Generalized Nevanlinna-Pick interpolation on the boundary. Application to impedance matching

Laurent Baratchart, Martine Olivi, Fabien Seyfert

Generalized Nevanlinna-Pick interpolation on the boundary. Application to impedance matching.

L. Baratchart, M. Olivi, F. Seyfert

December 31, 2015

1 Introduction

sec:intro

Nevanlinna-Pick interpolation is a classical topic from function theory on the circle that has undergone several generalizations and applications, in particular to control. A nice account of these may be found in [1]. In later years, the topic has also become insrumental in design of frequency devices, controlling the degree of a rational interpolant and imposing peak points for the modulus on the imaginary axis, see [10, 6].

BallHorst

Georgiou1987,BGL2001

In this work, we generalize this kind of interpolation further, by consider interpolation *on* the imaginary axis while imposing peak points for the modulus and controlling the degree. Our motivation comes from the broadband matching problem and described in details below.

Communication devices such as multiplexers, routers, power dividers, couplers or antenna receptor chains, are usually realized by connecting to each other elementary components of which filters and N-port junctions are the most common. Multiplexers, for example, are realized by plugging $N-1$ filters (one per channel) to a N-port junction. In fact, filters should be considered as typical elementary two-port components, to be present in almost every telecommunication device.

Now, when connecting a filter to some existing system, a recurring issue is to determine which frequencies will carry energy to the system across the filter, and which frequencies will bounce back. In this respect, the system L shown in Figure 1 (to be seen as the

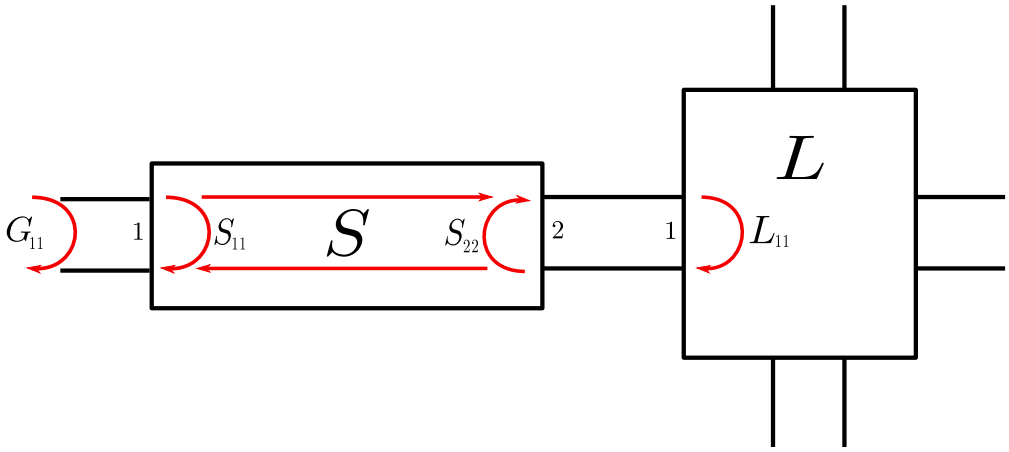

Figure 1: Filter plugged to a load L with reflexion coefficient L_{11} MatchTwoPort

load of the filter S) is characterized by its reflection coefficient L_{11}, which is a complex-valued function of the frequency ω as the latter ranges over real numbers. Hereafter, we abbreviate real and complex numbers by $\mathbb{R}$ and $\mathbb{C}$, respectively. As to the filter, its effect is described by a 2×2 scattering matrix, still denoted by S, whose entries are again $\mathbb{C}$-valued functions of $\omega\in\mathbb{R}$. If the filter is lossless, meaning that S is unitary at all frequencies:

$$S(\omega)^* S(\omega) = Id, \quad \omega\in\mathbb{R} \tag{1}$$

(here and below superscript "*" stands for "transpose-conjugate"), the reflection coefficient G_{11} of the overall system is easily computed at the frequency ω to be

$$\begin{aligned} G_{11}(\omega) &= S_{11}(\omega) + \frac{S_{12}(\omega)S_{21}(\omega)L_{11}(\omega)}{1-S_{22}(\omega)L_{11}(\omega)} \\ &= \frac{S_{11}(\omega) - L_{11}(\omega)\det(S(\omega))}{1-S_{22}(\omega)L_{11}(\omega)} \\ &= \det(S(\omega))\frac{\overline{S_{22}(\omega)} - L_{11}(\omega)}{1-S_{22}(\omega)L_{11}(\omega)}. \end{aligned} \tag{2}$$

A matching frequency is, by definition, a frequency ω for which $G_{11}(\omega)=0$. If $|L_{11}(\omega)|<1$, hence also $|S_{22}(\omega)L_{11}(\omega)|<1$, this reduces by (2) to the relation

$$S_{22}(\omega) = \overline{L_{11}(\omega)}. \tag{3}$$ defmatch

In contrast, a stopping frequency ω is defined by the property that $|G_{11}(\omega)|=1$. If $|L_{11}(\omega)|<1$, this amounts by (2) to $|S_{22}(\omega)|=1$ which is in turn equivalent by (1) to

$$S_{12}(\omega) = S_{21}(\omega) = 0. \tag{4}$$ stopfreq

The problem of synthesizing the filter S, or the matching network L, in such a way that G_{11} is lowest possible on a given frequency band is a very old one. When the filter is considered to be finite-dimensional, this issue gave rise to the so-called matching theory of Fano and Youla [7]. If the model for the load is likewise rational with 2 ports, this theory provides one with a parametrization of all global responses G associated with systems that can be realized by plugging a variable filter S of given degree to a given load L. Such G are characterized in terms of their transmission zeros, which account for the fact that the load L can be "extracted" from the overall response G. However, it is still unknown how to deduce matching filtering characteristics from this parametrization as soon as the load has degree greater that one. This may contribute to explain why the Fano-Youla theory has had little impact in practice. Also, the need to derive a rational model for the load, and to infer its transmission zeros, might have impeded its dissemination in the engineering community. Instead, system manufacturers often use brute force optimization with its usual drawbacks and uncertainties. Another approach was proposed by J. Helton [12] in an infinite dimensional setting, where the matching problem gets reformulated as a H^∞ approximation problem of Nehari type, the solution to which is elegantly produced in terms of the norm and maximizing vectors of a Hankel operator. This technique amounts to convexifying the problem, and it yields hard bounds on the achievable matching error no matter the degree of the filter, along with an optimal (non-rational) solution to match this bound. However, the "infinite degree" of this optimal filter makes it hardly realizable or even computable in practice. In the present paper, we propose an intermediate approach where a finite-dimensional filter response of prescribed degree is being synthesized by imposing matching and stopping frequencies with respect to some general frequency varying load.

2 An interpolation problem

Below, we shall regard the scattering matrix of a filter as a function of a *real* variable, namely the frequency. This differs from the somewhat more usual convention where the transfer function is defined on the imaginary axis rather than on the real line. In the present framework, when the filter is finite dimensional and stable, its scattering matrix is rational with poles in the open upper half-plane $\mathbb{C}^+$, and the degree of the numerator of each entry does not exceed the degree of the denominator; equivalently, the scattering matrix belongs to the Hardy space $H^\infty(\mathbb{C}^-)$ of bounded holomorphic functions in the open lower half-plane $\mathbb{C}^-$. A polynomial with no root in $\mathbb{C}^-$ is said to be *stable in the broad sense*. A polynomial is called *stable* if it has no root in $\overline{\mathbb{C}^-}$, the *closed* lower half-plane.

The scattering matrix S of a lossless filter is in fact *inner* in $H^\infty(\mathbb{C}^-)$, meaning that it satisfies (1). By the maximum principle, this entails that S is contractive in $\mathbb{C}^-$:

$$\|S(z)\xi\| \leq \|\xi\|, \quad \forall z \in \mathbb{C}^-,\ \xi \in \mathbb{C}^2,$$

where "$\|.\|$" indicates the Euclidean norm and the inequality is strict unless $S(z)\xi$ is constant.

We denote by $\mathbb{D}$ the open unit disk. For any complex matrix M, M^T denotes its transpose. For a rational matrix function $F(s)$, we define its *paraconjugate* $F^*(s)$ by

$$F^*(s) = F(\bar{s})^*, \quad s \in \mathbb{C}.$$

When F is constant, this notation agrees with the previously introduced M^* for the transpose conjugate of a complex matrix M. Note that $F^*(s)$ indeed takes transpose conjugate values of F on $\mathbb{R}$, and clearly $(F^*)^* = F$. If p is a polynomial, then its paraconjugate p^* is the polynomial obtained by conjugating the coefficients, in particular it has the same degree as p and roots conjugate to those of p.

KFA69,Fuhrmann
Recall (see *e.g.* [13, 8]) that the McMillan degree of a $\ell_1 \times \ell_2$ rational matrix R is the smallest non-negative integer ℓ for which one can write $R(s) = C(sI_\ell - A)^{-1}B + D$, where A, B, C and D are complex matrices of size $\ell \times \ell$, $\ell \times \ell_2$, $\ell_1 \times \ell$ and $\ell_1 \times \ell_2$ respectively, and I_ℓ is the identity matrix of size $\ell \times \ell$. Equivalently, the McMillan degree is the smallest degree possible for the determinant of an invertible polynomial matrix P such that PR is also a polynomial matrix. A function of the form q^*/q where q is a stable polynomial of degree d is called a *Blaschke product* of degree d. When a rational matrix R is inner, then its determinant is a Blaschke product whose degree is equal to the McMillan degree of R [3].
BU98

Each inner rational 2×2 matrix S of McMillan degree N such that $\lim_{s\to\infty} S(s) = I_2$ admits the following representation (Belevitch form [5]):
Belevitch56

$$S = \frac{1}{q}\begin{bmatrix} p^* & -r \\ r^* & p \end{bmatrix}, \tag{5}$$ Belevitch

where p, q are monic complex polynomials of degree N, r is a complex polynomial of degree at most $N-1$, the polynomials p, r have no common real zero, and q is computed from p and r to be the unique monic stable polynomial satisfying the Feldtkeller equation:

$$qq^* = pp^* + rr^*. \tag{6}$$ feld

Equation (6) expresses that q is a stable *spectral factor* of $pp^* + rr^*$, see Proposition 3.1 for more details about existence and uniqueness of q.

By definition (3), a set $\{x_1 \dots x_m\} \subset \mathbb{R}$ consists of matching frequencies for the filter (5) with respect to a load having reflection coefficient L_{11} if, and only if

$$\frac{p}{q}(x_k) = \overline{L_{11}(x_k)} \stackrel{def}{=} \gamma_k, \qquad 1 \leq k \leq m. \tag{7}$$

We will suppose that $|\gamma_k| < 1$, as the matching problem at fully reflecting frequencies for the load is not well defined (expression (2) is either of modulus 1 or indeterminate of the form $0/0$ when $|L_{11}(\omega)| = 1$).

Besides, it may be desirable for various purposes (*e.g.* preventing oscillations of the response or accounting for losses) to meet additional interpolation conditions inside the stability domain $\mathbb{C}^-$. To accomodate this case as well, we consider points $\{z_1 \dots z_l\}$ in $\mathbb{C}^-$ where following "complex" matching condition should hold:

$$\frac{p}{q}(z_k) = \overline{L_{11}(\Re(z_k))} \stackrel{def}{=} \beta_k, \tag{8}$$

where "$\Re$" indicates the real part.

