frac{(\hat{P}_{2L})_n(x)}{(\hat{h}_L)_n} = \Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_{n+1}\\\vdots\\(P_1)_{n+M}\end{bmatrix} & \begin{matrix}0\\\vdots\\1\end{matrix}\\\hline \Pi_r[K^{[n+1]}(x,\cdot)] & 0\end{array}\right)$$

$$\frac{(\hat{h}_L)_n}{h_n} =\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_n\\(P_1)_{n+1}\\\vdots\\(P_1)_{n+M-1}\end{bmatrix} & \begin{matrix}1\\0\\\vdots\\0\end{matrix}\\\hline \Pi_r[(P_1)_{n+M}] & 0\end{array}\right),$$

$$(\hat{P}_{2R})_n(x) = \frac{1}{R(x)}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_n\\(P_2)_{n+1}\\\vdots\\(P_2)_{n+M-1}\end{bmatrix} & \begin{matrix}(P_2)_n(x)\\(P_2)_{n+1}(x)\\\vdots\\(P_2)_{n+M-1}(x)\end{matrix}\\\hline \Pi_r[(P_2)_{n+M}] & (P_2)_{n+M}(x)\end{array}\right), \quad \frac{(\hat{P}_{1R})_n(x)}{(\hat{h}_R)_n} = \Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_{n+1}\\\vdots\\(P_2)_{n+M}\end{bmatrix} & \begin{matrix}0\\\vdots\\1\end{matrix}\\\hline \Pi_r[K^{[n+1]}(\cdot,x)] & 0\end{array}\right),$$

$$\frac{(\hat{h}_R)_n}{h_n} =\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_n\\(P_2)_{n+1}\\\vdots\\(P_2)_{n+M-1}\end{bmatrix} & \begin{matrix}1\\0\\\vdots\\0\end{matrix}\\\hline \Pi_r[(P_2)_{n+M}] & 0\end{array}\right).$$

Proof. We shall focus on the proof of the left-type deformation; the right-type one follows in a completely analogous way. Selecting the n-th component of the connection formula one gets

$$\begin{pmatrix}(\hat{\omega}_L)_{n,n} & (\hat{\omega}_L)_{n,n+1} & \dots & (\hat{\omega}_L)_{n,n+M-1} & 1\end{pmatrix}\begin{pmatrix}(P_1)_n(x)\\(P_1)_{n+1}(x)\\\vdots\\(P_1)_{n+M-1}(x)\\(P_1)_{n+M}(x)\end{pmatrix} = R(x)(\hat{P}_{1L})_n(x) .$$

Evaluating now in the zeroes of $R(x)$ it is easy to see that

$$\begin{pmatrix}(\hat{\omega}_L)_{n,n} & (\hat{\omega}_L)_{n,n+1} & \dots & (\hat{\omega}_L)_{n,n+M-1} & 1\end{pmatrix}\Pi\begin{bmatrix}(P_1)_n\\(P_1)_{n+1}\\\vdots\\(P_1)_{n+M-1}\\(P_1)_{n+M}\end{bmatrix} = \begin{pmatrix}0 & 0 & \dots & 0\end{pmatrix}$$

Therefore,

$$\begin{pmatrix}(\hat\omega_L)_{n,n} & (\hat\omega_L)_{n,n+1} & \dots & (\hat\omega_L)_{n,n+M-1}\end{pmatrix}\Pi_r\begin{bmatrix}(P_1)_n\\(P_1)_{n+1}\\\vdots\\(P_1)_{n+M-1}\end{bmatrix}=-\Pi_r[(P_1)_{n+M}]\ ,$$

i.e

$$\begin{pmatrix}(\hat\omega_L)_{n,n} & (\hat\omega_L)_{n,n+1} & \dots & (\hat\omega_L)_{n,n+M-1}\end{pmatrix}=-\Pi_r[(P_1)_{n+M}]\left(\Pi_r\begin{bmatrix}(P_1)_n\\(P_1)_{n+1}\\\vdots\\(P_1)_{n+M-1}\end{bmatrix}\right)^{-1}\ ,$$

from which the result for $\hat{P}_{1L}$ and $\hat{h}_L$ follow. In order to obtain the result for $\hat{P}_{2L}$, it is sufficient to start from the equation that relates the CD-Kernels, and to use the same procedure of evaluation on the zeroes of $R(x)$. □

Definition 22. *We introduce the* $(2M+1)$ *banded matrices*

$$(\hat\omega_L)(\hat\Omega_R):=\hat{J}_{1LR}\ ,\qquad\qquad (\hat\omega_R)(\hat\Omega_L):=\hat{J}_{2RL}\ .$$

We point out that a generalization of the notion of recurrence relation can be realized by allowing an intertwining of SBPS associated with different measure matrices instead of the same one. In this case we shall talk of a *quasi-recurrence relation.*

Proposition 19. *The right and left deformed SBPS satisfy the following* $(2M+1)$ *quasi–recurrence relation*

$$\begin{aligned}\hat{J}_{1LR}(\hat{P}_{R1})(x)&=R(x)(\hat{P}_{L1})(x)\ ,\\ \hat{J}_{2RL}(\hat{P}_{L2})(x)&=R(x)(\hat{P}_{R2})(x)\ ,\end{aligned}$$

with

$$\hat{J}_{1LR}=\hat{H}_L\left[\hat{J}_{2RL}\right]^\top\hat{H}_R^{-1}$$

Observe that if $\mathscr{W}\in\mathscr{W}_x$, then there would be no distinction between L or R sequences. In addition, if we choose $R(x)$ to be a polynomial of degree one, then $\hat\omega\cdot\hat\Omega$ is a $2(1)+1$-diagonal matrix and the standard three term recurrence relation is recovered.

6.4. **Geronimus-Sobolev transformations.** Let us now focus on the Geronimus transformation. To this aim, a polynomial $Q(x):=\prod_{i=1}^s(x-q_i)^{n_i}=Q_0+Q_1x+\dots+Q_{N-1}x^{N-1}+x^N$ of degree $\sum_{i=1}^s n_i=N$ is needed in order to define the left and right transformed measure matrices. We introduce the following auxiliary matrix, related to the polynomial $Q(x)$:

$$\mathbf{Q}:=\begin{pmatrix}Q_1 & Q_2 & Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots\\ Q_2 & Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots & \\ Q_3 & \dots & Q_{N-1} & 1 & 0 & \dots & & \\ \dots & Q_{N-1} & 1 & 0 & \dots & & & \\ Q_{N-1} & 1 & 0 & \dots & & & & \\ 1 & 0 & \dots & & & & & \\ 0 & \dots & & & & & & \end{pmatrix}.$$

Definition 23. *As long as $\{q_i\}_i \cap \Omega = \varnothing$ the Geronimus Sobolev deformed measure matrices are defined to be*

$$\check{\mathscr{W}}_L := [Q(\mathcal{X})]^{-1}\mathscr{W} + \sum_{i=1}^{s} \xi^{(i)}\delta(x-q_i)\mathrm{d}x \ , \qquad \check{\Omega}_L := \Omega \cup \{q_i\}_i$$

$$\check{\mathscr{W}}_R := \mathscr{W}\left[Q(\mathcal{X}^\top)\right]^{-1} + \sum_{i=1}^{s} \xi^{(i)}\delta(x-q_i)\mathrm{d}x \ , \qquad \check{\Omega}_R := \Omega \cup \{q_i\}_i$$

where $\xi^{(i)}$ are the $n_i \times n_i$ matrices of free parameters

$$\xi^{(i)} := \begin{pmatrix} \frac{\xi^{(i)}_{0,0}}{0!0!} & \frac{\xi^{(i)}_{0,1}}{0!1!} & \cdots & \frac{\xi^{(i)}_{0,n_i-1}}{(n_i-1)!(n_i-1)!} \\ \frac{\xi^{(i)}_{1,0}}{1!0!} & \ddots & & \\ \vdots & & \ddots & \\ \frac{\xi^{(i)}_{n_i-1,0}}{(n_i-1)!0!} & & & \frac{\xi^{(i)}_{n_i-1,n_i-1}}{(n_i-1)!(n_i-1)!} \end{pmatrix}, \qquad \mathring{\xi}^{(i)} := \begin{pmatrix} \xi^{(i)}_{0,0} & \xi^{(i)}_{0,1} & \cdots & \xi^{(i)}_{0,n_i-1} \\ \xi^{(i)}_{1,0} & \ddots & & \\ \vdots & & \ddots & \\ \xi^{(i)}_{n_i-1,0} & & & \xi^{(i)}_{n_i-1,n_i-1} \end{pmatrix}.$$

Proposition 20. *The transformed measure matrices and associated moment matrices are related to the original ones by the formulae*

$$\mathscr{W} := Q(\mathcal{X})\check{\mathscr{W}}_L \ , \qquad \mathscr{W} := \check{\mathscr{W}}_R Q(\mathcal{X}^\top) \ ,$$

$$G_{\mathscr{W}} = (Q(\Lambda))\, G_{\check{\mathscr{W}}_L} \ , \qquad G_{\mathscr{W}} = G_{\check{\mathscr{W}}_R}\,(Q(\Lambda))^\top \ .$$

The latter proposition and the assumption that the transformed moment matrices are LU-factorizable motivate the definition of the resolvents in terms of the following matrices.

Definition 24. *We introduce the matrices*

$$(\check{\omega}_L) := \check{H}_L\left(\check{S}_{1L}^{-1}\right)^\top Q(\Lambda^\top) S_1^\top H^{-1} = \check{S}_{2L} S_2^{-1} \ ,$$

$$(\check{\omega}_R) := \check{H}_R\left(\check{S}_{2R}^{-1}\right)^\top Q(\Lambda^\top) S_2^\top H^{-1} = \check{S}_{1R} S_1^{-1} \ .$$

The r.h.s. follow from the LU factorization of the transformed and non transformed moment matrices. It is not difficult to see that these equalities also imply that the resolvents are lower uni-triangular matrices with only N non-vanishing diagonals beneath the main one. Precisely:

$$\hat{\omega} = \begin{pmatrix} \check{\omega}_{0,0} & 0 & & & & & \\ \check{\omega}_{1,0} & \check{\omega}_{1,1} & & & & & \\ \vdots & \vdots & \ddots & & & & \\ \check{\omega}_{N,0} & \check{\omega}_{N,1} & & \check{\omega}_{N,N} & & & \\ 0 & \check{\omega}_{N+1,1} & & \check{\omega}_{N+1,N} & \check{\omega}_{N+1,N+1} & & \\ & & \ddots & & & \ddots & \\ & & & \check{\omega}_{k,k-N} & & & \check{\omega}_{k,k} \\ & & & & \ddots & & & \ddots \end{pmatrix},$$

where $\check{\omega}_{k,k-N} = \frac{\check{h}_k}{h_{k-N}}$ $\forall k > N$ and $\check{\omega}_{k,k} = 1$.

Proposition 21. *The Geronimus-Sobolev deformed polynomials and the associated second kind functions are related to the non transformed ones according to the formulae*

$$\check{\omega}_L P_2(x) = \check{P}_{2L}(x) \qquad \Longrightarrow \qquad \check{\omega}_L C_2(x) = Q(x)\check{C}_{2L}(x) - \check{H}_L\left(\check{S}_{1L}^{-1}\right)^\top \mathbf{Q}\chi(x)$$

$$\check{\omega}_R P_1(x) = \check{P}_{1R}(x) \qquad \Longrightarrow \qquad \check{\omega}_R C_1(x) = Q(x)\check{C}_{1R}(x) - \check{H}_R\left(\check{S}_{2R}^{-1}\right)^\top \mathbf{Q}\chi(x)$$

Proof. On the one hand the connection formulae for the polynomials follows straightforward remembering their definition in terms of the factorization matrices and from the definition of $\check{\omega}$. On the other hand the connection formulae for the second kind functions is a consequence of the former as we are about to prove. Let us prove it for the Right transformation, since the proof for the Left transformation needs of the same ideas. Firstly let us make the following definition and give a result that is easily verified

$$\Delta Q(x,y) := Q(x) - Q(y) \qquad\qquad \chi(x)^\top \mathbf{Q}\chi(y) = -\frac{\Delta Q(x,y)}{y-x}$$

Using this result the next chain of equalities can be followed

$$\begin{aligned}
&\check{\omega}_R C_1(y) - Q(y)\check{C}_{1R}(y) = \check{\omega}_R \left(P_1(x), \frac{1}{y-x}; \mathscr{W}\right) - Q(y)\left(\check{P}_{1R}(x), \frac{1}{y-x}; \check{\mathscr{W}}_R\right)\\
&= \left(\check{P}_{1R}(x), \frac{1}{y-x}; \mathscr{W}\right) - Q(y)\left(\check{P}_{1R}(x), \frac{1}{y-x}; \check{\mathscr{W}}_R\right) = \left(\check{P}_{1R}(x), \frac{1}{y-x}; \check{\mathscr{W}}_R\left[Q(\mathcal{X}^\top) - Q(y)\right]\right)\\
&= \int_\Omega \begin{pmatrix} \check{P}^{(0)}_{1R}(x) & \check{P}^{(1)}_{1R}(x) & \dots & \check{P}^{(k)}_{1R}(x) & \dots \end{pmatrix} \check{\mathscr{W}}_R \begin{pmatrix} \Delta Q(x,y) & & & \\ \frac{\partial}{\partial x}\Delta Q(x,y) & \Delta Q(x,y) & & \\ \frac{\partial^2}{\partial x^2}\Delta Q(x,y) & 2\frac{\partial}{\partial x}\Delta Q(x,y) & \Delta Q(x,y) & \\ \vdots & \vdots & & \ddots \end{pmatrix} \begin{pmatrix} \frac{1}{y-x} \\ \frac{\partial}{\partial x}\frac{1}{y-x} \\ \vdots \\ \frac{\partial^k}{\partial x^k}\frac{1}{y-x} \\ \vdots \end{pmatrix}\\
&= \left(\check{P}_{1R}(x), \frac{\Delta Q(x,y)}{y-x}; \check{\mathscr{W}}_R\right) = -\left(\check{P}_{1R}(x), \chi(x)^\top; \check{\mathscr{W}}_R\right)\mathbf{Q}\chi(y) = -\check{H}_R\left(\check{S}^{-1}_{2R}\right)^\top \mathbf{Q}\chi(y)
\end{aligned}$$

□

Let us now study the deformations of Christoffel–Darboux kernels.