Let us abbreviate $(z_1, \cdots, z_l)$ and $(\beta_1, \cdots, \beta_l)$ by Z and β respectively. We define $P(Z,\beta)$ to be the so-called Pick matrix associated with the interpolation data (z_k, β_k), namely the Hermitian $l \times l$ matrix defined by:

$$P_{k,j}(Z,\beta) = \frac{1 - \beta_k \overline{\beta_j}}{i(z_k - \overline{z_j})}. \tag{9}$$ Pickm

It is a classical fact (see [9, Ch. I, Cor. 2.3] for a version on the disk rather than the half-plane) that $P(Z,\beta)$ is positive semi-definite if and only if there is an analytic function f in $\mathbb{C}^-$ such that $|f(\xi)| \leq 1$ for all $\xi \in \mathbb{C}^-$ and f meets the interpolation conditions $f(z_k) = \beta_k$ for $1 \leq k \leq l$. When this holds, $P(Z,\beta)$ is actually positive definite unless the solution to this constrained interpolation problem is unique, in which case the unique solution is a Blaschke product of degree equal to the rank of $P(Z,\beta)$. Conversely, if there is a solution to the interpolation problem which is a Blaschke product of degree $\delta < l$ then $P(Z,\beta)$ has rank δ. Note that the positive definiteness of $P(Z,\beta)$ is equivalent to the existence of a solution to the corresponding interpolation problem whose trace on $\mathbb{R}$ has modulus strictly less than 1 on a set of positive measure. Indeed, if two distinct solutions have modulus 1 a.e. on $\mathbb{R}$, then any convex combination is another solution asuming modulus strictly less than 1 at every point where the two initial solutions take on distinct values.

We call $\mathbb{P}_Z^+$ the set of those $\beta \in \mathbb{C}^l$ such that $P(Z,\beta)$ is strictly positive. Clearly $\mathbb{P}_Z^+$ is open in $\mathbb{C}^l$. Moreover, the above-mentioned interpretation of positivity in terms of solutions to a constrained interpolation problem shows immediately that $\mathbb{P}_Z^+$ is convex, in particular it is connected. For simplicity, we sometimes drop the subscript Z when the interpolation points are understood, and write $\mathbb{P}^+$ instead $\mathbb{P}_Z^+$.

Finally, suppose that we are given a set of $N-1$ stopping frequencies (possibly lying at ∞), which are distinct from the x_k. In view of (4) and (5), this prescribes the roots of the transmission polynomial r of the filter (a zero at infinity means a drop in degree). We shall first study the situation where the leading coefficient of r is also prescribed, thus r itself is fixed. Then, we raise the following matching problem.

Problem $\mathscr{P}$: *Given* m distinct real frequencies $(x_1, x_2 \dots x_m)$, m interpolation conditions $(\gamma_1, \gamma_2 \dots \gamma_m)$ in $\mathbb{D}^m$, l distinct "complex frequencies" $(z_1, z_2 \dots z_l) \in (\mathbb{C}^-)^l$ associated to l interpolation values $(\beta_1, \beta_2 \dots, \beta_l) \in \mathbb{P}^+$ and $r \neq 0$ a complex polynomial of degree at most $m+l-1$, such that $r(x_k) \neq 0$, $k = 1, \dots, m$, *to find* (p,q) a pair of monic complex polynomials of degree $N = m+l$ such that,

$$\left\{ \begin{array}{rcll} \frac{p}{q}(x_k) & = & \gamma_k, & \text{for } k = 1,..,m \\ \frac{p}{q}(z_k) & = & \beta_k, & \text{for } k = 1,..,l \\ qq^* - pp^* & = & rr^* & \end{array} \right. \tag{10}$$ IntCond

and q has no root in the open lower half-plane $\mathbb{C}^-$ (*i.e.* is stable in the broad sense).

Nothing in the formulation of Problem $\mathscr{P}$ prevents the denominator polynomial q from vanishing at real points. Observe, however, that a real zero of q is a zero of both p and r with the same multiplicity (since $|p|^2 + |r|^2 = |q|^2$ on $\mathbb{R}$ by (10)), and that in this case the McMillan degree of the filter S will drop to some $d < N$, *cf* (5). Note also that, by the very assumption in Problem $\mathscr{P}$, a common zero to p and r cannot be one of the x_k, so that $\frac{p}{q}(x_k)$ in (10) is really equal to $p(x_k)/q(x_k)$.

Problem $\mathscr{P}$ may be viewed as an extension of the Nevanlinna-Pick interpolation problem with degree constraint solved in [10]. The latter is relevant to several applications such as sensitivity minimization or spectral estimation (see [6]). The extension presented here with Problem $\mathscr{P}$ consists in allowing the interpolation points to lie *on* the real axis, whereas in [10] they are confined to the stability domain $\mathbb{C}^-$. This difference is *crucial* to approach the matching problem with such techniques, which was the main incentive for the authors to carry out the present study. It generalizes the work presented in [4], allowing interpolation points to lie in the analyticity domain as well as on the boundary and the polynomial r to have real roots, except at the interpolation points, i.e. the x_i's. Note that the roots of r are then prescribed peak points for the modulus of

p/q, *i.e.* points where the maximum value 1 is attained.

3 A matching theorem.

secmatch

The analysis of problem $\mathscr{P}$ relies on the study of a specific evaluation map to be defined presently. According to the statement of the problem, we fix $(x_1, x_2 \dots x_m)^T \in \mathbb{R}^m$, $(z_1, z_2 \dots z_l)^T \in (\mathbb{C}^-)^l$, and a polynomial r of degree at most $m+l-1$ such that $r(x_k) \neq 0$ for all k. We let $\mathbf{PM}_N$ designate the set of monic polynomial of degree $N = m+l$ with complex coefficients. This set is topologized as $\mathbb{C}^N \sim \mathbb{R}^{2N}$, using coefficients as coordinates except for the leading one which is equal to 1 by definition. Specifically, if $p \in \mathbf{PM}_N$ and we write $p(z) = z^N + p_{N-1}z^{N-1} + \cdots + p_0$, then we identify p with the vector $(p_0, p_1, \cdots, p_{N-1})^T \in \mathbb{C}^N$.

Equation (6) associates to each $p \in \mathbf{PM}_N$ a unique polynomial $q = q(p) \in \mathbf{PM}_N$ which is stable in the broad sense, *cf* Proposition 3.1 to come. Since $|p|^2 \leq |p|^2 + |r|^2 = |q|^2$ on $\mathbb{R}$, the rational function p/q has no real pole and no pole in $\mathbb{C}^-$ either since q is stable in the broad sense. Thus, by the maximum principle, we conclude that $|p/q| \leq 1$ on $\overline{\mathbb{C}^-}$. In addition, since no x_k is a root of r by assumption, we have that $|p(x_k)/q(x_k)| < 1$ hence p/q is not a Blaschke product. Therefore the Pick matrix associated with the interpolation data $(z_k, p(z_k)/q(z_k))$ must be positive definite, and we can define an evaluation map $\psi : \mathbf{PM}_N \to \mathbb{D}^m \times \mathbb{P}_Z^+$ by the formula

$$\psi(p) = \begin{pmatrix} p(x_1)/q(x_1) \\ \vdots \\ p(x_m)/q(x_m) \\ p(z_1)/q(z_1) \\ \vdots \\ p(z_l)/q(z_l) \end{pmatrix} \in \mathbb{D}^m \times \mathbb{P}_Z^+. \tag{11}$$

psi

The main result of the paper may now be stated as follows.

thmain

Theorem 3.1 *ψ is a homeomorphism from $\mathbf{PM}_N$ onto the product space $\mathbb{D}^m \times \mathbb{P}^+$. If $p \in \mathbf{PM}_N$ has no real root in common with r, then ψ is continuously differentiable in a neighborhood of p with invertible derivative, hence ψ is a local diffeomorphism at p.*

The proof of Theorem 3.1 will be given in sections 3.5 and 3.6, after some preliminaries.

3.1 Continuity and differentiability of ψ.

For $k \geq 0$ an integer, we let $\mathbf{P}_k$ be the space of complex polynomials of degree at most k and $\mathbf{PE}_k$ the subset comprising polynomials of exact degree k. We occasionally write $\mathbf{P}_{\mathbb{R},k}$ for the real subspace of polynomials with real coefficients. The space $\mathbf{P}_k$ identifies with $\mathbb{C}^{k+1} \sim \mathbb{R}^{2k+2}$, using coefficients as coordinates; that is, $(p_0, p_1, \cdots, p_k)^T \in \mathbb{C}^{k+1}$ is regarded as the polynomial $p(z) = p_k z^k + p_{k-1} z^{k-1} + \cdots + p_0$. With this definition, $\mathbf{PM}_N \subset \mathbf{P}_N$ is the hyperplane $\{p_N = 1\}$ which in turn identifies with $\mathbb{C}^N$ as pointed out earlier. We further denote by $\mathbf{SB}_N$ the set of polynomials of degree at most N which are stable in the broad sense, and by $\mathbf{SBM}_N$ the subset of monic polynomials of degree N stable in the broad sense. The set of stable polynomials of degree at most N will likewise be denoted by $\mathbf{S}_N$, the subset of stable polynomial of exact degree N by $\mathbf{SE}_N$, and the subset of stable monic polynomials of degree N by $\mathbf{SM}_N$.

We write $\mathbf{P}_{2N}^+$ for the set of polynomials of degree at most $2N$ which are non-negative on $\mathbb{R}$. Such a polynomial must have real coefficients, even degree, positive dominant coefficient, and its real roots have even multiplicity. We put $\mathbf{PE}_{2N}^+$ for the subset of non-negative real polynomials of exact degree $2N$ and $\mathbf{PM}_{2N}^+$ for the subset of non-negative monic real polynomials of exact degree $2N$. The sets $\mathbf{P}_{2N}^+$ and $\mathbf{PE}_{2N}^+$ will be regarded as embedded in $\mathbb{R}^{2N+1}$, and again $\mathbf{PM}_{2N}^+$ will be seen as a subset of $\mathbb{R}^{2N}$ for it is the intersection of $\mathbf{P}_{2N}^+ \subset \mathbb{R}^{2N+1}$ with the hyperplane $\{p_{2N+1} = 1\}$.

The interior $\overset{\circ}{\mathbf{PM}}{}_{2N}^+$ of $\mathbf{PM}_{2N}^+ \subset \mathbb{R}^{2N}$ consists of real monic polynomials of degree $2N$ which are strictly positive on $\mathbb{R}$. Indeed, if $p \in \mathbf{PM}_{2N}^+$ is such that $p(x_0) = 0$, then adding a small negative constant to p will destroy positivity at x_0 and therefore p cannot lie interior to $\mathbf{PM}_{2N}^+$ in $\mathbb{R}^{2N}$. Conversely, let $p \in \mathbf{PM}_{2N}^+$ have no zero on $\mathbb{R}$. Then, there is $\varepsilon > 0$ such that $|p(x)| > \varepsilon$ for $x \in \mathbb{R}$. Let us write $p(x) = x^{2N+1} + p_{2N} x^{2N} + \cdots + p_0$ and put $a := \max\{1, \varepsilon + 2\Sigma_{j=0}^{2N} |p_j|\}$. If we let $(\delta_0, \cdots, \delta_{2N-1})^T \in \mathbb{R}^{2N}$ be such that $\Sigma |\delta_j| a^j < \varepsilon/2$, we get upon setting $\delta p(x) := \Sigma_{j=0}^{2N-1} \delta_j x^j$ that $|p + \delta p| > \varepsilon/2$ on $[-a, a]$ and that

$$|p(x) + \delta p(x)| \geq |x|^{2N} \left(|x| - \varepsilon/2 - \Sigma_{j=1}^{2N-1} |p_j| \right) > \frac{|x|^{2N+1}}{2}, \qquad |x| > a.$$

Hence p lies interior to $\mathbf{PM}_{2N}^+$. Likewise, the interior $\overset{\circ}{\mathbf{PE}}{}_{2N}^+$ of $\mathbf{PE}_{2N}^+ \subset \mathbb{R}^{2N+1}$ consists of real polynomials of exact degree $2N$ which are strictly positive on $\mathbb{R}$.

A completely similar argument shows that the interior of $\mathbf{SBM}_N$ is $\mathbf{SM}_N$.

After these preliminaries, we are in position to prove our first result:

sfn **Proposition 3.1** *To any non zero $P \in \mathbf{P}_{2N}^+$, one can associate $q \in \mathbf{SB}_N$ such that*

$$P(t) = |q(t)|^2 = q(t)q^*(t), \quad t \in \mathbb{R}. \tag{12}$$ SpectralFactor

The polynomial $q(s)$ is unique up to a multiplicative unimodular constant, and if P has exact degree $2N$ then q has exact degree N. If we define two maps (corresponding to two different normalizations) by:

a) $\varphi : \mathbf{P}_{2N}^+ \setminus \{0\} \mapsto \mathbf{SB}_N$, with $\varphi(P)$ the unique solution to (12) meeting $q(-i) > 0$,

b) $\varphi_N : \mathbf{PM}_{2N}^+ \mapsto \mathbf{SBM}_N$ with $\varphi_N(P)$ the unique monic solution to (12),

then φ, φ_N are continuous. Also, the restriction of φ (resp. φ_N) to $\overset{\circ}{\mathbf{PE}}{}_{2N}^+$ (resp. $\overset{\circ}{\mathbf{PM}}{}_{2N}^+$) is continuously differentiable with invertible derivative at every point. Specifically:

- *if $P \in \overset{\circ}{\mathbf{PE}}{}_{2N}^+$ and δP is a real polynomial of degree at most $2N$, then*

  $$D\varphi(P)[\delta P] = u$$

  where u is the unique polynomial such that

  $$u^*\varphi(P) + u\varphi^*(P) = \delta P, \qquad u \in \mathbf{P}_N, \quad u(-i) \in \mathbb{R}; \tag{13}$$ derphi

- *if $P \in \overset{\circ}{\mathbf{PM}}{}_{2N}^+$ and δP is a real polynomial of degree at most $2N-1$, then*

  $$D\varphi_N(P)[\delta P] = u$$

  where u is the unique polynomial such that

  $$u^*\varphi_N(P) + u\varphi_N^*(P) = \delta P, \qquad u \in \mathbf{P}_{N-1}. \tag{14}$$ derphiN

Proof. It is elementary to check that a polynomial q satisfies (12) and is stable in the broad sense if and only if its roots are the real roots of P with half their multiplicity and the non-real roots of P having strictly positive imaginary part with their multiplicity,

while its dominant coefficient has square modulus equal to the dominant coefficient of P. This shows the existence of q and its uniqueness up to a multiplicative unimodular constant. Alternatively, the result also follows upon applying to $P(i(e^{i\theta}+1)/e^{i\theta}-1))$ a classical result by Fejer and Riesz asserting that non-negative trigonometric polynomials are square moduli of algebraic polynomials on the unit circle [16, sec. 53].

Next, we prove that φ is continuous. Let (P_k) be a sequence in $\mathbf{P}^+_{2N}\setminus\{0\}$ converging to some $P\in\mathbf{P}^+_{2N}\setminus\{0\}$. We must show that $(q_k)=(\varphi(P_k))$ converges to $\varphi(P)$. As a basis of $\mathbf{P}_N$, we pick the Lagrange interpolation polynomials L_n, $n=0,1\ldots,N$, associated with the integer points $x=0,1,\ldots,N$. In other words, for each $n\in\{0,\cdots,N\}$, we have when $0\le j\le N$ that $L_n(j)=\delta_{n,j}$, the Kronecker delta function. The coordinates of q_k in this basis are $(q_k(0),q_k(1),\ldots,q_k(N))$. Since $|q_k(j)|=\sqrt{P_k(j)}$ and (P_k) is a fortiori bounded in $\mathbb{R}^{2N+1}$, the sequence (q_k) is in turn bounded. We may thus extract a convergent sub-sequence from any subsequence, and *we claim* that the limit is none but $\varphi(P)$; this will prove the announced continuity. Assume indeed that q_k converges to q. Since taking products and conjugates of polynomials are continuous operations $\mathbf{P}_N\times\mathbf{P}_N\to\mathbf{P}_{2N}$ and $\mathbf{P}_N\to\mathbf{P}_N$ respectively, we get in the limit from the relation $P_k=q_kq_k^*$ that $P=qq^*$. Moreover $q(-i)\ge 0$ because evaluation at a point is also continuous. Thus, in order to prove the claim, it only remains to show that $q\in\mathbf{SB}_N$. Proceeding by using reductio ad absurdum, we suppose that q has an unstable root, say $x_0\in\mathbb{C}^-$. As q is not identically zero, x_0 is an isolated root, and there exists $R>0$ such that the disk $D=\{x,|x-x_0|\le R\}$ is included in C^- and such that $q(x)\neq 0$ on $D's$ boundary. As the sequence (q_k) converges uniformly to q on every compact subset of $\mathbb{C}$, for k_0 large enough the inequality $|q(x)-q_{k_0}(x)|\le|q(x)|$ holds on $D's$ boundary. By Rouché's theorem this yields that the stable polynomial q_{k_0} has a root in D, and hence in $\mathbb{C}^-$. Hence a contradiction.

Consider now the map $\tilde\varphi:\mathbf{P}_N\to\mathbf{P}^+_{2N}$ defined by $\tilde\varphi(q)=qq^*$. It is continuously differentiable, and its differential $D\tilde\varphi(q)$ at q acts on $dq\in\mathbf{P}_N$ by the formula

$$D\tilde\varphi(q)[dq] \;=\; dqq^*+qdq^*. \tag{15}$$ diffspec

If $P\in\overset{\circ}{\mathbf{PE}}{}^+_{2N}$, then $q=\varphi(P)$ lies in $\mathbf{SE}_N\cap\mathfrak{V}$, where $\mathfrak{V}$ is the real subspace of $\mathbf{P}_N$ consisting of polynomials whose value at $-i$ is real. In fact, the restriction $\tilde\varphi_1$ of $\tilde\varphi$ to $\mathbf{SE}_N\cap\mathfrak{V}$ is inverse to the restriction of φ to $\overset{\circ}{\mathbf{PE}}{}^+_{2N}$. Note that $\mathbf{SE}_N\cap\mathfrak{V}$ is an open subset of the linear manifold $\mathbf{P}_N\cap\mathfrak{V}$ and that $\tilde\varphi_1$ is differentiable on this open set with derivative given by (15) restricted to $dq\in\mathbf{P}_N\cap\mathfrak{V}$ which is the tangent space. *We claim* that this derivative is injective. Assume indeed that $D\tilde\varphi_1(q)[dq]=0$. Then, since q and q^* are coprime polynomials (for their roots respectively lie in $\mathbb{C}^+$ and $\mathbb{C}^-$), we get from (15) that q divides dq so that $dq=\lambda\,q$ for some $\lambda\in\mathbb{C}$, because the degree of dq cannot exceed

N which is the degree of q. In view of (15), we conclude from $D\tilde{\varphi}_1(q)[dq] = 0$ that $(\lambda + \bar{\lambda})qq^* = 0$, and since $qq^* \neq 0$ (for it has exact degree $2N$) we see that λ is pure imaginary. As $q(-i) > 0$, this implies that $dq(-i) = \lambda q(-i)$ is pure imaginary, and since it is also real because $dq \in \mathbf{PE}_N \cap \mathfrak{V}$ by assumption, we necessarily get $\lambda = 0$ whence $dq = 0$. *This proves the claim.* Because $D\tilde{\varphi}_1(q)$ maps $\mathbf{P}_N \cap \mathfrak{V}$ into reals polynomials of degree at most $2N$ (*cf.* (15)) and both spaces have real dimension $2N+1$, we conclude since it is injective that it is in fact invertible. Now, the inverse function theorem asserts that $\tilde{\varphi}$ is locally invertible and that its inverse φ is continuously differentiable on $\overset{\circ}{\mathbf{PM}}{}^+_{2N}$, as announced. Moreover, the derivative of the inverse is the inverse of the derivative, so that (13) is clear from (15). This concludes the proof in case a.
Case b can be handled in a similar but simpler way for dq in (15) will have degree at most $N-1$, making it obvious that it must vanish if it is divisible by q. □

Remarks. Continuity of spectral factorization can be given other, more analytic proofs based on the Poisson representation of log-moduli of outer functions, see *e.g.* [2, Lemma 1] for an alternative argument on the disk that easily carries over to the half-plane. BCQ15

Keeping in mind notation from Proposition 3.1, we may now represent the map ψ introduced in (11) as the composition of two functions, namely the map from $\mathbf{PM}_N$ into $\mathbf{PM}_N \times \mathbf{SBM}_N$ given by

$$p \to (p, \varphi_N(pp^* + rr^*))$$

followed by the evaluation map

$$(p,q) \to \left(\frac{p}{q}(x_1), \ldots, \frac{p}{q}(x_m), \frac{p}{q}(z_1), \ldots, \frac{p}{q}(z_l)\right)^T$$

from $\mathbf{PM}_N \times \mathbf{SBM}_N$ into $\mathbb{D}^m \times \mathbb{P}^+_Z$. Thus, the proposition immediately yields:

ContDiff **Corollary 3.1** *The map ψ is continuous at every $p \in \mathbf{PM}_N$ and if p has no real common root with r, then ψ is differentiable at p.*

3.2 An excursion into positive real functions

A holomorphic function f on $\mathbb{C}^-$ is called a *Schur* function if $|f(\zeta)| \leq 1$ for every $\zeta \in \mathbb{C}^-$; it is called a *Carathéodory* function if $\mathrm{Re} f(\zeta) \geq 0$ for every $\zeta \in \mathbb{C}^-$. Note that Carathéodory functions may well have poles on $\mathbb{R}$, contrary to Schur functions.

If (p,q) is a solution to Problem $\mathscr{P}$, then p/q is a Schur function as explained in the paragraph above (11). Our proof that ψ is a homeomorphism rests on a fundamental link, to be stressed in this section, between problem $\mathscr{P}$ and its analog in terms of Carathéodory functions. Below, we establish properties of this Carathéodory version that will be of use to demonstrate theorem 3.1.

For $p \in \mathbf{PM}_N$ and $q = \varphi_N(pp^* + rr^*)$, we put $\Sigma = \Sigma(p) := p/q$.
By construction, this is a Schur rational function satisfying

$$1 - \Sigma^*\Sigma = \frac{rr^*}{qq^*}. \tag{16}$$

Using the conformal map $z \mapsto (1-z)/(1+z)$ from $\mathbb{D}$ onto the right half-plane $\mathbb{C}^+$, we define a Carathéodory function Y (the so-called Cayley transform of Σ) by the formula

$$Y := \frac{1-\Sigma}{1+\Sigma} = \frac{q-p}{q+p}, \tag{17}$$ CT

and a straightforward computation shows that

$$Y + Y^* = \frac{(q-p)(q^*+p^*) + (q^*-p^*)(q+p)}{(q+p)(q^*+p^*)} = \frac{2rr^*}{(q+p)(q^*+p^*)}. \tag{18}$$ lienpos

By definition, the dissipation polynomial of a rational Carathéodory function is the numerator of the fraction $Y + Y^*$ when the latter is written in irreducible form. To us, given a rational Carathéodory function π/χ with π, χ polynomials, it is more convenient to define the *dissipation polynomial of the pair* (π,χ) to be the polynomial $\pi\chi^* + \pi^*\chi$. Thus, by (18), $2rr^*$ is the dissipation polynomial of the pair $(q-p, q+p)$.