Proposition 22. *The deformed Christoffel–Darboux kernels are related to the original ones by means of the formulae*

$$\begin{aligned}
&\check{K}^{[k]}_R(x,y) = Q(x)K^{[k]}(x,y) - \begin{pmatrix} (\check{P}_{2R})_k(x) & \dots & (\check{P}_{2R})_{k+N-1}(x) \end{pmatrix} \cdot\\
&\cdot \begin{pmatrix} (\check{h}_R)^{-1}_k & & \\ & \ddots & \\ & & (\check{h}_R)^{-1}_{k+N-1} \end{pmatrix} \begin{pmatrix} (\check{\omega}_R)_{k,k-N} & \dots & (\check{\omega}_R)_{k,k-1} \\ & \ddots & \vdots \\ & & (\check{\omega}_R)_{k+N-1,k-1} \end{pmatrix} \begin{pmatrix} (P_1)_{k-N}(y) \\ (P_1)_{k+1-N}(y) \\ \vdots \\ (P_1)_{k-1}(y) \end{pmatrix},
\end{aligned}$$

$$\begin{aligned}
&\check{K}^{[k]}_L(x,y) = Q(y)K^{[k]}(x,y) - \begin{pmatrix} (\check{P}_{1L})_k(x) & \dots & (\check{P}_{1L})_{k+N-1}(x) \end{pmatrix} \cdot\\
&\cdot \begin{pmatrix} (\check{h}_L)^{-1}_k & & \\ & \ddots & \\ & & (\check{h}_L)^{-1}_{k+N-1} \end{pmatrix} \begin{pmatrix} (\check{\omega}_L)_{k,k-N} & \dots & (\check{\omega}_L)_{k,k-1} \\ & \ddots & \vdots \\ & & (\check{\omega}_L)_{k+N-1,k-1} \end{pmatrix} \begin{pmatrix} (P_2)_{k-N}(x) \\ (P_2)_{k+1-N}(x) \\ \vdots \\ (P_2)_{k-1}(x) \end{pmatrix}.
\end{aligned}$$

Similarly, the mixed kernels $\forall k \geq N$ are related as follows

$$
\begin{aligned}
&Q(x)\mathcal{K}_2^{[k]}(x,y) - \begin{pmatrix}(\check{P}_{2R})_k(x) & \dots & (\check{P}_{2R})_{k+N-1}(x)\end{pmatrix} \cdot \\
&\cdot \begin{pmatrix}(\check{h}_R)_k^{-1} & & \\ & \ddots & \\ & & (\check{h}_R)_{k+N-1}^{-1}\end{pmatrix} \begin{pmatrix}(\check{\omega}_R)_{k,k-N} & \dots & (\check{\omega}_R)_{k,k-1} \\ & \ddots & \\ & & (\check{\omega}_R)_{k+N-1,k-1}\end{pmatrix} \begin{pmatrix}(C_1)_{k-N}(y) \\ (C_1)_{k+1-N}(y) \\ \vdots \\ (C_1)_{k-1}(y)\end{pmatrix} \\
&= Q(y)\check{\mathcal{K}}_{2R}^{[k]}(x,y) - \left(\chi^{[N]}(x)\right)^\top \mathbf{Q}\chi^{[N]}(y) \ , \\
&Q(y)\mathcal{K}_1^{[k]}(x,y) - \begin{pmatrix}(\check{P}_{1L})_k(y) & \dots & (\check{P}_{1L})_{k+N-1}(y)\end{pmatrix} \cdot \\
&\cdot \begin{pmatrix}(\check{h}_L)_k^{-1} & & \\ & \ddots & \\ & & (\check{h}_L)_{k+N-1}^{-1}\end{pmatrix} \begin{pmatrix}(\check{\omega}_L)_{k,k-N} & \dots & (\check{\omega}_L)_{k,k-1} \\ & \ddots & \\ & & (\check{\omega}_L)_{k+N-1,k-1}\end{pmatrix} \begin{pmatrix}(C_2)_{k-N}(x) \\ (C_2)_{k+1-N}(x) \\ \vdots \\ (C_2)_{k-1}(x)\end{pmatrix} \\
&= Q(x)\check{\mathcal{K}}_{1L}^{[k]}(x,y) - \left(\chi^{[N]}(y)\right)^\top \mathbf{Q}\chi^{[N]}(x) \ .
\end{aligned}
$$

Proof. These expressions are a direct consequence of the connection formulae. □

We shall also introduce a couple of useful matrices, which will be relevant in the subsequent discussion.

Definition 25. *Let*

$$
Q_i(x) := \frac{Q(x)}{(x-q_i)^{n_i}} \qquad \eta_{n_i\times n_i} := \begin{pmatrix}0 & 0 & \dots & 1 \\ 0 & 0 & 1 & 0 \\ \vdots & \ddots & & \vdots \\ 1 & & & 0\end{pmatrix}_{n_i\times n_i} \qquad i = 1,2,\dots,s
$$

We define the $N \times N$ matrices

$$
\Xi_L := \begin{pmatrix}\Xi_{L1} & 0 & \dots & 0 \\ 0 & \Xi_{L2} & 0 & \\ & & \ddots & \\ & & & \Xi_{Ls}\end{pmatrix}, \qquad \Xi_R := \begin{pmatrix}\Xi_{R1} & 0 & \dots & 0 \\ 0 & \Xi_{R2} & 0 & \\ & & \ddots & \\ & & & \Xi_{Rs}\end{pmatrix},
$$

where

$$
\Xi_{Rj} := \left(\mathring{\xi}^{(j)}\right)\left(\eta_{n_j\times n_j}\right) \begin{pmatrix}\frac{Q_j^{(0)}(q_j)}{0!} & \frac{Q_j^{(1)}(q_j)}{1!} & \dots & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} & \frac{Q_j^{(n_j-1)}(q_j)}{(n_j-1)!} \\ & \frac{Q_j^{(0)}(q_j)}{0!} & & & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} \\ & & \ddots & & \vdots \\ & & & \ddots & \frac{Q_j^{(1)}(q_j)}{1!} \\ & & & & \frac{Q_j^{(0)}(q_j)}{0!}\end{pmatrix}
$$

and

$$\Xi_{Lj} := \left(\mathring{\xi}^{(j)}\right)^\top \left(\eta_{n_j\times n_j}\right)\begin{pmatrix} \frac{Q_j^{(0)}(q_j)}{0!} & \frac{Q_j^{(1)}(q_j)}{1!} & \cdots & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} & \frac{Q_j^{(n_j-1)}(q_j)}{(n_j-1)!} \\ & \frac{Q_j^{(0)}(q_j)}{0!} & & & \frac{Q_j^{(n_j-2)}(q_j)}{(n_j-2)!} \\ & & \ddots & & \vdots \\ & & & \ddots & \frac{Q_j^{(1)}(q_j)}{1!} \\ & & & & \frac{Q_j^{(0)}(q_j)}{0!} \end{pmatrix}.$$

We define a couple of matrices useful in the discussion of transformed Genonimus-Sobolev polynomials.

Definition 26. *We introduce the $N\times N$ matrices*

$$\mathring{\Pi}_R := \left(\Pi_q\begin{bmatrix}(C_1)_0\\ \vdots \\ (C_1)_{N-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_0\\ \vdots \\ (P_1)_{N-1}\end{bmatrix}\Xi_R\right)\left(\mathbf{Q}^{[N]}\Pi_q[\chi^{[N]}]\right)^{-1}$$

$$\mathring{\Pi}_L := \left(\Pi_q\begin{bmatrix}(C_2)_0\\ \vdots \\ (C_2)_{N-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_2)_0\\ \vdots \\ (P_2)_{N-1}\end{bmatrix}\Xi_L\right)\left(\mathbf{Q}^{[N]}\Pi_q[\chi^{[N]}]\right)^{-1},$$

Where $\Pi_q[f]$ is the vector of germs associated to the set $q := \{q_i, n_i\}$.

An interesting characterization of the class of Geronimus-type transformed polynomials can be obtained in terms of quasi-determinants, as clarified by the following

Proposition 23. *Geronimus Sobolev transformed polynomials are expressed $\forall k \geq N$ in terms of the original polynomials via the formulae*

$$(\check{P}_{1R})_k = \Theta_*\left(\begin{array}{c|c} \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots \\ (C_1)_{k-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ \vdots \\ (P_1)_{k-1}\end{bmatrix}\Xi_R & \begin{matrix}(P_1)_{k-N}\\ \vdots \\ (P_1)_{k-1}\end{matrix} \\ \hline \Pi_q[(C_1)_k] - \Pi_q[(P_1)_k]\Xi_R & (P_1)_k(x)\end{array}\right),$$

$$\frac{(\check{P}_{2R})_k(x)}{(\check{h}_R)_k} = \Theta_*\left(\begin{array}{c|c} \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots \\ (C_1)_{k-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ \vdots \\ (P_1)_{k-1}\end{bmatrix}\Xi_R & \begin{matrix}1\\ \vdots \\ 0\end{matrix}, \\ \hline Q(x)\left(\Pi_q[\mathcal{K}_2^{[k]}(x,\cdot)] - \Pi_q[K^{[k]}(x,\cdot)]\Xi_R\right) + \left(\chi^{[N]}(x)\right)^\top\mathbf{Q}\Pi_q[\chi^{[N]}] & 0\end{array}\right),$$

$$(\check{h}_R)_k(x) = h_{k-N}\Theta_*\left(\begin{array}{c|c} \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots \\ (C_1)_{k-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ (P_1)_{k+1-N}\\ \vdots \\ (P_1)_{k-1}\end{bmatrix}\Xi_R & \begin{matrix}1\\ 0\\ \vdots \\ 0\end{matrix} \\ \hline \Pi_q[(C_1)_k] - \Pi_q[(P_1)_k]\Xi_R & 0\end{array}\right),$$

$$
(\check{P}_{2L})_k = \Theta_* \left(\begin{array}{c|c} \Pi_q \begin{bmatrix} (C_2)_{k-N} \\ \vdots \\ (C_2)_{k-1} \end{bmatrix} - \Pi_q \begin{bmatrix} (P_2)_{k-N} \\ \vdots \\ (P_2)_{k-1} \end{bmatrix} \Xi_L & \begin{matrix} (P_2)_{k-N} \\ \vdots \\ (P_2)_{k-1} \end{matrix} \\ \hline \Pi_q[(C_2)_k] - \Pi_q[(P_2)_k]\Xi_L & (P_2)_k(x) \end{array}\right),
$$

$$
\frac{(\check{P}_{1L})_k(x)}{(\check{h}_L)_k} = \Theta_* \left(\begin{array}{c|c} \Pi_q \begin{bmatrix} (C_2)_{k-N} \\ \vdots \\ (C_2)_{k-1} \end{bmatrix} - \Pi_q \begin{bmatrix} (P_2)_{k-N} \\ \vdots \\ (P_2)_{k-1} \end{bmatrix} \Xi_L & \begin{matrix} 1 \\ \vdots \\ 0 \end{matrix} \\ \hline Q(x)\left(\Pi_q[\mathcal{K}_1^{[k]}(\cdot,x)] - \Pi_q[K^{[k]}(\cdot,x)]\Xi_L\right) + \left(\chi^{[N]}(x)\right)^\top \mathbf{Q}\Pi_q[\chi^{[N]}] & 0 \end{array}\right),
$$

$$
\check{h}_{Lk}(x) = h_{k-N}\Theta_* \left(\begin{array}{c|c} \Pi_q \begin{bmatrix} (C_2)_{k-N} \\ \vdots \\ (C_2)_{k-1} \end{bmatrix} - \Pi_q \begin{bmatrix} (P_2)_{k-N} \\ (P_2)_{k+1-N} \\ \vdots \\ (P_2)_{k-1} \end{bmatrix} \Xi_L & \begin{matrix} 1 \\ 0 \\ \vdots \\ 0 \end{matrix} \\ \hline \Pi_q[(C_2)_k] - \Pi_q[(P_2)_k]\Xi_L & 0 \end{array}\right).
$$