In view of (17)-(18), Problem $\mathscr{P}$ is equivalent to an interpolation problem for rational Carathédory functions of the form π/χ where $\pi \in \mathbf{P}_{N-1}$ and $\chi \in \mathbf{SBM}_N$, with prescribed dissipation polynomial rr^* for the pair (π,χ). For this equivalent problem, the analog of equation (6) is

$$\pi\chi^* + \pi^*\chi = rr^*. \tag{19}$$ specy

If r has no real root and (π,χ) is a solution to (19) then χ is stable, *i.e.* it lies in $\mathbf{SM}_N$ and not just in $\mathbf{SBM}_N$. This entails that χ, χ^* are coprime so that π is uniquely determined by r and χ through the linear equation (19). In this case the Carathéodory analog of Problem $\mathscr{P}$ is easier to study than $\mathscr{P}$ itself. The situation changes when r has a real root, say x_0. For if χ also has a root at x_0 and if π is a solution to (19) such that π/χ is a Carathéodory function, then the polynomial $\pi_a(s) := \pi(s) - ia\chi(s)/(s-x_0)$ is again a solution to (19) and π_a/χ is still a Carathéodory function for all $a > 0$. Thus,

if r and χ happen to have a common real root, then they are quite far from defining π. This discrepancy arises because the dissipation polynomial of the pair $(-ia,(z-x_0))$ is identically zero and still $z\mapsto -ia/(z-x_0)$ is a non-zero Carathéodory function. It illustrates the fact that real parts of Carathéodory functions, unlike those of Schur functions, may be Poisson integrals of measures rather than functions, and therefore cannot in general be recovered from their pointwise trace on $\mathbb{R}$.

Applications of Problem $\mathscr{P}$ to filter design, as discussed in Section 1, typically involve a transmission polynomial r having real zeros near the endpoints of the bandwidth, because these ensure stiffness of the filter. Hence, in view of the previous discussion, we will not translate and prove Theorem 3.1 all the way in terms of Carathéodory functions. Nevertheless, rational Carathéodory functions with no pole on $\mathbb{R}$ play an important role in our proof and we derive below some of their fundamental properties in connection with the corresponding analog of (10).

To state our results, it is convenient to introduce the Hardy space $H^2(\mathbb{C}^-)$ consisting of those holomorphic functions f in $\mathbb{C}^-$ satisfying

$$\sup_{y<0}\left(\int_{-\infty}^{+\infty}|f(x+iy)|^2\,dx\right)^{1/2}<+\infty. \tag{20}$$ defH2

Such a function has a nontangential limit at almost every $x\in\mathbb{R}$. We still denote the nontangential limit with $f(x)$, the argument being now in $\mathbb{R}$ and not in $\mathbb{C}^-$. This nontangential limit lies in the Lebesgue space $L^2(\mathbb{R})$, and in fact $\|f\|_{L^2(\mathbb{R})}$ is equal to the supremum in (20). Moreover, for $z\in\mathbb{C}^-$, $f(z)$ can be recovered from f on $\mathbb{R}$ through a Cauchy as well as a Poisson integral [9, Ch. II, sec. 3]. In particular, a rational function π/χ with $\pi\in\mathbf{PE}_k$ and $\chi\in\mathbf{PE}_N$ does lie in $H^2(\mathbb{C}^-)$ if and only if $k<N$ and it has no pole in $\overline{\mathbb{C}^-}$, in other words if it vanishes at infinity and if every zero of χ in $\overline{\mathbb{C}^-}$ is cancelled by a corresponding zero of π. It follows easily that a rational Carathéodory function lies in $H^2(\mathbb{C}^-)$ if and only if its restriction to the real line lies in $L^2(\mathbb{R})$. Every $f\in H^2(\mathbb{C}^-)$ is the Cauchy integral of the non-tangential limit of its real part:

$$f(z)=-\frac{1}{i\pi}\int_{\mathbb{R}}\frac{\Re f(t)}{t-z}\,dt,\qquad z\in\mathbb{C}^-, \tag{21}$$ Cauchy1

and the non-tangential limit of its imaginary part is the Hilbert transform of the nontangential limit of its real part [9, Ch. III, sec. 2]:

$$\Im f(x)=\frac{1}{\pi}\lim_{\varepsilon\to 0^+}\int_{|x-t|>\varepsilon}\frac{\Re f(t)}{t-x}\,dt,\qquad \text{a.e. } x\in\mathbb{R}. \tag{22}$$ HT

Consequently, the nontangential limit of f can be recovered from its real part as

$$f(x)=\Re f(x)+\frac{i}{\pi}\lim_{\varepsilon\to 0^+}\int_{|x-t|>\varepsilon}\frac{\Re f(t)}{t-x}\,dt,\qquad \text{a.e. } x\in\mathbb{R}. \tag{23}$$ Plemelj

When f is smooth on $\mathbb{R}$, in particular if it is rational, then the last formula is valid for all $x \in \mathbb{R}$ and not just almost every x.

theyth **Theorem 3.2** *Let $g \in \mathbf{PM}_{2N}^{+}$ and $d \in \mathbf{P}_{2K}^{+}$, with $K < N$ and $\frac{d}{g} \in L^2(\mathbb{R})$. Let further $(x_1, \cdots, x_m)^T \in \mathbb{R}^m$ and $(z_1, \cdots, z_l)^T \in (\mathbb{C}^-)^l$ with $m+l = N$, and assume that $d(x_k) \neq 0$ for all $k \in \{1, \cdots, m\}$. Then, the following three properties hold.*

1. *There exists a unique pair of polynomials $\chi_g \in \mathbf{SBM}_N$ and $\pi_{d,g} \in \mathbf{P}_{N-1}$, such that the rational function $Y_{d,g} = \frac{\pi_{d,g}}{\chi_g}$ satisfies:*

   1a (a) $Y_{d,g} \in H^2(\mathbb{C}^-)$

   1b (b) $\pi_{d,g}\chi_g^* + \pi_{d,g}^*\chi_g = d$

   1c (c) $Y_{d,g} + Y_{d,g}^* = \frac{d}{g}$

2 2. *Let g_1, g_2 in $\mathbf{PM}_{2N}^{+}$ such that $\frac{d}{g_1}$ and $\frac{d}{g_2}$ are in $L^2(\mathbb{R})$. If*

   2a (a) $\forall k \in \{1..m\}\ Y_{d,g_1}(x_k) = Y_{d,g_2}(x_k)$,

   2b (b) $\forall k \in \{1..l\}\ Y_{d,g_1}(z_k) = Y_{d,g_2}(z_k)$,

   then $g_1 = g_2$ so that $\pi_{d,g_1} = \pi_{d,g_2}$ and $\chi_{g_1} = \chi_{g_2}$ by property 1.

3. *For fixed $d \in \mathbf{P}_{2K}^{+}$, the evaluation map $\theta : \overset{\circ}{\mathbf{PM}}_{2N}^{+} \to \mathbb{C}^N$ given by*

$$\theta(g) = \begin{pmatrix} Y_{d,g}(x_1) \\ \vdots \\ Y_{d,g}(x_m) \\ Y_{d,g}(z_1) \\ \vdots \\ Y_{d,g}(z_l) \end{pmatrix} \tag{24}$$ deftheta

   is well-defined and a diffeomorphism onto its image.

Proof. Let $u(z) = \Pi_{j=1}^{\ell}(z - x_j)^{2\kappa_j}$ be the monic divisor of g comprising all its real roots (if g has no real roots, then $\ell = 0$ and $u \equiv 1$). If $\pi_{d,g}$, χ_g satisfy (1b) and (1c), a short computation yields that $\chi_g = \varphi_N(g)$, where φ_N was defined in Proposition 3.1. In particular χ_g is uniquely determined by g and of necessity $u^{1/2} := \Pi_{j=1}^{\ell}(z - x_j)^{\kappa_j}$ divides χ_g. Then, condition (1a) implies that $u^{1/2}$ also divides $\pi_{d,g}$. Moreover, u divides d since

$d/g \in L^2(\mathbb{R})$. After cancellation of the factor $u = u^{1/2}(u^{1/2})^*$ on both sides of (1b), we find that $\pi_{d,g}/u^{1/2}$ is uniquely determined in $\mathbf{P}_{N-1-\Sigma_j K_j}$ by an equation of the Bezout type involving the coprime polynomials $\chi_g/u^{1/2}$ and $(\chi_g/u^{1/2})^*$. This establishes the uniqueness part of property 1 and the existence part follows easily by reverting the computations.

Let $Y_{d,g}$ be as in property 1. It is a rational function in $H^2(\mathbb{C}^-)$ whose real part on $\mathbb{R}$ is $d/(2g)$ by (1c), therefore (23) implies for $k \in \{1 \dots m\}$ that

$$Y_{d,g}(x_k) = \frac{d}{2g}(x_k) + \frac{i}{2\pi} \lim_{\varepsilon \to 0} \int_{\varepsilon < |t-x_k|} \frac{d(t)}{g(t)} \frac{dt}{t-x_k} \tag{25}$$ hilb

and (21) entails that $\forall k \in \{1 \dots l\}$

$$Y_{d,g}(z_k) = \frac{i}{2\pi} \int_{-\infty}^{\infty} \frac{d(t)}{g(t)} \frac{dt}{t-z_k}. \tag{26}$$ herg

Suppose now that g_1, g_2 are as in property 2. Note that $g_j(x_k) \neq 0$ for $j = 1,2$ and $1 \leq k \leq m$, since $d(x_k) \neq 0$ and $d/g_j \in L^2(\mathbb{R})$. From (2a) and (25), it follows that $g_2(x_k) = g_1(x_k)$ for $1 \leq k \leq m$ and also that

$$\begin{aligned} J(x_k) &:= \lim_{\varepsilon \to 0} \int_{\varepsilon < |t-x_k|} d(t) \frac{g_2(t) - g_1(t)}{g_1(t) g_2(t)} \frac{dt}{t-x_k} \\ &= \int_{-\infty}^{\infty} d(t) \frac{g_2(t) - g_1(t)}{g_1(t) g_2(t)} \frac{dt}{t-x_k} = 0, \end{aligned} \tag{27}$$

where we omitted the principal value in the integral (27) because *we claim* that the integrand is in fact nonsingular. Indeed, even though g_1 and g_2 may have real zeros (some of which may be common to g_1 and g_2), the fraction $d(g_2 - g_1)/g_1 g_2$ has no pole on $\mathbb{R}$; for if λ is a zero of g_j with multiplicity μ_j and, say, $\mu_1 \geq \mu_2$, then λ is a zero of d with multiplicity at least μ_1 (as $d/g_1 \in L^2(\mathbb{R})$) and it is a zero of $(g_2 - g_1)$ of multiplicity at least μ_2. Moreover, λ cannot coincide with x_k by our assumption that $d(x_k) \neq 0$ and in adition we just observed that $g_2 - g_1$ vanishes at x_k. *This proves the claim.*

Similarly we get $\forall k \in \{1 \dots l\}$ that

$$I(z_k) = \int_{-\infty}^{\infty} d(t) \frac{g_2(t) - g_1 t)}{g_1(t) g_2(t)} \frac{dt}{t-z_k} = 0, \tag{28}$$ intatzk

and taking conjugates

$$\overline{I(z_k)} = \int_{-\infty}^{\infty} d(t) \frac{g_2(t) - g_1(t)}{g_1(t) g_2(t)} \frac{dt}{t-\overline{z_k}} = 0. \tag{29}$$ intatzkbar

We now combine linearly equations (28), (29) and (27) using arbitrary complex coefficients $a = (a_1, a_2, \ldots, a_l)^T$, $b = (b_1, b_2, \ldots, b_l)^T$ and $c = (c_1, c_2, \ldots, c_m)^T$ to obtain

$$\sum_{k=1}^{l} (a_k I(z_k)) + \sum_{k=1}^{l} (b_k \overline{I(z_k)}) + \sum_{k=1}^{m} c_k J(x_k) = 0.$$

Putting everything over a common denominator yields

$$\int_{-\infty}^{\infty} d(t) \frac{g_2(t) - g_1(t)}{g_1(t) g_2(t) \prod_{k=1}^{l} |t - z_k|^2} \frac{P_{a,b,c}(t) dt}{\prod_{k=1}^{m} (t - x_k)} = 0, \tag{30}$$

TheIntegral

where $P_{a,b,c}$ is the polynomial defined by

$$\begin{aligned} P_{a,b,c}(z) = & \sum_{k=1}^{l} a_k \prod_{j=1\ldots l, j\neq k} (z - z_j) \prod_{j=1\ldots l} (z - \overline{z_j}) \prod_{j=1\ldots m} (z - x_j) + \\ & \sum_{k=1}^{l} b_k \prod_{j=1\ldots l} (z - z_j) \prod_{j=1\ldots l, j\neq k} (z - \overline{z_j}) \prod_{j=1\ldots m} (z - x_j) + \\ & \sum_{k=1}^{m} c_k \prod_{j=1\ldots l} (z - z_j) \prod_{j=1\ldots l} (z - \overline{z_j}) \prod_{j=1\ldots m, j\neq k} (z - x_j). \end{aligned} \tag{31}$$

The family of $2l + m$ polynomials obtained by setting a_k, b_k, and c_k to 0 except for one of them which is set to 1 is but the Lagrange interpolating basis of $\mathbf{P}_{2l+m-1}$ at the points $\{x_j, z_k, \bar{z}_k\}$. Therefore $P_{a,b,c}$ ranges over $\mathbf{P}_{2l+m-1}$ as (a,b,c) ranges over $\mathbb{C}^m \times \mathbb{C}^l \times \mathbb{C}^l$. Observing now that $g_2 - g_1$ vanishes at the x_k, (30) can be rewritten as

$$\int_{-\infty}^{\infty} d(t) \frac{P(t) P_{a,b,c}(t)}{g_1(t) g_2(t) \prod_{k=1}^{l} |t - z_k|^2} dt = 0, \tag{32}$$

TheIntegralBis

where $P(t)$ is the polynomial $(g_2(t) - g_1(t)) / \prod_1^m (t - x_k)$. Note that P has degree at most $2N - 1 = 2l + 2m - 1$ (for g_1, g_2 are monic of degree N). Hence, we can choose (a,b,c) so that $P_{a,b,c} = P$, and then we conclude from (32) that

$$\frac{d(t) P^2(t)}{g_1(t) g_2(t) \Pi_{k=1}^{l} |t - z_k|^2} = 0, \qquad t \in \mathbb{R},$$

because it is everywhere non-negative and its integral is zero. Since d is not identically zero, we get that $P = 0$ and consequently that $g_2 = g_1$. This proves property 2.

As to property 3, observe first that the restriction of d/g to the real line lies in $L^2(\mathbb{R})$ as soon as $d \in \mathbf{P}_{2K}^+$ and $g \in \mathbf{PM}_{2N}^+$ has no real root, and that the set of such g is precisely the interior of $\mathbf{PM}_{2N}^+$ by the discussion before Proposition 3.1. Hence θ is well-defined by

(24). Next, the derivatives of $Y_{d,g}(x_k)$, $Y_{d,g}(z_k)$ with respect to the coefficients of $g \in \overset{\circ}{\mathbf{P}}{}^{+}_{2N}$ must be computed. Put

$$g(x) = x^{2N} + g_{2N-1}x^{2N-1} + \cdots + g_0.$$

Since $\chi_g = \varphi_N(g)$, we get from (14) that $\partial\chi_g/\partial g_j$ exists in $\mathbf{P}_{N-1}$, $0 \le j \le 2N-1$, and that

$$\chi_g^*(x)\frac{\partial\chi_g}{\partial g_j}(x) + \chi_g(x)\frac{\partial\chi_g^*}{\partial g_j}(x) = x^j \tag{33}$$

expderphi

(note that $(\partial\chi_g/\partial g_j)^* = \partial\chi_g^*/\partial g_j$ since $*$ is a linear operation). Moreover, by property 1 already proved, $\pi_{d,g}$ is the solution to (1b) which is a nonsingular linear equation whose coefficients depend linearly on those of χ_g. Hence $\partial\pi_{d,g}/\partial g_j$ also exists in $\mathbf{P}_{N-1}$, $0 \le j \le 2N-1$, and by the Leibnitz rule

$$\chi_g^*\frac{\partial\pi_{d,g}}{\partial g_j} + \pi_{d,g}\frac{\partial\chi_g^*}{\partial g_j} + \chi_g\frac{\partial\pi_{d,g}^*}{\partial g_j} + \pi_{d,g}^*\frac{\partial\chi_g}{\partial g_j} = 0. \tag{34}$$

varpi

From the differentiability of χ_g, $\pi_{d,g}$ just pointed out, we get since evaluation at x_k is a linear operation and because $\chi_g(x_k) \neq 0$ that

$$\frac{\partial}{\partial g_j}\left(Y_{d,g}(x_k)\right) = F_{d,g,j}(x_k) \tag{35}$$

diffYxk

where

$$F_{d,g,j} = \frac{(\partial\pi_{d,g}/\partial g_j)\chi_g - \pi_{d,g}(\partial\chi_g/\partial g_j)}{\chi_g^2} \tag{36}$$

diffY

is a rational function in $H^2(\mathbb{C}^-)$ as it is the ratio of a polynomial of degree at most $2N-1$ by a stable polynomial of degree $2N$ (namely χ_g^2). Using (36), (34), (1b), (33) and the fact that $\chi_g = \varphi_N(g)$, we compute

$$\begin{aligned}
F_{d,g,j}(x) + F_{d,g,j}^*(x) &= \frac{\left(\frac{\partial\pi_{d,g}}{\partial g_j}\chi_g - \pi_{d,g}\frac{\partial\chi_g}{\partial g_j}\right)(\chi_g^*)^2 + \left(\frac{\partial\pi_{d,g}^*}{\partial g_j}\chi_g^* - \pi_{d,g}^*\frac{\partial\chi_g^*}{\partial g_j}\right)\chi_g^2}{\chi_g^2(\chi_g^*)^2}(x) \\
&= \frac{\left(\frac{\partial\pi_{d,g}}{\partial g_j}\chi_g^* + \frac{\partial\pi_{d,g}^*}{\partial g_j}\chi_g\right)\chi_g\chi_g^* - \left(\pi_{d,g}\frac{\partial\chi_g}{\partial g_j}(\chi_g^*)^2 + \pi_{d,g}^*\frac{\partial\chi_g^*}{\partial g_j}\chi_g^2\right)}{g^2}(x) \\
&= \frac{-\left(\frac{\partial\chi_g}{\partial g_j}\pi_{d,g}^* + \frac{\partial\chi_g^*}{\partial g_j}\pi_{d,g}\right)\chi_g\chi_g^* - \left(\pi_{d,g}(\frac{\partial\chi_g}{\partial g_j}(\chi_g^*)^2 + \pi_{d,g}^*\frac{\partial\chi_g^*}{\partial g_j}\chi_g^2\right)}{g^2}(x) \\
&= \frac{-\frac{\partial\chi_g}{\partial g_j}\left(\pi_{d,g}^*\chi_g + \pi_{d,g}\chi_g^*\right)\chi_g^* - \frac{\partial\chi_g^*}{\partial g_j}\left(\pi_{d,g}\chi_g^* + \pi_{d,g}^*\chi_g\right)\chi_g}{g^2}(x) \\
&= -\frac{d\left(\frac{\partial\chi_g}{\partial g_j}\chi_g^* + \frac{\partial\chi_g^*}{\partial g_j}\chi_g\right)}{g^2}(x) = -\frac{d(x)x^k}{g^2(x)}.
\end{aligned}$$

Since $F_{d,g}+F_{d,g}^{*}=2\Re F_{d,g}$ on $\mathbb{R}$, we get from (23), (35) and the previous computation:

$$\frac{\partial Y_{d,g}(x_k)}{\partial g_j}=-\frac{d(x_k)x_k^j}{2g^2(x_k)}-\frac{i}{2\pi}\lim_{\varepsilon\to 0^+}\int_{|x_k-t|>\varepsilon}\frac{d(t)t^j}{g^2(t)(t-x_k)}\,dt, \tag{37}$$

decPF

and combining linearly these partial derivatives leads us to the formula

$$D\big(Y_{d,g}(x_k)\big)[\delta g]=\frac{-d(x_k)\delta g(x_k)}{2g^2(x_k)}-\frac{i}{2\pi}\lim_{\varepsilon\to 0}\int_{\varepsilon<|t-x_k|}\frac{d(t)\delta g(t)}{g(t)^2}\,\frac{dt}{t-x_k},\quad \forall \delta g\in \mathbf{P}_{\mathbb{R},2N-1}. \tag{38}$$

ker1

The companion formula

$$D\big(Y_{d,g}(z_k)\big)[\delta g]=\frac{-i}{\pi}\int_{-\infty}^{\infty}\frac{d(t)\delta g(t)}{g(t)^2}\,\frac{dt}{t-z_k},\quad \forall\ \delta g\in \mathbf{P}_{\mathbb{R},2N-1} \tag{39}$$

ker2

is obtained in the same manner, appealing to (21) rather than (23). Hereafter, we drop the dependence on d,g and we write for simplicity Y_{x_k} (resp Y_{z_k}) instead of $Y_{d,g}(x_k)$ (resp. $Y_{d,g}(z_k)$). Altogether we find that the application θ is differentiable with derivative

$$D\theta(g):\delta g\in \mathbf{P}_{\mathbb{R},2N-1}\to\begin{pmatrix} DY_{x_1}(\delta g))\\ \vdots\\ DY_{x_1}(\delta g)\\ DY_{z_1}(\delta g)\\ \vdots\\ DY_{z_l}(\delta g)\end{pmatrix}\in\mathbb{C}^N, \tag{40}$$

where $DY_{x_k}(\delta g)$ is given by (38) and $DY_{z_k}(\delta g)$ by (39).