For $k < N$ the following expressions hold

$$
(\check{P}_{1R})_k(x) = \Theta_* \left(\begin{array}{ccc|c} & \mathring{\Pi}_R^{[k]} & & \begin{matrix} (P_1)_0(x) \\ \vdots \\ (P_1)_{k-1}(x) \end{matrix} \\ \hline \left(\mathring{\Pi}_R\right)_{k,0} & \dots & \left(\mathring{\Pi}_R\right)_{k,k-1} & (P_1)_k(x) \end{array}\right),
$$

$$
(\check{P}_{2R})_k(x) = \Theta_* \left(\begin{array}{ccc|c} & \left(\mathring{\Pi}_R^\top\right)^{[k]} & & \begin{matrix} 1 \\ \vdots \\ x^{k-1} \end{matrix} \\ \hline \left(\mathring{\Pi}_R^\top\right)_{k,0} & \dots & \left(\mathring{\Pi}_R^\top\right)_{k,k-1} & x^k \end{array}\right)
$$

$$
(\check{h}_R)_k = -\Theta_* \left(\begin{array}{ccc|c} & \mathring{\Pi}_R^{[k]} & & \begin{matrix} \left(\mathring{\Pi}_R\right)_{0,k} \\ \vdots \\ \left(\mathring{\Pi}_R\right)_{k-1,k} \end{matrix} \\ \hline \left(\mathring{\Pi}_R\right)_{k,0} & \dots & \left(\mathring{\Pi}_R\right)_{k,k-1} & \left(\mathring{\Pi}_R\right)_{k,k} \end{array}\right),
$$

$$
(\check{P}_{2L})_k(x)=\Theta_*\left(\begin{array}{ccc|c}
 & \mathring{\Pi}_L^{[k]} & & \begin{matrix}(P_2)_0(x)\\ \vdots\\ (P_2)_{k-1}(x)\end{matrix}\\ \hline
\left(\mathring{\Pi}_L\right)_{k,0} & \dots & \left(\mathring{\Pi}_L\right)_{k,k-1} & (P_2)_k(x)
\end{array}\right),
$$

$$
(\check{P}_{1L})_k(x)=\Theta_*\left(\begin{array}{ccc|c}
 & \left(\mathring{\Pi}_L^\top\right)^{[k]} & & \begin{matrix}1\\ \vdots\\ x^{k-1}\end{matrix}\\ \hline
\left(\mathring{\Pi}_L^\top\right)_{k,0} & \dots & \left(\mathring{\Pi}_L^\top\right)_{k,k-1} & x^k
\end{array}\right),
$$

$$
(\check{h}_L)_k=-\Theta_*\left(\begin{array}{ccc|c}
 & \mathring{\Pi}_L^{[k]} & & \begin{matrix}\left(\mathring{\Pi}_L\right)_{0,k}\\ \vdots\\ \left(\mathring{\Pi}_L\right)_{k-1,k}\end{matrix}\\ \hline
\left(\mathring{\Pi}_L\right)_{k,0} & \dots & \left(\mathring{\Pi}_L\right)_{k,k-1} & \left(\mathring{\Pi}_L\right)_{k,k}
\end{array}\right).
$$

Proof. We shall focus on the case of right transformations. We start looking at the Geronimus transformed second kind functions

$$
(\check{C}_{1R})_k(y)=\left((\check{P}_{1R})_k,\frac{1}{y-x}\right)_{\check{\mathscr{W}}_R}=\int\begin{pmatrix}(\check{P}_{1R})_k & (\check{P}_{1R})_k' & \dots\end{pmatrix}\mathscr{W}\left[Q(\mathcal{X}^\top)\right]^{-1}\begin{pmatrix}\frac{1}{y-x}\\ \frac{\partial}{\partial x}\left(\frac{1}{y-x}\right)\\ \vdots\end{pmatrix}
$$

$$
+\sum_{j=1}^{s}\begin{pmatrix}\frac{(\check{P}_{1R})_k^{(0)}(q_j)}{0!} & \frac{(\check{P}_{1R})_k^{(1)}(q_j)}{1!} & \dots & \frac{(\check{P}_{1R})_k^{(n_j-1)}(q_j)}{(n_j-1)!}\end{pmatrix}\begin{pmatrix}\xi^{(j)}_{0,0} & \xi^{(j)}_{0,1} & \dots & \xi^{(j)}_{0,2} & \xi^{(j)}_{0,n_j-1}\\ \xi^{(j)}_{1,0} & \xi^{(j)}_{1,1} & & & \xi^{(j)}_{1,n_j-1}\\ \vdots & & \ddots & & \vdots\\ \xi^{(j)}_{n_j-1,n_j-1} & & & \ddots & \xi^{(j)}_{n_j-1,n_j-1}\end{pmatrix}(\eta)_{n_j\times n_j}\begin{pmatrix}\left(\frac{1}{y-q_j}\right)^{n_j}\\ \left(\frac{1}{y-q_j}\right)^{n_j-1}\\ \vdots\\ \frac{1}{y-q_j}\end{pmatrix}.
$$

Therefore, multiplying the previous expression by $Q(y)$ and letting $y\to q_j$, we obtain the Taylor expansion

$$
Q(y)(\check{C}_{1R})_k(y)=\begin{pmatrix}\frac{(\check{P}_{1R})_k^{(0)}(q_j)}{0!} & \frac{(\check{P}_{1R})_k^{(1)}(q_j)}{1!} & \dots & \frac{(\check{P}_{1R})_k^{(n_j-1)}(q_j)}{(n_j-1)!}\end{pmatrix}\Xi_{Rj}\begin{pmatrix}1\\ (y-q_j)\\ \vdots\\ (y-q_j)^{n_j-1}\end{pmatrix}+O(y-q_j)^{n_j}.
$$

The previous reasoning can be repeated for each j. Consequently, collecting all the information in the same matrix we can write the relation

$$
\Pi_q[Q(\check{C}_{1R})_k]=\Pi_q[(\check{P}_{1R})_k]\Xi_R.
$$

By using the connection formula for the second kind functions, and applying Π_q to both sides, if we also take into account the previous relation, we get the equations

$$\begin{aligned}\check{\omega}_R\Pi_q[C_1] &= \Pi_q[Q\check{C}_{1R}] - \check{H}_R\left(\check{S}_{2R}^{-1}\right)^\top \mathbf{Q}\Pi_q[\chi(x)] \ ,\\ \check{\omega}_R\Pi_q[C_1] &= \Pi_q[\check{P}_{1R}]\Xi_R - \check{H}_R\left(\check{S}_{2R}^{-1}\right)^\top \mathbf{Q}\Pi_q[\chi(x)] \ .\end{aligned}$$

Rearranging terms and using the connection formula for the polynomials, we arrive at the expression

$$\check{\omega}_R\left(\Pi_q[C_1] - \Pi_q[P_1]\Xi_R\right) = -\check{H}_R\left(\check{S}_{2R}^{-1}\right)^\top \mathbf{Q}\Pi_q[\chi(x)] \ .$$

This result can be made more explicit once written in the form

$$\begin{pmatrix}(\check{\omega}_R)_{k,k-N} & \dots & (\check{\omega}_R)_{k,k-1} & 1\end{pmatrix}\left(\Pi_q\begin{bmatrix}(C_1)_{k-N}\\(C_1)_{k-N+1}\\ \vdots\\ (C_1)_k\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\(P_1)_{k-N+1}\\ \vdots\\ (P_1)_k\end{bmatrix}\Xi_R\right) = 0, \qquad \forall k \geq N \ ,$$

whence, the expression for the first right-family and their norms follows straightforwardly. In order to obtain the expression for the second right-family, a similar approach can be used, based now on the relations between CD kernels and their mixed versions. For $k < N$ the expression for both families and norms is a consequence of the following LU-factorization of the matrix $\mathring{\Pi}_R$

$$(\check{\omega}_R)^{[N]}\left(\Pi_q\begin{bmatrix}(C_1)_0\\(C_1)_1\\ \vdots\\ (C_1)_N\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_0\\(P_1)_1\\ \vdots\\ (P_1)_N\end{bmatrix}\Xi_R\right) = -\check{H}_R^{[N]}\left(\left(\check{S}_{2R}^{-1}\right)^\top\right)^{[N]}\mathbf{Q}^{[N]}\Pi[\chi^{[N]}(x)] \qquad \Longrightarrow$$

$$\mathring{\Pi}_R = -\left(\check{\omega}_R^{-1}\right)^{[N]}\check{H}_R^{[N]}\left(\left(\check{S}_{2R}^{-1}\right)^\top\right)^{[N]} \ .$$

The proof for the case of the left deformation is completely analogous and is left to the reader.

□

We shall conclude this section with an observation on the recurrence relations for Geronimus-type polynomials arising from our transformation approach.

Definition 27. *Let us define the following matrices*

$$\check{J}_{1RL} := \check{S}_{1R}Q(\Lambda)\check{S}_{1L}^{-1} \qquad\qquad \check{J}_{2LR} := \check{S}_{2L}Q(\Lambda)\check{S}_{2R}^{-1} \ .$$

Proposition 24. *The matrices $\check{J}_{1RL}$ and $\check{J}_{2LR}$ possess a $2N+1$ diagonal structure and are related to each other according to the formulae*

$$\check{J}_{1RL}\check{H}_L = \check{H}_R\check{J}_{2LR}^\top \ .$$

These induce a left and right $2N+1$ term recurrence relation involving the Geronimus transformed polynomials:

$$\check{J}_{1RL}\check{P}_{1L} = Q(x)\check{P}_{1R} \qquad\qquad \check{J}_{2LR}\check{P}_{2R} = Q(x)\check{P}_{2L} \ .$$

Proof. The proposition is a consequence of the relation

$$Q(\Lambda)\check{G}_L = \check{G}_R Q(\Lambda^\top)$$

combined with a LU-factorization of the moment matrices. □

6.5. **Sobolev–linear spectral transformations.** After the previous discussion concerning both Christoffel and Geronimus Sobolev transformations, the successive composition of the last two follows straightforwardly. For this reason, proofs will be summarized or omitted in case they provide no new insight.

We start with the selection of two (co-prime) polynomials in order to deform an initial $\mathscr{W}(\Omega)$. Let these be $R(x) := \prod_{i=1}^{d}(x-r_i)^{m_i}$ of degree $\sum_{i=1}^{d} m_i = M$, and $Q(x) := \prod_{i=1}^{s}(x-q_i)^{n_i}$ of degree $\sum_{i=1}^{s} n_i = N$, where again we require that $\{q_i\}\cap\Omega=\varnothing$ in order to define what we understand for Sobolev linear spectral transformations.

Definition 28. *The Sobolev linear spectral deformed measure matrices are defined to be the composition of both a Geronimus and Christoffel transformation*

$$\begin{aligned}
\tilde{\mathscr{W}}_{RL} &:= \widehat{(\check{\mathscr{W}}_R)}_L = R(\mathcal{X})\mathscr{W}\left[Q(\mathcal{X}^\top)\right]^{-1} + \sum_{i=1}^{s} R(\mathcal{X})\xi^{(i)}\delta(x-q_i)\\
\tilde{\mathscr{W}}_{LR} &:= \widehat{(\check{\mathscr{W}}_L)}_R = \left[Q(\mathcal{X})\right]^{-1}\mathscr{W}R(\mathcal{X}^\top) + \sum_{i=1}^{s} \xi^{(i)}R(\mathcal{X}^\top)\delta(x-q_i)
\end{aligned}$$

Therefore transformed and non transformed moment matrices are related according to the formulae

$$R(\Lambda)G_{\mathscr{W}} = G_{\tilde{\mathscr{W}}_{RL}}Q(\Lambda^\top) \qquad\qquad Q(\Lambda)G_{\mathscr{W}} = G_{\tilde{\mathscr{W}}_{LR}}R(\Lambda^\top)\ .$$

After performing a LU-factorization of the moment matrices we are led to the following expressions.

Definition 29. *The resolvents and adjoint resolvents are defined as*

$$\begin{aligned}
(\tilde\omega_{RL}) &:= (\tilde S_{RL1})R(\Lambda)S_1^{-1} & \qquad (\tilde\Omega_{RL}) &:= S_2Q(\Lambda)(\tilde S_{RL2})^{-1}\\
(\tilde\omega_{LR}) &:= (\tilde S_{LR2})R(\Lambda)S_2^{-1} & \qquad (\tilde\Omega_{LR}) &:= S_1Q(\Lambda)(\tilde S_{LR1})^{-1}
\end{aligned}$$

and are related as follows

$$(\tilde\omega_{RL}) = (\tilde H_{RL})(\tilde\Omega_{RL})^\top H^{-1} \qquad\qquad (\tilde\omega_{LR}) = (\tilde H_{LR})(\tilde\Omega_{LR})^\top H^{-1}$$

The last relation induces a $N+M+1$ diagonal structure for them. For example $\tilde\omega$ has only non zero terms along the main diagonal together with N sub-diagonals and M super-diagonals. It also follows that $\tilde\omega_{k,k-N} = \frac{\tilde h_k}{h_{k-N}}$ and $\tilde\omega_{k,k+M}=1$.