Now, suppose that $\delta g\in\ker(D\theta)$. Separating real and imaginary parts in (38), we see that δg vanishes at every x_k. Consequently the principal part of the integral in (38) can be omitted, and this integral is zero for all x_k. Moreover, the integrals in (39) vanish for all z_k. Thus, equating to zero an arbitrary linear combination of the integrals in (38) and those in (39) together with their conjugates, as x_k ranges over $\{x_1,\cdots,m\}$ and z_k ranges over $\{z_1,\cdots,z_l\}$, we get in the same manner as we got (32) that

$$\forall P_{a,b,c}\in\mathbf{P}_{2l+m-1},\ \int_{-\infty}^{\infty}d(t)\frac{\hat{\delta}g(t)P_{a,b,c}(t)}{g^2(t)\prod_{k=1}^{l}|t-z_k|^2}\,dt=0, \tag{41}$$

variatdg

where $\hat{\delta}g$ is the real polynomial which is the quotient of δg by $\prod_1^m(t-x_k)$. Picking $P_{a,b,c}=\hat{\delta}g$ in (41), we conclude since the integrand is nonnegative that $\hat{\delta}g=0$, hence

also $\delta g = 0$. Therefore $D\theta(g)$ is injective, thus it is invertible and θ is a local diffeomorphism. Finally, we know from property 2 that θ is injective, therefore it is a diffeomorphism from $\overset{\circ}{\mathbf{PM}}{}^{+}_{2N}$ onto its image. □

3.3 Injectivity of ψ.

inj

We can now establish that the map ψ introduced in (11) is one-to-one.

inject **Proposition 3.2** *The map ψ is injective.*

Proof. Let $v = (\gamma_1, \ldots, \gamma_m, \beta_1, \ldots, \beta_m) \in \mathbb{D}^m \times \mathbb{P}_Z^+$ and assume that there exist distinct polynomials $p_1(z)$ and $p_2(z)$ in $\mathbf{PM}_N$ such that $\psi(p_1) = \psi(p_2)$. Put $q_j = \varphi_N(p_j p_j^* + rr^*)$ for $j = \{1,2\}$, so that our assumption means:

$$\frac{p_1}{q_1}(x_k) = \frac{p_2}{q_2}(x_k),\ 1 \le k \le m, \quad \text{and} \quad \frac{p_1}{q_1}(z_\ell) = \frac{p_2}{q_2}(z_\ell),\ 1 \le l \le l. \tag{42}$$ noninja

By the Feldtkeller equation (6), $|p_j(t)/q_j(t)| \le 1$ for $t \in \mathbb{R}$, and $|p_j(t)/q_j(t)| = 1$ exactly when t is a real zero of r with multiplicity $d \ge 1$ which is not a zero of p_j of multiplicity greater than, or equal to d; here, when p_j and q_j both vanish at t, the value $p_y(t)/q_j(t)$ is understood as the limit of $p_j(\tau)/q_j(\tau)$ when $\tau \to t$. In particular, there are at most $\deg r$ real numbers t for which $|p_j(t)/q_j(t)| = 1$, hence we can find a complex number ξ of modulus 1, distinct from -1, such that $1 + \xi\, p_j/q_j$ is never zero on $\mathbb{R}$ for $j = \{1,2\}$. We introduce following fractions,

$$\begin{aligned} G_j(z) = \frac{1 - \xi\frac{p_j(z)}{q_j(z)}}{1 + \xi\frac{p_j(z)}{q_j(z)}} &= \frac{1-\xi}{1+\xi} + \left(\frac{2\xi}{1+\xi}\right)\frac{q_j(z) - p_j(z)}{q_j(z) + \xi p_j(z)} \\ &\overset{def}{=} \frac{1-\xi}{1+\xi} + Y_j(z) \end{aligned} \tag{43}$$ cayley

Being the Cayley transform of a Schur function, G_j is a Carathédory function, and so is Y_j as it differs from G_j by the pure imaginary constant $\frac{1-\xi}{1+\xi}$. The choice of ξ ensures the continuity of G_j and hence of Y_j on the real axis, hence we conclude that Y_j lies in $H^2(\mathbb{C}^-)$ by noting that its numerator has degree at most $N-1$ while its denominator has degree N. By a computation similar to (18) we obtain

$$\begin{aligned} G_j + G_j^* &= \frac{2rr^*}{(q_j + \xi p_j)(q_j + \xi p_j)^*} \\ &= Y_j + Y_j^* \end{aligned} \tag{44}$$ contdiss

We now apply Theorem 3.2 with $g_j = \frac{1}{|1+\xi|^2}(q_j+\xi p_j)(q_j+\xi p_j)^*$ and $d = \frac{2}{|1+\xi|^2} rr^*$ after checking that d/g_j, being the real part on $\mathbb{R}$ of the H^2-function Y_j, belongs to $L^2(\mathbb{R})$. Setting,

$$\chi_j = \frac{q_j+\xi p_j}{1+\xi} \quad \text{and} \quad \pi_j = \left(\frac{2\xi}{(1+\xi)^2}\right)(q_j - p_j). \tag{45}$$ deterchipi

we verify by using (44) and $Y_j = \pi_j/\chi_j$ that the pair of polynomials χ_j, π_j verifies all the assertions 1a,b,c of Theorem 3.2 verified by χ_{g_j} and π_{d,g_j}. Using the second assertion of the theorem yields $\pi_1 = \pi_1$ and $\chi_1 = \chi_2$ and therefore $p_1 = p_2$.

□

3.4 Properness of ψ

proper **Proposition 3.3** *The map ψ defined in* (11) *is proper: for any compact $K \subset \mathbb{D}^m \times \mathbb{P}^+_Z$, the set $\psi^{-1}(K)$ is compact in* $\mathbf{PM}_N$.

Proof. Let K be a compact subset of the topological space $\mathbb{D}^m \times \mathbb{P}^+$ and denote $W = \psi^{-1}(K)$. By $\psi's$ continuity, W is closed. It therefore remains to prove that W is bounded.

Assume for a contradiction that there is an unbounded sequence p_n in $\psi^{-1}(K)$, and let us write $\psi(p_n) = (\gamma_1^{\{n\}}, \cdots, \gamma_m^{\{n\}}, \beta_1^{\{n\}}, \cdots \beta_l^{\{n\}})$. By definition of ψ, it holds that $\gamma_j^{\{n\}} = p_n(x_k)/q_n(x_k)$ and $\beta_\ell^{\{n\}} = p_n(z_\ell)/q_n(z_\ell)$, with $q_n = \varphi_N(p_n p_n^* + rr^*)$, where φ_N was defined in Proposition 3.1, item b). Extracting a subsequence if necessary, we may assume that $\psi(p_n)$ converges to some $(\gamma_1, \cdots, \gamma_m, \beta_1, \cdots, \beta_l) \in K$ in $\mathbb{D}^m \times \mathbb{P}^+_Z$. Thus, our assumption is that $\gamma_k^{\{n\}} \to \gamma_k$ and $\beta_\ell^{\{n\}} \to \beta_\ell$ as $n \to \infty$ for $1 \le k \le m$ and $1 \le \ell \le l$.

By Euclidean division of $p_n(t)$ by $L(t) := \Pi_{k=1}^m (t - x_k)$, we can write

$$p_n(t) = \sum_{k=1}^m p_n(x_k) L_{x_k}(t) + L(t) h_n(t) \tag{46}$$ decomp

where $L_{x_k}(t) = \prod_{\substack{1 \le j \le m \\ j \ne k}} \frac{t - x_j}{x_k - x_j}$ is the k-th Lagrange interpolation polynomial of the set $\{x_1, \ldots x_m\}$ and h_n is a monic polynomial of degree $N - m = l$. It may of course

happen that $m = 0$ (if there is no x_k), in which case we set $L \equiv 1$ and $L_{x_k} \equiv 0$; then $h_n = p_n$. To the opposite, it may be that $l = 0$ (if there is no z_ℓ) in which case $h_n = 1$.

Let $\|p_n\|$ indicate the norm of p_n in $\mathbf{PM}_N \sim \mathbb{C}^N$. The precise norm we use is irrelevant for they are all equivalent. Since $p_n/\|p_n\|$ is bounded whereas $\|p_n\|$ is not, we may assume upon taking another subsequence if necessary that $\|p_n\| \to +\infty$ and $p_n/\|p_n\| \to g$ where $g \in \mathbf{P}_N$ is such that $\|g\| = 1$. In another connection, using (6), one easily checks that

$$\forall k \in \{1 \dots m\} \; |p_n(x_k)|^2 = \frac{\left|\gamma_k^{\{n\}}\right|^2}{1 - \left|\gamma_k^{\{n\}}\right|^2} |r(x_k)|^2, \tag{47}$$ bound

and since $\gamma_k^{\{n\}} \to \gamma_k \in \mathbb{D}$ we conclude that $p_n(x_k)$ is bounded independently of n. Thus, dividing (46) by $\|p_n\|$ and letting $n \to \infty$, we get $g = Lh$ where h is the limit of $h_n/\|p_n\|$. Observe that $h \in \mathbf{P}_{l-1}$ for $h_n/\|p_n\| \in \mathbf{PE}_l$ has leading coefficient $1/\|p_n\|$ which tends to 0 as $n \to \infty$. If $l = 0$ the proof is finished because then $h = 0$, contradicting that $g \neq 0$.

Suppose next that $l > 0$ and rewrite the Feldtkeller equation after division by $\|p_n\|^2$ as

$$\frac{p_n p_n^*}{\|p_n\|^2} + \frac{rr^*}{\|p_n\|^2} = \varphi\left(\frac{p_n p_n^*}{\|p_n\|^2} + \frac{rr^*}{\|p_n\|^2}\right)\left(\varphi\left(\frac{p_n p_n^*}{\|p_n\|^2} + \frac{rr^*}{\|p_n\|^2}\right)\right)^*, \tag{48}$$ nfeld

see Proposition 3.1 for the definition of φ. By continuity of the latter shown in that proposition, and since $rr^*/\|p\|_n^2 \to 0$, we get from (48) that $\varphi((p_n p_n^* + rr^*)/\|p_n\|^2)$ converges to $\varphi(gg^*)$ in $\mathbf{P}_N$ as $n \to \infty$. Moreover, as $q_n/\|p_n\| = a_n \varphi\left((p_n p_n^* + rr^*)/\|p_n\|^2\right)$ for some $a_n \in \mathbb{C}$ with $|a_n| = 1$ by Proposition 3.1, we may assume upon extracting another subsequence that $a_n \to a$ with $|a| = 1$ and therefore that $q_n/\|p_n\| \to a\varphi(gg^*)$. In addition, since $L = L^*$ has only real roots, it holds that $\varphi(gg^*) = bL\varphi(hh^*)$ for some $b \in \mathbb{C}$ with $|b| = 1$. Therefore, because convergence in $\mathbf{P}_N$ implies pointwise convergence on $\mathbb{C}$ and since $\varphi(gg^*)$ has no zeros in $\mathbb{C}^-$ by definition of φ, we have that

$$\beta_\ell = \lim_{n\to\infty} \beta_\ell^{\{n\}} = \lim_{n\to\infty} \frac{p_n(z_\ell)}{q_n(z_\ell)} = \lim_{n\to\infty} \frac{p_n(z_\ell)/\|p_n\|}{q_n(z_\ell)/\|p_n\|} = \frac{g(z_\ell)}{a\varphi(gg^*)(z_\ell)} = \frac{h(z_\ell)}{ab\varphi(hh^*)(z_\ell)}.$$

Hence the $l \times l$ matrix $P(Z,\beta)$ defined by (9) is the Pick matrix corresponding to the interpolation data $(z_\ell, (h/(ab\varphi(hh^*)))(z_\ell))$, and since $h/(ab\varphi(hh^*))$ is a Blaschke product of degree at most $l-1$ it cannot have full rank, see discussion after (9). This, however, contradicts the fact that $P(Z,\beta)$ is nonsingular by definition of $\mathbb{P}_Z^+$. □

3.5 ψ is a homeomorphism from $\mathbf{PM}_N$ onto $\mathbb{D}^m \times \mathbb{P}^+$

mainstep1

We are now in position to prove the first claim of Theorem 3.1. It will be convenient to invoke a famous result by Brouwer, known as *invariance of the domain* [15, chap. 10, sect. 62]. If $\Omega \subset \mathbb{R}^n$ is open and $f : \Omega \to \mathbb{R}^n$ is continuous and injective, it says that f is an open map, meaning that it maps open sets to open sets. Hence $f(\Omega)$ is open and the inverse map $f^{-1} : f(\Omega) \to \Omega$ is continuous, that is: f is a homeomorphism onto its image.

Munkres

psihomeom

Proposition 3.4 *ψ defined in* (11) *is a homeomorphism from $\mathbf{PM}_N$ onto $\mathbb{D}^m \times \mathbb{P}^+$.*

Proof. We may regard ψ as a map from $\mathbb{R}^{2N}$ into $\mathbb{R}^{2N}$. By Corollary 3.1 and Proposition 3.2 it is continuous and injective, hence the image $\psi(\mathbf{PM}_N)$ is open and ψ is a homeomorphism onto this image by invariance of the domain. In another connection, the properness of ψ implies that $\psi(\mathbf{PM}_N)$ is closed in $\mathbb{D}^m \times \mathbb{P}^+$. Indeed, suppose that $\psi(p_n)$ is a sequence in $\psi(\mathbf{PM}_N)$ that converges to some $\nu \in \mathbb{D}^m \times \mathbb{P}^+$. Because the union of a convergent sequence and its limit is compact, properness entails that we can extract a subsequence (p_{n_k}) converging to some $p \in \mathbf{PM}_N$, and then $\psi(p) = \nu$ by continuity. Hence $\psi(\mathbf{PM}_N)$ contains its limit point ν, thereby showing that it is closed.

Now, being the product of two connected topological spaces, the space $\mathbb{D}^m \times \mathbb{P}^+_Z$ is connected. Consequently $\psi(\mathbf{PM}_N)$, which is both open and closed in $\mathbb{D}^m \times \mathbb{P}^+_Z$, is either empty or the whole space and in fact it must be the whole space because it is certainly not empty. This concludes the proof.

□

3.6 ψ is a local diffeomorphism wherever it is differentiable

mainstep2

We established through Proposition 3.4 and Corollary 3.1 that $p \mapsto \psi(p)$ is a homeomorphism which is differentiable at every p having no common real root with r. A pending question is whether ψ^{-1} is differentiable at $\psi(p)$ for such p. This issue is of practical importance because, as we shall see, computationally efficient algorithms for the numerical inversion of ψ can be based on continuation techniques, which rely on the differentiability of ψ^{-1}. As for the proof that $D\psi$ is injective, we will use an equivalent formulation of that question in terms of positive real functions.

As an extra-piece of notation, we define $\mathbf{PM}_N(r)$ to be the open subset of $\mathbf{PM}_N$ comprised of those polynomials that have no common real root with r.

prop_diff **Proposition 3.5** *The map ψ is a diffeomorphism from $\mathbf{PM}_N(r)$ onto its image.*

Proof. The proof consists in showing that, locally, ψ can be seen as a composition of diffeomorphisms involving the map θ defined in Theorem 3.2. This will ensure that ψ is a local diffeomorphism, and since we know from Proposition 3.4 that it is a (global) homeomorphism $\mathbf{PM}_N \to \mathbb{D}^m \times \mathbb{P}^+$ the proof will be complete.
Let $p_0 \in \mathbf{PM}_N(r)$. By (6), the polynomial $q_0 := \varphi_N(p_0 p_0^* + rr^*) \in \mathbf{SBM}_N$ is devoid of real roots and the fraction p_0/q_0 assumes modulus 1 at the real zeros of r (if any) while it has modulus strictly less than 1 elsewhere on $\mathbb{R} \cup \mathbb{C}^-$. Thus, we can thus choose $\xi \in \mathbb{C}$ of unit modulus, $\xi \neq -1$, such that $(\xi p_0 + q_0)/(1+\xi)$ lies in $\mathbf{SM}_N$.

Since $\mathbf{SM}_N$ is open in $\mathbf{PM}_N$ (*cf.* discussion before Proposition 3.1), the continuity and differentiability of φ_N ensures the existence of a neighborhood V of p_0 in $\mathbf{PM}_N(r)$ such that the map $\eta(p) := (\xi p + \varphi_N(p))/(1+\xi)$ is defined and differentiable on V with $\eta(V) \subset \mathbf{SM}_N$. *We claim* that its differential $D\eta$ is invertible at every $p \in V$. Indeed, it is enough to show that $D\eta$ is injective and its kernel consists of those dp for which

$$\xi dp + dq = 0 \tag{49}$$

where $dq = D\varphi_N(pdp^* + p^*dp)$ satisfies (*cf.* (14))

$$q^*dq + qdq^* = pdp^* + p^*dp. \tag{50}$$

Combining the two last equations yields

$$\overline{\xi}(\xi p + q)dp^* + \xi(\xi p + q)^* dp = 0. \tag{51}$$ diffetaz

The polynomial $(\xi p + q)$ is strictly stable and therefore it is coprime with its paraconjugate, hence it must divide dp by (51). Since dp has degree at most $N-1$ while $(\xi p + q)$ has degree N (remember $\xi \neq -1$), this yields $dp = 0$, this *proves the claim* and shows that η is a diffeormophism when restricted to V. In particular, $\eta(V)$ is open in $\mathbf{SM}_N$.

Next, consider the map $m : \eta(V) \to \overset{\circ}{\mathbf{PM}}{}^+_{2N}$ given by $m(v) = vv^*$; to check that m indeed maps $\eta(V)$ into the interior of $\mathbf{PM}^+_{2N}$, simply observe that $(\xi p + q)(\xi p + q)^*$ has no real root because so does $(\xi p + q)$ as it is strictly stable. Shrinking V if necessary, we get from Proposition 3.1 that m is the restriction to $\eta(V)$ of φ_N^{-1} and therefore a diffeomorphism onto its image.

Then, putting $g = \eta(p)\eta(p)^*$ and $d = 2rr^*/|1+\xi|^2$, we see from Theorem 3.2 that the map θ defined there allows us to evaluate at the interpolation points $(x_a, \cdots, x_m, z_1, \cdots, z_l)$ the positive real function

$$Y = (q - \xi p)/(\xi p + q) = \frac{1+\xi}{1-\xi} + Y_{d,g}$$

in a smooth and diffeomorphic manner with respect to $m(\eta(p))$.

Eventually we need to come back to the "scattering domain", that is, we must compute the $p(x_k)/q(x_k)$, $p(z_\ell)/q(z_\ell)$ in terms of the $Y(x_k)$, $Y(z_\ell)$ in a diffeomorphic manner. This is easily accomplished by smoothly inverting the correspondence $p_j/q_j \mapsto Y_j$ in equation (43) upon defining $\tau : \mathbb{C}^+ \to \mathbb{D}$ by

$$\tau(z) = \overline{\xi}\frac{1 - \left(\frac{1-\xi}{1+\xi} + z\right)}{1 + \left(\frac{1-\xi}{1+\xi} + z\right)}, \qquad z \in \mathbb{C}^+, \tag{52}$$

and observing that $\tau(Y_j) = p_j/q_j$. Altogether, letting $\tau_N : (\mathbb{C}^+)^N \to \mathbb{D}^N$ act componentwise as τ, we find that on V

$$\psi = \tau_N \circ \theta \circ \varphi_N^{-1} \circ \eta \tag{53}$$

decomppsi

which expresses ψ as a composition of local diffeomorphisms, thereby concluding the proof. □

Remark: in the decomposition (53), the maps τ_N and η depend on ξ and therefore on the point p_0 around which we carry out the local analysis of ψ. In fact, there is no global decomposition of ψ in terms of θ, but merely a collection of local ones, taylored so as to associate a non singular Carathéodory function Y (*i.e.* one having no pole on $\mathbb{R}$) to the initial Schur function (*i.e.* scattering element) p_0/q_0.

We now come to an application of Proposition 3.5 that enables us to use continuation techniques in order to practically solve for Problem $\mathscr{P}$.

Proposition 3.6 *$\psi(\mathbf{PM}_N(r))$ is an open, dense and connected subset of $\mathbb{D}^m \times \mathbb{P}^+$. Suppose that ν_0, ν_1 both lie in $\psi(\mathbf{PM}_N(r))$, and that γ is a continuous path from ν_0 to ν_1 in $\mathbb{D}^m \times \mathbb{P}^+$. Then, for every $\varepsilon > 0$ there exists a continuous path $\hat{\gamma}$ from ν_0 to ν_1 in $\psi(\mathbf{PM}_N(r))$ such that*

$$\sup_{t\in[0,1]} \|\hat{\gamma}(t) - \gamma(t)\| \leq \varepsilon,$$

where $\|.\|$ designates an arbitrary but fixed norm on $\mathbb{R}^{2N} \sim \mathbb{C}^N \supset \mathbb{D}^m \times \mathbb{P}^+$.

Proof. By Proposition 3.4 ψ is a homeomorphism $\mathbf{PM}_N \to \mathbb{D}^m \times \mathbb{P}^+$. Openness, density and connectedness of $\psi(\mathbf{PM}_N(r))$ in $\mathbb{D}^m \times \mathbb{P}^+$ will thus follow from the corresponding properties of $\mathbf{PM}_N(r)$ in $\mathbf{PM}_N$. These are easily verified, for if $\{\zeta_1, \cdots, \zeta_\mu\}$ are the real roots of r then $\mathbf{PM}_N(r)$ consists of those monic polynomials no root of which coincides with a ζ_j. This is clearly an open condition. Moreover, given any $p(z) = \Pi_{k=1}^N (z - \xi_k)$ in $\mathbf{PM}_N$, we can find ξ_k' arbitrary close to ξ_k which is not a ζ_j, thereby showing the density of $\mathbf{PM}_N(r)$. In addition, two polynomials $\Pi_{j=1}^N (z - \xi_k^{(1)})$ and $\Pi_{j=1}^N (z - \xi_k^{(2)})$ such that neither $\xi_k^{(1)}$ nor $\xi_k^{(2)}$ is a ζ_j can be deformed into each other within $\mathbf{PM}_N(r)$ by a map $t \mapsto \Pi_{j=1}^N (z - \xi_k(t))$ where $t \mapsto \xi_k(t)$, $t \in [0,1]$, is a continuous path from $\xi_k^{(1)}$ to $\xi_k^{(2)}$ in $\mathbb{C}$ which does not meet any ζ_j; hence $\mathbf{PM}_N(r)$ is connected.