Proposition 25. *The Sobolev linear spectral deformed polynomials and the associated second kind functions are related to the non transformed ones according to the formulae*

$$\begin{aligned}
\tilde\omega_{RL}P_1(x) &= R(x)\tilde P_{1RL}(x) & \qquad \tilde\omega_{RL}C_1(x) &= Q(x)\tilde C_{1RL}(x) - \tilde H_{RL}\left(\tilde S_{2RL}^{-1}\right)^\top \mathbf{Q}\chi(x)\\
\tilde\omega_{LR}P_2(x) &= R(x)\tilde P_{2LR}(x) & \qquad \tilde\omega_{LR}C_2(x) &= Q(x)\tilde C_{2LR}(x) - \tilde H_{LR}\left(\tilde S_{1LR}^{-1}\right)^\top \mathbf{Q}\chi(x)
\end{aligned}$$

Let us use the notation

$$A = \left(\begin{array}{c|c} A^{[k]} & A^{[k,\geq k]}\\ \hline A^{[\geq k,k]} & A^{[\geq k]}\end{array}\right)$$

in order to state the following

Definition 30. *We define the $(N+M)\times(N+M)$ matrices*

$$(\Upsilon_{RL})_k := \left(\begin{array}{c|c} 0_{M\times N} & -\left(\tilde h_{RL}\tilde\omega_{RL}\right)^{[k,\geq k]}\\ \hline \left(\tilde h_{RL}\tilde\omega_{RL}\right)^{[\geq k,k]} & 0_{N\times M}\end{array}\right)\ ,\quad (\Upsilon_{LR})_k := \left(\begin{array}{c|c} 0_{M\times N} & -\left(\tilde h_{LR}\tilde\omega_{LR}\right)^{[k,\geq k]}\\ \hline \left(\tilde h_{LR}\tilde\omega_{LR}\right)^{[\geq k,k]} & 0_{N\times M}\end{array}\right)\ .$$

Proposition 26. *The deformed Christoffel–Darboux kernels are related to the original ones by means of the formulae*

$$R(y)\tilde{K}^{[k]}_{RL}(x,y) = Q(x)K^{[k]}(x,y) - \big((\check{P}_{2RL})_{k-M}(x) \quad \dots \quad (\check{P}_{2RL})_{k+N-1}(x)\big)\,(\Upsilon_{RL})_k \begin{pmatrix} (P_1)_{k-N}(y) \\ (P_1)_{k+1-N}(y) \\ \vdots \\ (P_1)_{k+M-1}(y) \end{pmatrix}$$

$$R(y)\tilde{K}^{[k]}_{LR}(y,x) = Q(x)K^{[k]}(y,x) - \big((\check{P}_{1LR})_{k-M}(x) \quad \dots \quad (\check{P}_{1LR})_{k+N-1}(x)\big)\,(\Upsilon_{LR})_k \begin{pmatrix} (P_2)_{k-N}(y) \\ (P_2)_{k+1-N}(y) \\ \vdots \\ (P_2)_{k+M-1}(y) \end{pmatrix}$$

Similarly, the mixed kernels are related by means of the formulae

$$Q(y)\tilde{\mathcal{K}}^{[k]}_{2RL}(x,y) = Q(x)\mathcal{K}^{[k]}_2(x,y) - \big((\check{P}_{2RL})_{k-M}(x) \quad \dots \quad (\check{P}_{2RL})_{k+N-1}(x)\big)\,(\Upsilon_{RL})_k \begin{pmatrix} (C_1)_{k-N}(y) \\ (C_1)_{k+1-N}(y) \\ \vdots \\ (C_1)_{k+M-1}(y) \end{pmatrix}$$
$$+ \Big(\chi^{[N]}(x)\Big)^\top \mathbf{Q}\chi^{[N]}(y)$$

$$Q(y)\tilde{\mathcal{K}}^{[k]}_{1LR}(y,x) = Q(x)\mathcal{K}^{[k]}_1(y,x) - \big((\check{P}_{1LR})_{k-M}(x) \quad \dots \quad (\check{P}_{1LR})_{k+N-1}(x)\big)\,(\Upsilon_{LR})_k \begin{pmatrix} (C_2)_{k-N}(y) \\ (C_2)_{k+1-N}(y) \\ \vdots \\ (C_2)_{k+M-1}(y) \end{pmatrix}$$
$$+ \Big(\chi^{[N]}(x)\Big)^\top \mathbf{Q}\chi^{[N]}(y)$$

Since in the linear spectral type transformations two polynomials are involved, the presence of two vectors of germs is expected. As was done previously, we denote by $\Pi_r[f]$ the one related to the set $r := \{r_i, m_i\}_{i=1}^d$ and by $\Pi_q[f]$ the one related to $q := \{q_i, n_i\}_{i=1}^s$.

Proposition 27. *Sobolev linear spectral transformed polynomials are expressed $\forall k \geq N$ in terms of the original polynomials via the formulae*

$$(\tilde{P}_{1RL})_k(x) = \frac{1}{R(x)}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots\\ (C_1)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}\Xi_R & \begin{matrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{matrix}\\ \hline \Pi_r[(P_1)_{k+M}], \Pi_q[(C_1)_{k+M}] - \Pi_q[(P_1)_{k+M}]\Xi_R & (P_1)_{k+M}(x)\end{array}\right),$$

$$\frac{(\tilde{P}_{2RL})_k(x)}{(\tilde{h}_{RL})_k} = \Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots\\ (C_1)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}\Xi_R & \begin{matrix}1\\0\\ \vdots\\ 0\end{matrix}\\ \hline Q(x)\Pi_r[K^{[k]}(x,\cdot)], Q(x)\left(\Pi_q[\mathcal{K}_2^{[k]}(x,\cdot)] - \Pi_q[K^{[k]}(x,\cdot)]\Xi_R\right) + \left(\chi^{[N]}(x)\right)^\top \mathbf{Q}\Pi_q[\chi^{[N]}] & 0\end{array}\right),$$

$$(\tilde{h}_{RL})_k(x) = h_{k-N}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_1)_{k-N}\\ \vdots\\ (C_1)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_1)_{k-N}\\ \vdots\\ (P_1)_{k+M-1}\end{bmatrix}\Xi_R & \begin{matrix}1\\0\\ \vdots\\ 0\end{matrix}\\ \hline \Pi_r[(P_1)_{k+M}], \Pi_q[(C_1)_{k+M}] - \Pi_q[(P_1)_{k+M}]\Xi_R & 0\end{array}\right),$$

$$(\tilde{P}_{2LR})_k(x) = \frac{1}{R(x)}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_2)_{k-N}\\ \vdots\\ (C_2)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}\Xi_L & \begin{matrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{matrix}\\ \hline \Pi_r[(P_2)_{k+M}], \Pi_q[(C_2)_{k+M}] - \Pi_q[(P_2)_{k+M}]\Xi_L & (P_2)_{k+M}(x)\end{array}\right),$$

$$\frac{(\tilde{P}_{1LR})_k(x)}{(\tilde{h}_{LR})_k} = \Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_2)_{k-N}\\ \vdots\\ (C_2)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}\Xi_L & \begin{matrix}1\\0\\ \vdots\\ 0\end{matrix}\\ \hline Q(x)\Pi_r[K^{[k]}(\cdot,x)], Q(x)\left(\Pi_q[\mathcal{K}_1^{[k]}(\cdot,x)] - \Pi_q[K^{[k]}(\cdot,x)]\Xi_L\right) + \left(\chi^{[N]}(x)\right)^\top \mathbf{Q}\Pi_q[\chi^{[N]}] & 0\end{array}\right),$$

$$(\tilde{h}_{LR})_k(x) = h_{k-N}\Theta_*\left(\begin{array}{c|c} \Pi_r\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}, \Pi_q\begin{bmatrix}(C_2)_{k-N}\\ \vdots\\ (C_2)_{k+M-1}\end{bmatrix} - \Pi_q\begin{bmatrix}(P_2)_{k-N}\\ \vdots\\ (P_2)_{k+M-1}\end{bmatrix}\Xi_L & \begin{matrix}1\\0\\ \vdots\\ 0\end{matrix}\\ \hline \Pi_r[(P_2)_{k+M}], \Pi_q[(C_2)_{k+M}] - \Pi_q[(P_2)_{k+M}]\Xi_L & 0\end{array}\right).$$

For $k<N$ the following expressions hold

$$(\tilde{P}_{1RL})_k(x)=\frac{1}{R(x)}\Theta_*\left(\begin{array}{ccc|c} & \mathring{\Pi}_{RL}^{[k]} & & \begin{matrix}(P_1)_0(x)\\ \vdots \\ (P_1)_{k+N-1}(x)\end{matrix}\\ \hline \left(\mathring{\Pi}_{RL}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{RL}\right)_{k,k-1} & (P_1)_{k+N}(x)\end{array}\right),$$

$$(\tilde{P}_{2RL})_k(x)=\Theta_*\left(\begin{array}{ccc|c} & \left(\mathring{\Pi}_{RL}^{\top}\right)^{[k]} & & \begin{matrix}1\\ x\\ \vdots \\ x^{k-1}\end{matrix}\\ \hline \left(\mathring{\Pi}_{RL}^{\top}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{RL}^{\top}\right)_{k,k-1} & x^k\end{array}\right),$$

$$(\tilde{h}_{RL})_k=-\Theta_*\left(\begin{array}{ccc|c} & \mathring{\Pi}_{RL}^{[k]} & & \begin{matrix}\left(\mathring{\Pi}_{RL}\right)_{0,k}\\ \vdots \\ \left(\mathring{\Pi}_{RL}\right)_{k-1,k}\end{matrix}\\ \hline \left(\mathring{\Pi}_{RL}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{RL}\right)_{k,k-1} & \left(\mathring{\Pi}_{RL}\right)_{k,k}\end{array}\right),$$

and

$$(\tilde{P}_{2LR})_k(x)=\frac{1}{R(x)}\Theta_*\left(\begin{array}{ccc|c} & \mathring{\Pi}_{LR}^{[k]} & & \begin{matrix}(P_2)_0(x)\\ \vdots \\ (P_2)_{k+N-1}(x)\end{matrix}\\ \hline \left(\mathring{\Pi}_{LR}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{LR}\right)_{k,k-1} & (P_2)_{k+N}(x)\end{array}\right),$$

$$(\tilde{P}_{1LR})_k(x)=\Theta_*\left(\begin{array}{ccc|c} & \left(\mathring{\Pi}_{LR}^{\top}\right)^{[k]} & & \begin{matrix}1\\ x\\ \vdots \\ x^{k-1}\end{matrix}\\ \hline \left(\mathring{\Pi}_{LR}^{\top}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{LR}^{\top}\right)_{k,k-1} & x^k\end{array}\right),$$

$$(\tilde{h}_{LR})_k=-\Theta_*\left(\begin{array}{ccc|c} & \mathring{\Pi}_{LR}^{[k]} & & \begin{matrix}\left(\mathring{\Pi}_{LR}\right)_{0,k}\\ \vdots \\ \left(\mathring{\Pi}_{LR}\right)_{k-1,k}\end{matrix}\\ \hline \left(\mathring{\Pi}_{LR}\right)_{k,0} & \dots & \left(\mathring{\Pi}_{LR}\right)_{k,k-1} & \left(\mathring{\Pi}_{LR}\right)_{k,k}\end{array}\right),$$

where the $(N+M)\times(N+M)$ matrices $\mathring{\Pi}$ are defined by

$$\mathring{\Pi}_{RL} := \left[\Pi_r \begin{bmatrix}(P_1)_0\\ \vdots \\ (P_1)_{N+M-1}\end{bmatrix}, \left(\Pi_q \begin{bmatrix}(C_1)_0\\ \vdots \\ (C_1)_{N-1}\end{bmatrix} - \Pi_q \begin{bmatrix}(P_1)_0\\ \vdots \\ (P_1)_{N-1}\end{bmatrix} \Xi_R\right)\left(\mathbf{Q}^{[N]}\Pi_q[\chi^{[N]}]\right)^{-1}\right],$$

$$\mathring{\Pi}_{LR} := \left[\Pi_r \begin{bmatrix}(P_2)_0\\ \vdots \\ (P_2)_{N+M-1}\end{bmatrix}, \left(\Pi_q \begin{bmatrix}(C_2)_0\\ \vdots \\ (C_2)_{N-1}\end{bmatrix} - \Pi_q \begin{bmatrix}(P_2)_0\\ \vdots \\ (P_2)_{N-1}\end{bmatrix} \Xi_L\right)\left(\mathbf{Q}^{[N]}\Pi_q[\chi^{[N]}]\right)^{-1}\right].$$

7. Deformations arising from the action of linear differential operators

In this Sobolev context, where derivatives are ubiquitous, the polynomial deformation theory seems to be missing something. For that reason, in this section we will now discuss a different, more general class of deformations, obtained when a differential operator acts on one of the entries of the bilinear form. Although a general theory like the one for Darboux–Sobolev deformations is not available yet, some steps and results in that direction, together with some easy examples, can be proposed. To address this question, let us start with the derivative operator. We have

$$DG_{\mathscr{W}} = D\boldsymbol{D}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top = \boldsymbol{D}\begin{pmatrix}0 & 0 & 0 & 0 & \dots\\ \mathbb{I} & 0 & 0 & 0 & \dots\\ 0 & \mathbb{I} & 0 & 0 & \dots\\ 0 & 0 & \mathbb{I} & 0 & \dots\\ 0 & 0 & 0 & \mathbb{I} & \ddots\\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}\left(\int_\Omega \chi(x)\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top =$$

$$\boldsymbol{D}\left(\int_\Omega \chi(x)\begin{pmatrix}0 & 0 & 0 & 0 & \dots\\ 1 & 0 & 0 & 0 & \dots\\ 0 & 1 & 0 & 0 & \dots\\ 0 & 0 & 1 & 0 & \dots\\ 0 & 0 & 0 & 1 & \ddots\\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}\mathscr{W}\chi(x)^\top\right)\boldsymbol{D}^\top.$$

Consequently, we can obtain immediately the following result.

Theorem 4. *The relations*

$$(f',h;\mathscr{W}) = (f,h;\Lambda^\top\mathscr{W}) \qquad DG_{\mathscr{W}} = G_{\Lambda^\top\mathscr{W}}$$
$$(f,h';\mathscr{W}) = (f,h;\mathscr{W}\Lambda) \qquad G_{\mathscr{W}}D^\top = G_{\mathscr{W}\Lambda}$$

hold.

By linearity, we deduce that given any linear differential operator $\boldsymbol{L} := \sum_{n,r=0}^\infty a_{n,r}x^n\frac{\mathrm{d}^r}{\mathrm{d}x^r}$, acting on one of the entries of our inner product, we can translate its action into a matrix multiplying the initial moment matrix $L := \sum_{n,r=0}^\infty a_{n,r}D^r\Lambda^n$ or into a matrix multiplying the initial measure matrix $\mathcal{L} = \sum_{n,r=0}^\infty a_{n,r}(\Lambda^\top)^r\mathcal{X}^n$.

The interplay among the three different actions $\boldsymbol{L}, L, \mathcal{L}$ is clarified in the next

Proposition 28. *We have*

$$(\boldsymbol{L}_1[f],\boldsymbol{L}_2[h];\mathscr{W}) = (f,h:\mathcal{L}_1\mathscr{W}\mathcal{L}_2^\top), \qquad L_1G_{\mathscr{W}}(L_2)^\top = G_{\mathcal{L}_1\mathscr{W}(\mathcal{L}_2)^\top}.$$

This is a direct generalization of Proposition 15. Provided both $G_{\mathscr{W}}$ and $G_{\mathcal{L}_1\mathscr{W}(\mathcal{L}_2)^\top}$ are LU-factorizable, this proposition could allow us, in some particular cases, to relate the SBPS associated to each of the two moment matrices.

A couple of interesting, nontrivial problems arise from the last discussion.

- Determine a pair $(\boldsymbol{L}_1, \boldsymbol{L}_2)$ of linear differential operators with associated $(\mathcal{L}_1, \mathcal{L}_2)$ such that $\mathcal{L}_1\mathscr{W} \sim \mathscr{W}\mathcal{L}_2^\top$ (and therefore $L_1 G_{\mathscr{W}} = G_{\mathscr{W}} L_2^\top$).
- Determine a pair of operators $(\boldsymbol{L}_1, \boldsymbol{L}_2)$ with associated $(\mathcal{L}_1, \mathcal{L}_2)$ such that $\mathcal{L}_1\mathscr{W}_1\mathcal{L}_2^\top \sim \mathscr{W}_2$ and $\mathscr{W}_1, \mathscr{W}_2$ have some "suitable" properties.

An answer to the first problem would ensure that the associated SBPS possess many interesting properties. For instance, the special case where the usual three term recurrence relation holds is just a particular answer to this question for $\boldsymbol{L}_1 = \boldsymbol{L}_2 = x$. Another example of this kind was given in proposition 14 with the operator $\mathbf{F}$.
We will devote the next section to a partial answer to the second problem.

7.1. **Orthogonal polynomials with respect to differential operators.** For the second problem some simple cases can be tackled. The idea behind it is to start with a simple measure matrix $\mathscr{W}_1$ and deform it by means of differential operators into a new one $\mathscr{W}_2 \sim \mathcal{L}_1\mathscr{W}_1(\mathcal{L}_2)^\top$ so that we can establish explicit relations between $G_{\mathscr{W}_1}$ and $G_{\mathscr{W}_2}$. If both moment matrices are LU-factorizable, they may lead to relations between their associated SBPS. For example, one can start with the standard (non Sobolev) matrix $\mathscr{W}_1 = E_{00}\omega$. This case deserves special attention since it connects usual moment matrices with certain Sobolev moment matrices in a direct way. This entails the possibility to relate the associated OPS and SBPS as well. This section is intimately related to the notion of orthogonality with respect to a differential operator (OPDO) [2]. Here we start from the standard orthogonality, in order to obtain connections between standard and Sobolev polynomials. A similar approach could be used in the more general case of a diagonal matrix $\mathscr{W}$. In that case, we would be able to relate Sobolev orthogonal polynomials associated to different measure matrices.

Proposition 29. *Given two linear differential operators $\boldsymbol{L}_\alpha := \sum_k p_{\alpha,k}(x)\frac{\mathrm{d}^k}{\mathrm{d}x^k}$, $\alpha = 1,2$, with $p_{\alpha,k}(x)$ polynomials of any degree for all k, the following relation between the standard inner product involving these differential operators and a Sobolev bilinear function exists*

$$\langle \boldsymbol{L}_1[f], \boldsymbol{L}_2[h]\rangle_\mu = \big(f, h; \mathscr{W}_{\boldsymbol{L}_{1,2}}\big) \ .$$

The relation between the associated Sobolev moment matrix and the standard one reads

$$L_1 g_\mu (L_2)^\top = G_{\mathscr{W}_{\boldsymbol{L}_{1,2}}},$$

and the measure matrix is

$$\mathscr{W}_{\boldsymbol{L}_{1,2}} = \begin{pmatrix} p_{1,0}p_{2,0} & p_{1,0}p_{2,1} & p_{1,0}p_{2,2} & \cdots \\ p_{1,1}p_{2,0} & p_{1,1}p_{2,1} & p_{1,1}p_{2,2} & \cdots \\ p_{1,2}p_{2,0} & p_{1,2}p_{2,1} & p_{1,2}p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \mathrm{d}\mu(x) \ .$$

Proof. Since g_μ is the usual moment matrix associated to the measure $\mathrm{d}\mu(x)$ we have

$$\langle \boldsymbol{L}_2[f], \boldsymbol{L}_1[h]\rangle_\mu = \big(f, h; [\mathcal{L}_1 E_{0,0}(\mathcal{L}_2)^\top \mathrm{d}\mu]\big) \qquad\qquad L_1 g (L_2)^\top = G_{[\mathcal{L}_1 E_{0,0}(\mathcal{L}_2 E_{0,0})^\top \mathrm{d}\mu]} \ .$$

Note that the shape of $[\mathcal{L}_1 E_{0,0}(\mathcal{L}_2 E_{0,0})^\top \mathrm{d}\mu]$ is particularly simple: it is quite straightforward to see that

$$[\mathcal{L}_1 E_{0,0}(\mathcal{L}_2 E_{0,0})^\top \mathrm{d}\mu] = \begin{pmatrix} p_{1,0}(x) \\ p_{1,1}(x) \\ p_{1,2}(x) \\ \vdots \end{pmatrix} \cdot \begin{pmatrix} p_{2,0}(x) & p_{2,1}(x) & p_{2,2}(x) & \ldots \end{pmatrix} \mathrm{d}\mu(x) = \begin{pmatrix} p_{1,0}p_{2,0} & p_{1,0}p_{2,1} & p_{1,0}p_{2,2} & \cdots \\ p_{1,1}p_{2,0} & p_{1,1}p_{2,1} & p_{1,1}p_{2,2} & \cdots \\ p_{1,2}p_{2,0} & p_{1,2}p_{2,1} & p_{1,2}p_{2,2} & \cdots \\ \vdots & \vdots & \vdots & \end{pmatrix} \mathrm{d}\mu(x) \ .$$

□

Definition 31. *Given two families of linear differential operators* $S=\{\{\boldsymbol{L}_k\},\{\boldsymbol{U}_k\}\}_{k=0}^{\mathcal{N}}$ *with*

$$\boldsymbol{L}_k=\frac{\mathrm{d}^k}{\mathrm{d}x^k}+\sum_{j=k+1}l_{jk}(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}, \qquad \boldsymbol{U}_k=\frac{\mathrm{d}^k}{\mathrm{d}x^k}+\sum_{j=k+1}u_{kj}(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}$$

and a set of measures $\{\mathrm{d}\mu_k(x)\}_{k=0}^{\mathcal{N}}$*, we shall call the function*

$$(f,h)_S:=\sum_{k=0}^{\mathcal{N}}\langle\boldsymbol{L}_k[f],\boldsymbol{U}_k[h]\rangle_{\mu_k}$$

the generalized diagonal Sobolev bilinear function.

Shall we had $l_{jk}(x)=0=u_{kj}(x)\ \forall k,j$ the generalized diagonal Sobolev bilinear function would be indeed the usual diagonal Sobolev bilinear function.

Proposition 30. *Given a* $(\mathcal{N}+1)\times(\mathcal{N}+1)$ *measure matrix satisfying* $\det\mathscr{W}^{[k]}(x)\neq 0\ \forall x\in\Omega$ *and* $k=0,1\dots,\mathcal{N}$*, then the Sobolev bilinear function* $(f,h;\mathscr{W})$ *is equivalent to a generalized diagonal Sobolev bilinear function* $(f,h)_S$*. The pair* $S=\{\{\boldsymbol{L}_k\},\{\boldsymbol{U}_k\}\}_{k=0}^{\mathcal{N}}$ *with*

$$\boldsymbol{L}_k=\frac{\mathrm{d}^k}{\mathrm{d}x^k}+\sum_{j=k+1}l_{jk}(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j} \qquad \boldsymbol{U}_k=\frac{\mathrm{d}^k}{\mathrm{d}x^k}+\sum_{j=k+1}u_{kj}(x)\frac{\mathrm{d}^j}{\mathrm{d}x^j}$$

is determined by the LU factorization of $\mathscr{W}$ *by means of the relations*

$$\mathscr{W}(x)=\begin{pmatrix}1&&&\\ l_{10}(x)&1&&\\ l_{20}(x)&l_{21}(x)&1&\\ \vdots&\vdots&&\ddots\\ l_{\mathcal{N}0}(x)&l_{\mathcal{N}1}(x)&&&1\end{pmatrix}\begin{pmatrix}\mathrm{d}\mu_0(x)&&&&\\ &\mathrm{d}\mu_1(x)&&&\\ &&\ddots&&\\ &&&\ddots&\\ &&&&\mathrm{d}\mu_{\mathcal{N}}(x)\end{pmatrix}.$$

$$\cdot\begin{pmatrix}1&u_{01}(x)&u_{02}(x)&\dots&u_{0\mathcal{N}}(x)\\ &1&u_{12}(x)&\dots&u_{1\mathcal{N}}(x)\\ &&1&&\\ &&&\ddots&\\ &&&&u_{\mathcal{N}-1\mathcal{N}}(x)\\ &&&&1\end{pmatrix}$$

In addition, if each $\mathrm{d}\mu_k(x)$ *is positive definite and* $l_{j,k}(x),u_{k,j}(x)$ *are polynomials satisfying the relations*

$$j-\deg[u_{k,j}(x)]>k \qquad \text{and} \qquad j-\deg[l_{j,k}(x)]>k,$$

then $G_{\mathscr{W}}$ *is LU-factorizable and therefore has an associated SBPS.*

Proof. The first part of the proposition is an easy generalization of Proposition 29, since the LU factorization of $\mathscr{W}$ can be understood as follows

$$\mathscr{W}=\left[\mathcal{L}_0E_{0,0}(\mathcal{U}_0E_{0,0})^\top\omega_0\right]+\left[\mathcal{L}_1E_{0,0}(\mathcal{U}_1E_{0,0})^\top\omega_1\right]+\dots+\left[\mathcal{L}_{\mathcal{N}}E_{0,0}(\mathcal{U}_{\mathcal{N}}E_{0,0})^\top\omega_{\mathcal{N}}\right] .$$

Therefore, we have that $(f,h;\mathscr{W})=\sum_{k=0}^{\mathcal{N}}\langle\boldsymbol{L}_k[f],\boldsymbol{U}_k[h]\rangle_{\mu_k}$ or equivalently $G_{\mathscr{W}}=\sum_{k=0}^{N}L_kg_{\mu_k}(U_k)^\top$. This expression, together with the fact that the condition on the degrees of $u_{k,j}(x)$ and $l_{k,j}(x)$ is equivalent to requiring that

L_k and U_k have the shape of D^k + *diagonals beneath this one* (also equivalent to $\chi^{[k]} \in ker\boldsymbol{U}_k, \chi^{[k]} \in ker\boldsymbol{L}_k$), make the reasoning of the positive definiteness of $G_{\mathscr{W}}$ exactly the same as the one we used for the positive definite diagonal case. □

7.2. **Examples where SBPS and OPS can be related in terms of differential operators.** Let us show in more detail some examples where the relation between OPS and SBPS can be explicitly constructed. Assume that $\boldsymbol{L}_\alpha$ satisfy the two conditions

- $\deg[p_{\alpha,k} \leq k]$, $\forall k$. This implies that $L_\alpha \in \mathscr{L}$.
- both $\boldsymbol{L}_\alpha$ are invertible operators.

For these cases the LU factorization of $L_1 g(L_2)^\top$ is trivial. If $g = S^{-1}h\left(S^{-1}\right)^\top$ it is easy to see that

$$L_1 g(L_2)^\top = [S(L_1)^{-1}]^{-1}h\left([S(L_2)^{-1}]^{-1}\right)^\top \ .$$

This means that we can write the SBPS from the OPS. Indeed,

$$P_1(x) = SL_1^{-1}\chi(x), \qquad\qquad P_2(x) = S(L_2)^{-1}\chi(x) \ .$$

Let us discuss a couple of examples of this kind.

(1) Consider a $\mathscr{W}$ of the form

$$\mathscr{W}(x) := \begin{pmatrix} 1 & -1 & 0 & 0 & \dots \\ -1 & 1 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \mathrm{d}\mu(x) \ .$$

This measure matrix comes from the operator $\boldsymbol{L}_\alpha = 1 - \frac{\mathrm{d}}{\mathrm{d}x}$, which of course satisfies the two conditions above. The related moment matrix reads

$$G_{\mathscr{W}} = (\mathbb{I}-D)g(\mathbb{I}-D)^\top \qquad\qquad \text{where}(\mathbb{I}-D)^{-1} = \sum_{n=0}^{\infty} D^n \ .$$

Thus,

$$G_{\mathscr{W}} = \left[S(\mathbb{I}-D)^{-1}\right]^{-1} H\left(\left[S(\mathbb{I}-D)^{-1}\right]^{-1}\right)^\top \ .$$

We conclude that the SOPS associated with $G_{\mathscr{W}}$ is related to the OPS associated to ω as follows

$$P(x) = S(\mathbb{I}-D)^{-1}\chi(x) = S(\sum_{n=0}^{\infty} D^n)\chi(x) = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ S_{1,0} & 1 & 0 & 0 & \dots \\ S_{2,0} & S_{2,1} & 1 & 0 & \dots \\ S_{3,0} & S_{3,1} & S_{3,2} & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} 1 \\ x+1 \\ x^2+2x+2 \\ x^3+3x^2+6x+6 \\ \vdots \end{pmatrix} .$$

(2) We start with a $\mathscr{W}$ of the form

$$\mathscr{W}(x) := \begin{pmatrix} 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ 1 & 1 & 1 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix} \mathrm{d}\mu(x) \ .$$

It is not hard to see that the two previous conditions are fulfilled. This allows us to write explicitly

$$G_{\mathscr{W}}=\sum_{k=0}^{\infty}D^kS^{-1}H(S^{-1})^\top(\sum_{k=0}^{\infty}D^k)^\top=[S(\mathbb{I}-D)]^{-1}H\left[[S(\mathbb{I}-D)]^{-1}\right]^\top .$$

Thus, the associated OPS is nothing but

$$P(x)=S(\mathbb{I}-D)\chi(x)=\begin{pmatrix}1&0&0&0&\dots\\S_{1,0}&1&0&0&\dots\\S_{2,0}&S_{2,1}&1&0&\dots\\S_{3,0}&S_{3,1}&S_{3,2}&1&\dots\\\vdots&\vdots&\vdots&\vdots&\ddots\end{pmatrix}\begin{pmatrix}1\\x-1\\x^2-2x\\x^3-3x^2\\\vdots\end{pmatrix}.$$

(3) Now we shall consider a matrix measure of the kind

$$\mathscr{W}(x):=\begin{pmatrix}\frac{a^0}{0!0!}&\frac{a^1}{0!1!}&\frac{a^2}{0!2!}&\frac{a^3}{0!3!}&\cdots\\\frac{a^1}{1!0!}&\frac{a^2}{1!1!}&\frac{a^3}{1!2!}&\frac{a^4}{1!3!}&\cdots\\\frac{a^2}{2!0!}&\frac{a^3}{2!1!}&\frac{a^4}{2!2!}&\frac{a^5}{2!3!}&\cdots\\\frac{a^3}{3!0!}&\frac{a^4}{3!1!}&\frac{a^5}{3!2!}&\frac{a^6}{3!3!}&\cdots\\\vdots&\vdots&\vdots&\vdots&\ddots\end{pmatrix}\mathrm{d}\mu(x) .$$

Remarkably, $\mathscr{W}(x)\in\mathscr{W}_x$. Its expression corresponds to the one in eq. (21) by choosing $\mathrm{d}\mu_k=\frac{a^k\mathrm{d}\mu}{k!}$. The previous theory allows us to write

$$G_{\mathscr{W}}=\sum_{k=0}^{\infty}\frac{a^kD^k}{k!}g\left(\sum_{r=0}^{\infty}\frac{a^rD^r}{r!}\right)^\top=\exp\{aD\}g\exp\{aD^\top\}=[S\exp\{-aD\}]^{-1}H\left[[S\exp\{-aD\}]^{-1}\right]^\top .$$

This expression implies that the associated SOPS is nothing but the usual one OPS associated with ω but with an shift by a in the independent variable, i.e.

$$P=S\exp\{-aD\}\chi(x)=S\chi(y), \qquad y=(x-a) .$$

Let us mention here that when $a=1$, this example establishes a connection between "Hankel transforms" (as defined in [16]) and Sobolev Polynomials to light. One can show that the matrices that act to the left and right of the initial sequence (the initial moment matrix g) are $\left(\frac{D^k}{k!}\right)_{l,j}=\binom{l}{j}$. In other words, we recover the so called "Binomial transform" of the initial sequence, under which the Hankel transform remains invariant.

Appendix A. A relation with integrable hierarchies of Toda type

The purpose of this final section is to clarify the connection of the present theory of Sobolev bi-orthogonal polynomials with the theory of integrable systems.

As usual in this context, one can start from a suitable deformation of the moment matrix with certain appropriate matrices. These matrices involve the exponential of a linear combination of two set of times and the powers of the matrices Λ. Inspired by this approach, we shall generalize to our framework some well-known results.
To this aim, let us introduce two different sets of real deformation parameters $t_a=\{t_{a,0}=0,t_{a,1},t_{a,2},\dots\}$ for $a=1,2$, which will allow us to deform the moment matrix according to the following prescription.

Definition 32. *We define the time-deformed moment matrix*

$$G_{\mathscr{W}}(t)=W_{1,0}(t_1)G_{\mathscr{W}}[W_{2,0}(t_2)]^{-1} \tag{22}$$

where the deformation matrices $W_{1,0}(t_1)$ *and* $W_{1,0}(t_2)$ *are given by*

$$W_{1,0}(t_1) = \exp\left(\sum_{j=0}^{\infty} t_{1,j}\Lambda^j\right) \qquad\qquad W_{2,0}(t_2) = \exp\left(\sum_{j=0}^{\infty} t_{2,j}\left(\Lambda^\top\right)^j\right)$$

As the following result shows, the reason for this deformation of the moment matrix is that it can be directly translated into a deformation of the corresponding measure matrix.

Theorem 5. *The deformed moment matrix* $G_{\mathscr{W}}(t)$ *can be written as the moment matrix associated to a time dependent measure matrix, this is*

$$G_{\mathscr{W}}(t) = G_{\mathscr{W}(t)}$$

where the new time dependent measure matrix is given by the following expression

$$\mathscr{W}(t) := [\mathscr{W}_{1,0}(t_1, x)]\, \mathscr{W}\, [\mathscr{W}_{2,0}(t_2, x)]^{-1} = \left[\exp\left(\sum_{j=0}^{\infty} t_{1,j}\mathcal{X}^j\right)\right] \mathscr{W} \left[\exp\left(-\sum_{j=0}^{\infty} t_{2,j}\left(\mathcal{X}^\top\right)^j\right)\right] .$$

It is worth pointing out that $\mathscr{W}_{1,0}(t_1, x)$ is upper triangular while $\mathscr{W}_{2,0}(t_2, x)$ is lower triangular. As an example

$$\exp(t\mathcal{X}) = \begin{pmatrix} \binom{0}{0} t^0 & \binom{1}{0} t^1 & \binom{2}{0} t^2 & \binom{3}{0} t^3 & \dots \\ & \binom{1}{1} t^{1-1} & \binom{2}{1} t^{2-1} & \binom{3}{1} t^{3-1} & \dots \\ & & \binom{2}{2} t^{2-2} & \binom{3}{2} t^{3-2} & \dots \\ & & & \binom{3}{3} t^{3-3} & \dots \\ & & & & \ddots \end{pmatrix} \exp(tx) .$$

Once the moment matrix is deformed, in case we can still LU-factorize it we can write

$$G_{\mathscr{W}}(t) = S_1(t)\left(S_2(t)\right)^{-1} , \tag{23}$$

which leads to the time dependent Sobolev orthogonal polynomial sequences. This factorization also is the key for the following

Definition 33. *The wave semi-infinite matrices are*

$$W_1(t) := S_1(t) W_{1,0}(t_1) \qquad\qquad W_2(t) := S_2(t) W_{2,0}(t_2) .$$

These are indeed related to the initial moment matrix.

Proposition 31. *The following relation hold*

$$G_{\mathscr{W}} = \left(W_1(t)\right)^{-1} W_2(t)$$

Proof. From eqs. (22) and (23) we can see that

$$G_{\mathscr{W}} = \left(W_{1,0}(t_1)\right)^{-1}\left(S_1(t)\right)^{-1} S_2(t) W_{2,0}(t_2) = \left(W_1(t)\right)^{-1} W_2(t) . \tag{24}$$

□

We shall introduce two operators that will be relevant hereon.

Definition 34. *The Lax operators associated with our moment matrix are*

$$L_1 := S_1 \Lambda S_1^{-1} \qquad\qquad L_2 := S_2 \Lambda^\top S_2^{-1} .$$

It is important to remark here that in contrast with what happens in the standard theory of deformation of moment matrices, where $L_1 = L_2$ (because both coincide with the tri-diagonal Jacobi matrix responsible for the usual three term recurrence relation), this is no longer the case in the Sobolev context. Indeed, $\Lambda G_{\mathscr{W}} \neq G_{\mathscr{W}}\Lambda^\top$. Thus $L_1 \neq L_2$ and we can only infer that L_1 is a lower triangular matrix with an extra diagonal over the main one, while L_2 is an upper triangular matrix with an extra diagonal beneath the main one.

Proposition 32. *For $a = 1, 2$ we have the following differential equations for the wave semi infinite matrices*

$$\frac{\partial W_a}{\partial t_{1,j}} W_a^{-1} = (L_1^j)_+ \qquad \frac{\partial W_a}{\partial t_{2,j}} W_a^{-1} = (L_2^j)_- \ .$$

Here $(A)_-$ is the projection of the matrix A onto the space of strictly lower triangular matrices while $(A)_+$ is its projection onto the space of upper triangular matrices.

Proof. Deriving eq. (24), on one hand we can obtain that

$$\frac{\partial W_1}{\partial t_{a,j}} W_1^{-1} = \frac{\partial W_2}{\partial t_{a,j}} W_2^{-1} \qquad a = 1, 2; \quad j = 1, 2, 3, \dots \ .$$

On the other hand,

$$\frac{\partial S_1}{\partial t_{1,j}} S_1^{-1} + S_1 \Lambda^j S_1^{-1} = \frac{\partial S_2}{\partial t_{1,j}} S_2^{-1} \qquad \frac{\partial S_2}{\partial t_{2,j}} S_2^{-1} + S_2 \left(\Lambda^\top\right)^j S_2^{-1} = \frac{\partial S_1}{\partial t_{2,j}} S_1^{-1} \ .$$

Decomposing them in their upper and strictly lower projections leads to the result of the proposition. □

The results of these proof can also be used to prove the next interesting result.

Proposition 33. *The following Lax equations hold*

$$\frac{\partial L_a^j}{\partial t_{b,r}} = \left[(L_b^j)_{(-1)^{b+1}}, L_a^j\right]$$

or explicitly

$$\frac{\partial L_1^j}{\partial t_{1,r}} = \left[(L_1^j)_+, L_1^j\right] \qquad \frac{\partial L_1^j}{\partial t_{2,r}} = \left[(L_2^j)_-, L_1^j\right]$$

$$\frac{\partial L_2^j}{\partial t_{1,r}} = \left[(L_1^j)_+, L_2^j\right] \qquad \frac{\partial L_2^j}{\partial t_{2,r}} = \left[(L_2^j)_-, L_2^j\right]$$

The compatibility equations of these give rise to the classical Zakharov–Shabat equations.

Proposition 34. *Wave functions evaluated at different times t and t' satisfy the relation*

$$W_1(t)W_1(t')^{-1} = W_2(t)W_2(t')^{-1}.$$

Proof. From Proposition 31 we derive the equality

$$(W_1(t))^{-1}W_2(t) = G = (W_1(t'))^{-1}W_2(t') \ ,$$

from which the result follows immediately. □

Acknowledgments.
The research of G. A. and M. M. has been supported by the research project [MTM2015-65888-C4-3-P] *"Ortogonalidad, teoría de la aproximación y aplicaciones en fsica matemática"'* MINECO, Spain. G. A. also thanks the Program *"Ayudas para Becas y Contratos Complutenses Predoctorales en Espaa"' 2011*, Universidad Complutense de Madrid, Spain.
The research of P. T. has been partly supported by the research project FIS2015-63966, MINECO, Spain, and by the ICMAT Severo Ochoa project SEV-2015-0554 (MINECO).

References

[1] P. Althammer, Eine Erweiterung des Orthogonalitätsbegriffes bei Polynomen und deren Anwendung auf die beste Approximation, *Journal für die Reine und Angewandte Mathematik*, **211**, 192–204 (1962).

[2] A.I. Aptekarev, G. T. López Lagomasino and F. Marcellán, Orthogonal polynomials with respect to a differential operator. Existence and uniqueness. *Rocky Mountain Journal of Mathematics* **32**, 467-481 (2002).

[3] C. Álvarez Fernández, G. Ariznabarreta, J. C. Garcia-Ardilla, M. Mañas and F. Marcellán, Christoffel transformations for matrix orthogonal polynomials in the real line and the non-Abelian 2D Toda lattice hierarchy, *International Mathematical Research Notices* doi:10.1093/imrn/rnw027 (2016).

[4] C. Álvarez Fernández, G. Ariznabarreta, J. C. Garcia-Ardilla, M. Mañas and F. Marcellán, Transformation theory and Christoffel formulas for matrix Bi orthogonal polynomials on the real line. *Arxiv:1605.04617v6* [math.CA] (2016).

[5] G. Ariznabarreta, M. Mañas, Matrix orthogonal matrix polynomials on the unit circle and Toda type integrable systems, *Advances in Mathematics* **264**, 396-463 (2014).

[6] D. Barrios Rolania, G. López Lagomasino, H. Pijeira Cabrera, The moment problem for a Sobolev inner product, *Journal of Approximation Theory* **100**, 364-380 (1999).

[7] E. B. Christoffel, Über die Gaussische Quadratur und eine Verallgemeinerung derselben. *Journal für die Reine und Angewandte Mathematik* (Crelle's journal) **55**, 61-82 (1858) (in German).

[8] A. J. Durań, A generalization of Favard's theorem for polynomials satisfying a recurrence relation, *Journal of Aproximation Theory* **74**, 83-109 (1993).

[9] A. J. Durán and M. D. de la Iglesia, Differential equations for discrete Laguerre-Sobolev orthogonal polynomials, *Journal of Approximation Theory*, in press. arXiv:1309.6259.

[10] W. D. Evans, L. L. Littlejohn, F. Marcellán, C. Markett and A. Ronveaux, On recurrence relations for Sobolev orthogonal polynomials, *SIAM Journal of Mathematical Analysis* **26** No.2, 446-467 (1995).

[11] J. Geronimus, On polynomials orthogonal with regard to a given sequence of numbers and a theorem by W. Hahn, *Izvestiya Akademii Nauk SSSR* **4**, 215228 (1940) (in Russian).

[12] L. Golinskii, On the scientific legacy of Ya. L. Geronimus (to the hundredth anniversary), in *Self-Similar Systems* (Proceedings of the International Workshop (July 30 - August 7, Dubna, Russia, 1998)), 273-281, Edited by V.B. Priezzhev and V. P. Spiridonov, Publishing Department, Joint Institute for Nuclear Research, Moscow Region, Dubna.

[13] A. Iserles, P. E. Koch, S. P. Norsett and J. M. Sanz Serna, On Polynomials Orthogonal with respect to Certain Sobolev Inner Products, *Journal of Approximation Theory* **65**, 151-175 (1991).

[14] S. Kakutani, Concrete representation of abstract (M)-spaces. (A characterization of the space of continuous functions.), Ann. of Math. (2), 42: 994-1024, (1941).

[15] R. Koekoek, The search for differential equations for certain sets of orthogonal polynomials, *Journal of Computational and Applied Mathematics* **49**, 111119 (1993).

[16] J. W. Layman, The Hankel transform and some of its properties, *Journal of integer Sequences* 4, 01.1.5 (2001).

[17] F. Marcellán, R. Álvarez-Nodarse, On the Favard theorem and its extensions, *Journal of Computational and Applied Mathematics* **127**, 231-254 (2001).

[18] F. Marcellán and J. J. Moreno-Balcázar, Asymptotics and zeros of Sobolev orthogonal polynomials on unbounded supports, *Acta Applicandae Mathematicae* **94**, 163192 (2006).

[19] F. Marcellán, T. E. Pérez, M. A. Pinar, Laguerre Sobolev orthognal polynomials, *Journal of Computational and Applied Mathematics* **71**, 245-265 (1996).

[20] F. Marcellán, T. E. Pérez, M. A. Pinar and A. Ronveaux, General Sobolev Orthogonal Polynomials, *Journal of Mathematical Analysis and Applications* **200**, 614-634 (1996).

[21] F. Marcellán, M. F. Pérez-Valero, Y. Quintana, A. Urieles, Recurrence relations and outer relative asymptotics of orthogonal polynomials with respect to a discrete Sobolev type inner product, *Bulletin of Mathematical Sciences* **4**, 83-97 (2014).

[22] F. Marcellán, J. C. Petronilho. Orthogonal polynomials and coherent pairs: The classical case. *Indagationes Mathematicae* **3**, 287-307 (1995).

[23] F. Marcellán, F. H. Szafraniec, A matrix algorithm towards solving the moment problem of Sobolev type, *Linear Algebra and its applications* **331**, 155-164 (2001).

[24] F. Marcellán and Yuan Xu, On Sobolev Orthogonal Polynomials, *Arxiv:1403.6249v1* [math.CA] (2014).

[25] H. G. Meijer, A short history of orthogonal polynomials in a Sobolev space I. The non-discrete case, Report 95-51, Technische Universiteit Delft (1995).

[26] H. G. Meijer, Determination of all coherent pairs, *Journal of Approximation Theory* **89**, 321-343 (1997).

[27] P. J. Olver, On multivariate interpolation, *Studies in Applied Mathematics* **116**, 201-240 (2006).

[28] H. E. Pijeira Cabrera, Teoria de Momentos y Propiedades Asintoticas para Polinomios Ortogonales de Sobolev, Doctoral Dissertation, Universidad Carlos III de Madrid, (1998).

[29] F. W. Schäfke, Zu den Orthogonalpolynomen von Althammer, *Journal für die Reine und Angewandte Mathematik* **252**, 195-199 (1972).

[30] F. W. Schäke and G. Wolf, Einfache verallgemeinerte klassische Orthogonal polynome, *Journal für die Reine und Angewandte Mathematik* **262/263**, 339355 (1973).

[31] M. Z. Spivey, Laura L. Steil, The k-Binomial transforms and the Hankel transforms, *Journal of Integer Sequences* **9**, 06.1.1 (2006).

[32] I. M. Gel'fand, G. E. Shilov, Generalized functions, *Academic press inc.* New York, (1964).

Departamento de Física Teórica II, Métodos y modelos matemáticos, Facultad de Físicas, Universidad Complutense, 28040 – Madrid, Spain

Departamento de Física Teórica II, Métodos y modelos matemáticos, Facultad de Físicas, Universidad Complutense, 28040 – Madrid, Spain

Departamento de Física Teórica II, Métodos y modelos matemáticos, Facultad de Físicas, Universidad Complutense, 28040 – Madrid, Spain and Instituto de Ciencias Matemáticas, C/ Nicolás Cabrera, No 13–15, 28049 Madrid, Spain

Bibliografía

[1] M. J. Ablowitz, D. J. Kaup, A. C. Newell and H. Segur, Method for solving the sine-Gordon equation, *Physical Review Letters*, **30**, 1262-1264, (1963).

[2] N. Abramesco, Résumé des principales propriétés des polynômes orthogonaux, *Nouvelles annales de mathématiques, journal des candidats aux Écoles Polytechnique et Normale*, **5**, 166-177, (1923).

[3] N. Abramesco, Les polynômes orthogonaux, *Annales de la Faculté des Sciences de Toulouse : Mathématiques*, **3**, 67-87, (1932).

[4] M. Adler and P. van Moerbeke, Generalized orthogonal polynomials, discrete KP and Riemann–Hilbert problems, *Communications in Mathematical Physics*, **207**, 589-620, (1999).

[5] M. Adler and P. van Moerbeke, Vertex operator solutions to the discrete KP hierarchy, *Communications in Mathematical Physics*, **203**, 185-210, (1999).

[6] M. Adler and P. van Moerbekee, The spectrum of coupled random matrices, *Annals of Mathematics*, **149**, 921-976, (1999).

[7] M. Adler and P. van Moerbeke, Hermitian, symmetric and symplectic random ensembles: PDEs for the distribution of the spectrum, *Annals of Mathematics*, **153**, 149-189, (2001).

[8] M. Adler and P. van Moerbeke, Darboux transforms on band matrices, weights and associated polynomials, *International Mathematics Research Notices*, **18**, 935-984, (2001).

[9] M. Adler, P. van Moerbeke, and P. Vanhaecke, Moment matrices and multi- component KP, with applications to random matrix theory, *Communications in Mathematical Physics*, **286**, 1-38, (2009).

[10] P. Althammer, Eine Erweiterung des Orthogonalitätsbegriffes bei Polynomen und deren Anwendung auf die beste Approximation, *Journal für die Reine und Angewandte Mathematik*, **211**, 192-204, (1962).

[11] C. Álvarez-Fernández, U. Fidalgo, M. Mañas, Multiple orthogonal polynomials of mixed type: Gauss–Borel factorization and the multi-compoment 2D Toda hierarchy, *Advances in Mathematics*, **227**, 1451-1525, (2011).

[12] C. Álvarez-Fernández, M. Mañas, Orthogonal Laurent polynomials on the unit circle, extended CMV ordering and 2D Toda type integrable hierarchies, *Advances in Mathematics*, **240**, 132-193, (2013).

[13] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas and F. Marcellán, Transformation theory and Christoffel formulas for matrix biorthogonal polynomials on the real line (with applications to the non-Abelian Toda lattice and nonconmmutative KP hierarchies), arXiv:1605.04617.

[14] C. Álvarez-Fernández, G. Ariznabarreta, J. C. García-Ardila, M. Mañas and F. Marcellán, Christoffel Transformations for Matrix Orthogonal Polynomials in the Real line and the non-Abelian 2D Toda Lattice Hierarchy, *International Mathematics Research Notices.* **5**, 1285-1341, (2017).

[15] P. Apell and J. K. Fériet, Fonctions hypergéométriques et hypershpériques, polynômes d'Hermite, Gauthiers-Villars, Paris, (1926).

[16] A. I. Aptekarev and E. M. Nikishin, The scattering problem for a discrete Sturm-Liouville operator, *Matematicheskii Sbornik*, **121**, 327-358, (1983).

[17] A. I. Aptekarev, Multiple orthogonal polynomials, *Journal of Computational and Applied Mathematis*,**99**, 423-447, (1998).

[18] A.I. Aptekarev, P.M. Bleher and A.B.J. Kuijlaars, Large n limit of Gaussian random matrices with external source, Part II, *Communications in Mathematical Physics*, **259**, 367-289, (2005).

[19] G. Ariznabarreta and M. Mañas, Matrix orthogonal Laurent polynomials on the unit circle and Toda type integrable systems, *Advances in Mathematics*, **264**, 396-463, (2014).

[20] G. Ariznabarreta and M. Mañas, A Jacobi type Christoffel-Darboux formula for multiple orthogonal polynomials of mixed type, *Linear Algebra and its Applications* **468**, 154-170, (2014).

[21] G. Ariznabarreta and M. Mañas, Multivariate orthogonal polynomials and integrable systems, *Advances in Mathematics* **302**,628-739, (2016)

[22] G. Ariznabarreta and M. Mañas, Darboux transformations for multivariate orthogonal polynomials, arXiv:1503.04786.

[23] G. Ariznabarreta and M. Mañas, Linear spectral transformations for multivariate orthogonal polynomials and multispectral Toda Hierarchies, arXiv:1511.09129.

[24] G. Ariznabarreta and M. Mañas, Multivariate orthogonal Laurent polynomials and integrable systems, arXiv:1506.08708.

[25] G. Ariznabarreta, M. Mañas and A. Toledano, CMV biorthogonal Laurent polynomials: Christoffel formulas for Christoffel and Geronimus perturbations, arXiv:1610.02008.

[26] G. Ariznabarreta, M. Mañas and A. Toledano, CMV biorthogonal Laurent polynomials II: Christoffel formulas for Geronimus-Uvarov transformations, arXiv:1611.03547.

[27] G. Ariznabarreta, M. Mañas and P. Tempesta Generalized Sobolev orthogonal polynomials, matrix moment problems and integrable systems, arXiv:1612.07229.

[28] R. Askey, Orthogonal Polynomials and Special Functions, *Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, **21** (1975).

[29] Yu. M. Berezanskiî, Expansions in Eigenfunctions of Self-adjoint Operators, *Translations of Mathematical Monographs, American Mathematical Society*, **17**, (1968).

[30] C. Berg, Fibonacci numbers and orthogonal polynomials, *Journal of Computational and Applied Mathematics*, **17**, 75-88, (2011).

[31] S. Bochner, Über Sturm-Liouvillesche Polynomsysteme, *Mathematische Zeitschrift*, **29**, 730-736, (1929).

[32] E. Bour, Théorie de la déformation des surfaces, *Journal de l'École Imperiale Polytechnique*, **19**, 1-48, (1862).

[33] J. Boussinesq, Essai sur la theorie des eaux courantes, *Mémoires présentées par divers savants à l'Académie des Sciences de l'Institut de France*, **23**, 1-680, (1877).

[34] M. G. de Bruin, Simultaneous Padé approximants and orthogonality, in "Polynômes Orthogonaux et Applications" (C. Brezinski et al., eds.), *Lecture Notes in Mathematics 1171, Springer-Verlag*, 74-83, (1985).

[35] A. Bultheel, P. González-Vera, E. Hendriksen and O. Njästad, Orthogonal Rational Functions, Cambridge University Press, (1999).

[36] M. Cafasso, Matrix Biorthogonal Polynomials on the unit circle and the non-Abelian Ablowitz-Ladik hierarchy, *Journal of Physics A: Mathematical and Theoretical*, **42**, 211-365, (2009).

[37] M.J. Cantero, L. Moral and L. Velázquez, Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle, *Linear Algebra and its Applications*, **362**, 29-56, (2003).

[38] M. M. Castro and F. A. Grünbaum, Orthogonal matrix polynomials satisfying first order differential equations: a collection of instructive examples, *Journal of Nonlinear Mathematical Physics*, **4**, 63-76, (2005).

[39] P. L. Chebyshev, Sur les values limitées des intégrales, *Journal des Mathématiques Pures et Appliquées*, **19**, 157-170, (1874).

[40] T. S. Chihara, An introduction to orthogonal polynomials, *Mathematics and its applications*, **213**, Gordon and Breach Science Publishers, (1978).

[41] E. B. Christoffel, Über die Gaussische Quadratur und eine Verallgemeinerung derselben, *Journal für die Reine und Angewandte Mathematik* (Crelle's journal), **55**, 61-82, (1858) (in German).

[42] A. R. Collar, On the reciprocation of certain matrices, *Proceedings of the Royal Society of Edinburgh Section A Mathematics.*, **59**, 195-206, (1939).

[43] N. Copernicus, De revolutionibus orbium coelestium, libri VI, *Johhanes Petreius*, Nurenberg, (1543).

[44] E. Daems and A. B. J. Kuijlaars, A Christoffel-Darboux formula for multiple orthogonal polynomials, *Journal of Approximation Theory*, **130**, 188-200, (2004).

[45] E. Daems and A. B. J. Kuijlaars, Multiple orthogonal polynomials of mixed type and non-intersecting Brownian motions, *Journal of Approximation Theory*, **146**, 91-114, (2007).

[46] D. Damanik, A. Pushnitsky and B. Simon, The Analytic Theory of Matrix Orthogonal Polynomials, *Surveys in Approximation Theory*, **4**, 1-85, (2008).

[47] G. Darboux, Sur une proposition relative aux équations linéaires, *Comptes rendus hebdomadaires des séancesde l'Académie des Sciences*, **94**, 1456-1459, Paris, (1882).

[48] G.Darboux, Leçons sur la Théorie Générale des Surfaces, 2éme partie, Gauthier-Villars, Paris, (1889).

[49] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa, Operator approach to the Kadomtsev-Petviashvili equation. Transformation groups for soliton equations. III, *Journal of the Physical Society of Japan*, **50**, 3806-3812, (1981).

[50] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa, Transformation groups for soliton equations. Euclidean Lie algebras and reduction of the KP hierarchy, *Publications of the Research Institute for Mathematical Sciences*, **18**, 1077-1110, (1982).

[51] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa, Transformation groups for soliton equations in *Nonlinear Integrable Systems-Classical Theory and Quantum Theory, M. Jimbo and T. Miwa (eds.). World Scientific*, Singapore, (1983).

[52] P. Delsarte, Y. V. Genin, and Y. G. Kamp, Orthogonal polynomial matrices on the unit circle, *IEEE Transactions of Circuits and Systems CAS-25*, **3**, 149-160, (1978).

[53] C.F. Dunkl and Y. Xu, Orthogonal polynomials of several variables, *Encyclopedia of Mathematics and its Applications, Cambridge University Press*, **155**, New York, (2014).

[54] A. J. Durán, On orthogonal polynomials with respect to positive definite matrix of measures, *Canadian Journal of Mathematics*, **47**, 88-112, (1995).

[55] A. J. Durán and W. Van Assche, Orthogonal matrix polynomials and higher-order recurrence relations, *Linear Algebra and its Applications*, **219**, 261-280, (1995).

[56] A. J. Durán and F. A. Grünbaum, A survey on orthogonal matrix polynomials satisfying second-order differential equations, *Journal of Computational and Applied Mathematics*, **178**, 169-90, (2005).

[57] L. P. Eisenhart, Transformations of Surfaces, Princeton University Press, Hamburg, (1943).

[58] A. Erdélyi, W. Magnus, F. Oberhettinger and F. G. Tricomi, Higher transcendental functions, *McGraw-Hill*, **2**, New York, (1953).

[59] L. D. Faddeev, Properties of the S-matrix of the one-dimensional Schrödinger equation, *American Mathematical Society Translation Series*, **65**, 139-166, (1967).

[60] J. Favard, Sur les polynômes de Tchebicheff, *Comptes Rendus de l'Académie des Sciences*, Paris, **200**, 2052-2053, (1935).

[61] E. Fermi, J. Pasta and S. Ulam, Studies of non linear problems, *Document LA-1940* Los Alamos National Laboratory, (1955).

[62] H. Flaschka, The Toda lattice. II. Existence of integrals, *Physical Review B, (Solid State)*, **9**, (1974).

[63] G. Freud, Orthogonal Polynomials, Pergamon Press, Oxford, (1971).

[64] C. S. Gardner, J. M. Greene, M. D. Kruskal and R. M. Miura, Method for solving the Korteweg-de Vries equation, *Physical Review Letters*, **19**, 1095-1097, (1967).

[65] C. F. Gauss, Methodus nova integralium valores per approximationem inveniendi, *Commentationes Societatis regiae scientarium Gottingensis recentiores*, **3** 39-76, (1814).

[66] W. Gautschi, Orthogonal Polynomials: Computation and Approximation, *Numerical Mathematics and Scientific Computation*, Oxford University Press, New York, (2004).

[67] I. M. Gel'fand and V. S. Retakh, Determinants of matrices over nonconmmutative rings, *Functional Analysis and Its Applications*, **25**, 91-102, (1991).

[68] I. M. Gel'fand, S. Gel'fand, V. S. Retakh and R. Wilson, Quasideterminants, *Advances in Mathematics* **193**, 56-141, (2005).

[69] J. S. Geronimo, Matrix orthogonal polynomials on the unit circle, *Journal of Mathematical Physics, 22*, **7**, 1359-1365, (1981).

[70] J. S. Geronimo, Scattering theory and matrix orthogonal polynomials on the real line, *Circuits Systems Signal Process 1*, **3-4**, 471-495, (1982).

[71] Y. L. Geronimus, Generalized orthogonal polynomials and Cristoffel–Darboux formula, *Doklady Akademii Nauk SSSR*, **26**, 843-846, (1940).

[72] Y. L. Geronimus, On polynomials orthogonal with regard to a given sequence of numbers and a theorem by W. Hahn, *Izvestiya Akademii Nauk SSSR*, **4**, 215-228, (1940).

[73] Y. L. Geronimus, Polynomials Orthogonal on a Circle and Their Applications. Series and Approximations, *American Mathematical Society Translation Series* **3**, 1-78, Providence, Rhode Island, (1962).

[74] W. Hahn, Über die Jacobischen Polynome and zwei verwandte Polynomklassen, *Mathematische Zeitschrift*, **39**, 634-638, (1935).

[75] W. Hahn, Über höhere Abeitungen von Orthogonalpolynomen, *Mathematische Zeitschrift*, **43**, 101, (1937).

[76] H. Hamburger, Uber eine Erweiterung des Stieltjesschen Momentproblem, *Journal Mathematische Annalen* **81**, 235-318, (1920), **82**, 168-187, (1921).

[77] G. H. Hardy, A Mathematician's Apology (1940), *Cambridge University press*, (2004).

[78] E. V. Haynsworth, On the Schur Complement, *Basel Mathematical Notes* **20**, University of Basel, (1968).

[79] F. Hausdorff, Momentprobleme fur ein endliches intervall, *Journal Mathematische Zeitschrift* **16**, 220-248, (1923).

[80] C. Hermite, Sur la fonction exponentielle, *Comptes Rendus de l'Académie des Sciences* **77**, 18-24, 74-79, 226-233, 285-293 (1873); reprinted in his Oeuvres, Tome III, Gauthier-Villars, Paris, 150-181, (1912).

[81] A. Iserles, P. E. Koch, S. P. Norsett and J. M. Sanz Serna, On Polynomials Orthogonal with respect to Certain Sobolev Inner Products, *Journal of Approximation Theory* **65**, 151-175, (1991).

[82] M. E. H. Ismail, Classical and Quantum orthogonal polynomials in one variable, *Encyclopedia of Mathematics and its Applications*, **98**, Cambridge University Press, (2005).

[83] C. G. J. Jacobi, Über Gauss' neue Methode, die Werthe der Integrale näherungsweise zu finden, *Journal für die reine und angewandte Mathematik (Crelle's Journal)*, **1**, 301-308, (1826).

[84] D. Jackson, Formal properties of orthogonal polynomials in two variables, *Duke Mathematical Journal*, **2**, 423-434, (1936).

[85] W. B. Jones, W. J. Thron and H. Waadeland, A strong Stieltjes moment problem, *Transactions of the American Mathematical Society*, **261**, 503-528, (1980).

[86] W. B. Jones, W. J. Thron, Orthogonal Laurent polynomials and Gaussian quadrature, in: K. E. Gustafson, W. P. Reinhardt (Eds.), Quantum Mechanics in Mathematics Chemistry and Physics, Plenum, New York, 449-455, (1981).

[87] S. Kakutani, Concrete representation of abstract (M)-spaces. (A characterization of the space of continuous functions.), *Annals of Mathematics*, **42**, 994-1024, (1941).

[88] T. H. Koornwinder, Two variable analogues of the classical orthogonal polynomials, *Theory and Applications of Special Functions,* R. Askey, ed., Academic Press, NY, 435-495, (1975).

[89] D. J. Korteweg and G. de Vries, On the change of form of long waves advancing in a rectangular channel, and a new type of long stationary waves, *Philosophical Magazine*, **39**, 422-443, (1985).

[90] H. L. Krall and I. M. Sheffer, Orthogonal polynomials in two variables, *Annali di Matematica Pura ed Applicata*, **76**, 325-376, (1967).

[91] M. G. Krein, The fundamental propositions of the theory of representations Hermitian operators with deficiency index (m, m), *Ukranian Mathematical Journal*, **1** Providence, RI, 3-66, (1949).

[92] P. D. Lax, Integrals of nonlinear equations of evolution and solitary waves, *Communications in Pure and Applied Mathematics*, **21**, 467-490, (1968).

[93] D.C. Lewis, Polynomial least square approximations, *American Journal of Mathematics*, **69**, 273-278, (1947).

[94] F. Marcellán and J. C. Petronilho, Orthogonal polynomials and coherent pairs: The classical case, *Indagationes Mathematicae* **3**, 287-307, (1995).

[95] F. Marcellán and R. Álvarez Nodarse, On the “Favard theorem” and its extensions, *Journal of Computational and Applied Mathematics* **127**, 231-254, (2001).

[96] F. Marcellán and Yuan Xu, On Sobolev Orthogonal Polynomials, *Expositiones Mathematicae*, **33**, 308-352, (2015).

[97] V. A. Marchenko, Sturm-Liouville operators and applications, Birkhäuser, Basel, (1986).

[98] A. Markov, On mean values and exterior densities, *Matematicheskii Sbornik 4* **46**, 165-191, (1938).

[99] P. Maroni, Sur la suite de polynômes orthogonaux associée à la forme $u = \delta_c + \lambda(x-c)^{-1}$, *Periodica Mathematica Hungarica*, **21**, 223-248, (1990).

[100] V. B. Matveev and M. A. Salle, Differential-difference evolution equations. II: Darboux Transformation for the Toda lattice, *Letters in Mathematical Physics*, **3**, 425-429, (1979).

[101] H. G. Meijer, A short history of orthogonal polynomials in a Sobolev space I. The non-discrete case, *Nieuw Archief voor Wiskunde (4)*, **14**, 93-112, (1996).

[102] L. Miranian, Matrix valued orthogonal polynomials on the real line: some extensions of the classical theory, *Journal of Physics A: Mathematical and General*, **38**, 5731-49, (2005).

[103] M. Mulase, Complete integrability of the Kadomtsev–Petviashvili equation, *Advances in Mathematics*, **54**, 57-66, (1984).

[104] P. G. Nevai, Orthogonal polynomials, *Memoirs American Mathematical Society*, **213**, Providence RI, (1979).

[105] E. M. Nikishin and V. N. Sorokin, Rational Approximations and Orthogonality, *Translations of Mathematical Monographs 92, American Mathematical Society*, Providence, RI, (1991).

[106] P. J. Olver, On multivariate interpolation, *Studies in Applied Mathematics*, **116**, 201-240, (2006).

[107] F. Riesz, Sur les opérations fonctionnelles linéaires, *Comptes Rendus de l'Académie des Sciences* , **149**, 974-977, (1909).

[108] M. Sato, Soliton equations as dynamical systems on infinite dimensional Grassmann manifolds (Random Systems and Dynamical Systems), *RIMS Kokyuroku*, **439**, 30-46 (1981).

[109] M. Sato and Y. Sato, Soliton equations as dynamical systems on infinite-dimensional Grassmann manifold, *Nonlinear partial differential equations in applied science, Mathematical Studies*, **81**, 259-271, (Tokyo, 1982), (Amsterdam, 1983).

[110] F. W. Schäfke, Zu den Orthogonalpolynomen von Althammer, *Journal für die Reine und Angewandte Mathematik*, **252**, 195-199, (1972).

[111] F. W. Schäfke and G. Wolf, Einfache verallgemeinerte klassische Orthogonal polynome, *Journal für die Reine und Angewandte Mathematik*, **262/263**, 339-355, (1973).

[112] J. Shohat, Théorie générale des polynômes orthogonaux de Tchebichef, *Mémorial des Sciences Mathématiques*, **66**, 1-69, (1934).

[113] B. Simon, Orthogonal Polynomials on the unit circle. Part 1, Vol. 54, *American Mathematical Society Colloquium series* Providence RI, (2005).

[114] B. Simon, Orthogonal Polynomials on the unit circle. Part 2, Vol. 54, *American Mathematical Society Colloquium series* Providence RI, (2005).

[115] V. N. Sorokin, On simultaneous approximation of several linear forms, *Vestnik Moskovskogo Universiteta. Seriya I. Matematika, Mekhanika*, **1**, 44-47, (1983).

[116] T. J. Stieltjes, Recherches sur les fractions continues, *Annales de la Faculté de Sciences de Toulouse*, **8**, (1894).

[117] P. K. Suetin, Orthogonal Polynomials in Two Variables, *Analytical Methods and Special Functions*, **3**, Gordon and Breach Science Publishers, Amsterdam, (1999).

[118] G. Szegő, Ein Grenzwertsatz über die Toeplitzschen Determinanten einer reellen positiven Funktion, *Mathematische Annalen*, **76**, 490-503, (1915).

[119] G. Szegő, Orthogonal Polynomials, *American Mathematical Society Colloquium Publications*, **23**, Providence RI, (1939) (4th edition 1975).

[120] O. Toeplitz, Zur Transformation der Scharen bilinearer Formen von unendlichvielen Veränderlichen, Nachr. der Kgl. Gessellschaft der Wissenschaften zu Göttingen, *Mathematischphysikalische Klasse*, 110-115, (1907).

[121] O. Toeplitz, Zur Theorie der quadratischen und bilinearen Formen von unendlichvielen Veränderlichen. I. Teil: Theorie der L-Formen, *Mathematische Annalen*, **70**, 351-376, (1911).

[122] M. Toda, Vibration of a chain with nonlinear interaction, *Journal of the Physical Society of Japan*, **22**, 431-436, (1967).

[123] F. Tricomi, Equazioni differentiali, Torino, (1948).

[124] V. B. Uvarov, The connection between systems of polynomials that are orthogonal with respect to different distribution function, *Computational Mathematics and Mathematical Physics USSR*, **9**, 25-36, (1969).

[125] W. Van Assche, J.S. Geronimo and A.B.J. Kuijlaars, Riemann-Hilbert problems for multiple orthogonal polynomials, in: "Special Functions 2000: Current Perspectives and Future Directions" (J. Bustoz et al., eds.), Kluwer, Dordrecht, 23-59, (2001).

[126] S. Verblunsky, On Positive Harmonic Functions: A Contribution to the Algebra of Fourier Series, *Proceedings of the London Mathematical Society*, **1**, 290-320, (1935).

[127] S. Verblunsky, On Positive Harmonic Functions, *Proceedings of the London Mathematical Society*, **1**, 125-157, (1936).

[128] N. J. Zabusky and M. D. Kruskal, Interaction of "solitons" in a collisionless plasma and initial states, *Physical Review Letters*, **15**, 240-243, (1965).

[129] V. E. Zakharov and A. B. Shabat, Exact theory of two-dimensional self-focusing and one-dimensional self-modulation of waves in nonlinear media, *Soviet Physics Uspekhi*, JETP **34**, 62-69, (1972).

[130] F. Zhang, The Schur Complement and its Applications, Springer, New York, (2005).

[131] A. Zhedanov, Rational spectral transformations and orthogonal polynomials, *Journal of Computational and Applied Mathematics*, **85**, 67-86, (1997).