Next, let $\gamma : [0,1] \to \mathbb{D}^m \times \mathbb{P}^+$ be a continuous map such that $\gamma(0) = \nu_0$ and $\gamma(1) = \nu_1$ with $\nu_0, \nu_1 \in \psi(\mathbf{PM}_N(r))$. By continuity, each $t \in [0,1]$ lies in a relatively open interval $I_t \subset [0,1]$ such that $\|\gamma(t') - \gamma(t)\| < \varepsilon/8$ for all $t' \in I_t$. By compactness, we can cover $[0,1]$ with finitely many such intervals and discard some if necessary so as to get a covering family no interval of which is included in the union of the others. Shrinking these remaining intervals if needed, we may arrange things so that any three have empty intersection, thereby giving rise to a finite sequence of real numbers:

$$0 = a_0 < a_1 < b_0 < a_2 < b_1 < a_3 < b_2 < \cdots b_{K-2} < a_K < b_{K-1} < b_K = 1$$

such that $[0,1] = [0,b_0) \cup_{j=1}^{K-1} (a_j, b_j) \cup (a_K, 1]$ and any t, t' belonging to one of these intervals satisfy $\|\gamma(t) - \gamma(t')\| < \varepsilon/4$. Pick $t_j \in (a_{j+1}, b_j)$ for each $j \in \{0, \cdots, K-1\}$, and let $t_{j_1}, \cdots, t_{j_L}$ be those t_j, if any, whose image under γ lies in $\mathbb{D}^m \times \mathbb{P}^+ \setminus \psi(\mathbf{PM}_N(r))$. We modify γ slightly in a neighborhood of the t_{j_ℓ} as follows. We let $\eta > 0$ be so small that $[t_{j_\ell} - \eta, t_{j_\ell} + \eta] \subset (a_{j_\ell+1}, b_{j_\ell})$ for $1 \le \ell \le L$; in particular these intervals are all disjoint. We pick for each ℓ some $\nu_\ell \in \psi(\mathbf{PM}_N(r))$ with $\|\gamma(t_{j_\ell}) - \nu_\ell\| < \varepsilon/8$, and we put

$$\gamma_1(t) = \begin{cases} \gamma(t) & \text{if} \quad |t - t_{j_\ell}| > \eta \text{ for } 1 \le \ell \le L, \\ \frac{t_{j_\ell} - t}{\eta} \gamma(t) + (1 - \frac{t_{j_\ell} - t}{\eta}) \nu_\ell & \text{if} \quad t \in [t_{j_\ell} - \eta, t_{j_\ell}], \\ (1 - \frac{t - t_{j_\ell}}{\eta}) \nu_\ell + \frac{t - t_{j_\ell}}{\eta} \gamma(t) & \text{if} \quad t \in [t_{j_\ell}, t_{j_\ell} + \eta]. \end{cases}$$

Clearly γ_1 is continuous and takes its values in $\mathbb{D}^m \times \mathbb{P}^+$ because the latter is convex. Moreover, we get by construction that

$$\|\gamma_1(t) - \gamma(t)\| < \varepsilon/8, \qquad t \in [0,1], \tag{54}$$ distgg1

so by the triangle inequality it holds whenever t, t' belong to (a_j, b_j) with $1 \le j \le K-1$ or to $[a_0, b_0)$ or else to $(a_K, b_K]$ that

$$\begin{array}{rcl} \|\gamma_1(t) - \gamma_1(t')\| & \le & \|\gamma_1(t) - \gamma(t)\| + \|\gamma(t) - \gamma(t')\| + \|\gamma(t') - \gamma_1(t')\| \\ & < & \frac{\varepsilon}{8} + \frac{\varepsilon}{4} + \frac{\varepsilon}{8} = \varepsilon/2. \end{array} \tag{55}$$ estg1

Also, by construction, we have that $\gamma_1(t_j) \in \psi(\mathbf{PM}_N(r))$ for $0 \leq j \leq K-1$.

Set by convention $t_{-1} = 0$ and $t_K = 1$, so that we may regard γ_1 as the concatenation of $K+1$ maps $\gamma_{1,j} : [t_{j-1}, t_j] \to \mathbb{D}^m \times \mathbb{P}^+$ for $0 \leq j \leq K$, where $\gamma_{1,j}$ indicates the restriction of γ_1 to $[t_{j-1}, t_j]$. Define

$$V_j := \{\nu \in \mathbb{D}^m \times \mathbb{P}^+, \inf_{t \in [t_{j-1}, t_j]} \|\nu - \gamma_1(t)\| < 3\varepsilon/8\}, \qquad 0 \leq j \leq K. \tag{56}$$ defVj

Suppose for a while that to each index $j \in \{0, \cdots K\}$ we can associate a continuous path $\gamma_{2,j} : [t_{j-1}, t_j] \to V_j \cap \psi(\mathbf{PM}_N(r))$ linking $\gamma_1(t_{j-1})$ and $\gamma_1(t_j)$. Then, *we claim* that the concatenation $\hat{\gamma}$ of all these paths will satisfy our requirements. Indeed, it is clear by construction that $\hat{\gamma}$ is continuous $[0,1] \to \psi(\mathbf{PM}_N(r))$ with $\hat{\gamma}(0) = \nu_0$ and $\hat{\gamma}(1) = \nu_1$. Moreover, if we pick $t \in [t_{j-1}, t_j]$ for some $j \in \{0, \cdots, K\}$ and if we let $\tau_j \in [t_{j-1}, t_j]$ be such that

$$\|\gamma_{2,j}(t) - \gamma_1(\tau_j)\| = \min_{\tau \in [t_{j-1}, t_j]} \|\gamma_{2,j}(t) - \gamma_1(\tau)\|, \tag{57}$$ minVj

we find by (57), (56), (55) and (54), since $[t_{j-1}, t_j] \subset (a_j, b_j)$ whenever $1 \leq j \leq K-1$ while $[t_{-1}, t_0] \subset [a_0, b_0)$ and $[t_{K-1}, t_K] \subset (a_K, b_K]$, that

$$\|\gamma_{2,j}(t) - \gamma(t)\| \leq \|\gamma_{2,j}(t) - \gamma_1(\tau_j)\| + \|\gamma_1(\tau_j) - \gamma_1(t)\| + \|\gamma_1(t) - \gamma(t)\| < \frac{3\varepsilon}{8} + \frac{\varepsilon}{2} + \frac{\varepsilon}{8} = \varepsilon.$$

This proves the claim.

It remains for us to show the existence of $\gamma_{2,j}$. Since connected open sets in $\mathbb{R}^{2N}$ are arcwise connected and $\gamma_1(t_{j-1})$, $\gamma_1(t_j)$ lie in $V_j \cap \psi(\mathbf{PM}_N(r))$, it is enough to prove that the latter is a connected set (the fact that $\gamma_{2,j}$ is defined over $[t_{j-1}, t_j]$ is irrelevant for one can always reparametrize a continuous path). Now, V_j is connected because each $\nu \in V_j$ can be connected to some $\gamma_1(t)$ with $t \in [t_{j-1}, t_j]$ by a line segment included in V_j (remember $\mathbb{D}^m \times \mathbb{P}^+$ is convex), while the collection of all such $\gamma_1(t)$ is obviously connected and included in V_j. Therefore, it is sufficient to establish that a connected open set included in $\mathbb{D}^m \times \mathbb{P}^+$ has an intersection with $\psi(\mathbf{PM}_N(r))$ which is connected. As ψ is a homeomorphism $\mathbf{PM}_N \to \mathbb{D}^m \times \mathbb{P}^+$, this is tantamount to show if $U \subset \mathbf{PM}_N$ is a connected open set, then $U \cap \mathbf{PM}_N(r)$ is still connected. Now, $\mathbf{PM}_N \setminus \mathbf{PM}_N(r)$ is the union of μ complex affine hyperplanes A_k defined for $k \in \{1, \cdots, \mu\}$ by

$$A_k := \{P \in \mathbf{PM}_N : \ P(\zeta_k) = 0\}.$$

Thus, by induction, it is enough to show that $U \setminus A_k$ is connected.

Since it has complex dimension $N-1$, clearly A_k is a C^∞-smooth closed manifold of real dimension $2N-2$ in $\mathbf{PM}_N \sim\sim \mathbb{R}^{2N}$. Let $P_0, P_1 \in U \setminus A_k$ and $F : [0,1] \to U$ a continuous

path with $F(0) = P_0$ and $F(1) = P_1$; such a path exists since U is connected. The argument below has two steps: first we replace F by a smooth map $\mathfrak{P}$ and second we show that suitable perturbations of $\mathfrak{P}$ do not meet A_k.

By compactness of the image $F([0,1])$, there is $d > 0$ such that $\|F(t) - P\| > d$ whenever $t \in [0,1]$ and $P \in \mathbf{PM}_N \setminus U$. Thanks to the stone-Weierstrass theorem, there is a polynomial map $\mathfrak{P} : [0,1] \to \mathbf{PM}_N$ such that $\|F(t) - \mathfrak{P}(t)\| < d/4$ for all $t \in [0,1]$; here, by a polynomial map, we mean that each component is a polynomial in t. Then, the map $G : [0,1] \to \mathbf{PM}_N$ defined as

$$G(t) := \mathfrak{P}(t) + (1-t)(F(0) - \mathfrak{P}(0)) + t((F(1) - \mathfrak{P}(1))$$

is C^∞-smooth (in fact: a polynomial in t) with $G(0) = P_0$ and $G(1) = P_1$. Moreover since $\|G(t) - \mathfrak{P}(t)\| < d/4$ for all t, the map G is valued in U and $\|G(t) - P\| > d/2$ whenever $t \in [0,1]$ and $P \in \mathbf{PM}_N \setminus U$.

Next, let B_ρ be the open disk of radius $\rho > 0$ centered at 0 in $\mathbb{R}^2$. For simplicity, we shall abbreviate a vector $((x_0, y_0), \cdots, (x_{N-1}, y_{N-1})) \in \mathbb{R}^{2N}$ as X. Denoting by z the variable in $\mathbf{P}_{N-1} \sim \mathbb{R}^{2N}$, consider the map $Q : [0,1] \times B_\rho^N \to \mathbf{P}_{N-1}$ given by

$$Q\big(t, X\big) = t(1-t)\Big((x_{N-1} + iy_{N-1})z^{N-1} + \cdots + (x_1 + iy_1)z + (x_0 + iy_0)\Big).$$

Clearly the map Q is C^∞-smooth with $Q(0,X) = Q(1,X) = 0$, and its derivative at every (t,X) with $t \neq 0, 1$ is surjective since $\partial Q/\partial x_k = t(1-t)z^k$ and $\partial Q/\partial y_k = it(1-t)z^k$. Moreover, we can pick ρ so small that $\|Q(t,X)\| < d/2$ for all (t,X). Then, the map $\mathfrak{T} : [0,1] \times B_\rho^N \to \mathbf{PM}_N$ given by

$$\mathfrak{T}(t,X) = G(t) + Q(t,X)$$

is U-valued with $\mathfrak{T}(0,X) = P_0$ and $\mathfrak{T}(1,X) = P_1$ for all X. Let us denote with $\operatorname{Im} D\mathfrak{T}(t,X)$ the image of the derivative $D\mathfrak{T}(t,X)$ at (t,X); when $t = 0$ or $t = 1$, this is understood as the image of the derivative with respect to X alone. We put also TA_k (resp. $T_{\mathbf{PM}_N}$) for the tangent space to A_k (resp. $\mathbf{PM}_N$) which is the same at every point since A_k (resp. $\mathbf{PM}_N$) is affine. Note that, with the identification we made, $T_{\mathbf{PM}_N} = \mathbf{P}_{N-1} \sim \mathbb{R}^{2N}$ while TA_k is the subspace of $\mathbf{P}_{N-1}$ consisting of polynomials vanishing at ζ_k. Now, the map $\mathfrak{T}$ is tranversal to $A_k \cap U$, meaning that whenever $\mathfrak{T}(t,X) \in A_k$ we have:

$$\operatorname{Im} D\mathfrak{T}(t,X) + TA_k = T_{\mathbf{PM}_N}.$$

Indeed, $\mathfrak{T}(t,X) \notin A_k$ when $t = 0$ or $t = 1$ since $P_0, P_1 \notin A_k$, and if $t \neq 0, 1$ then the surjectivity of $DQ(t,X)$ makes $\operatorname{Im} D\mathfrak{T}(t,X)$ the full space $T_{\mathbf{PM}_N}$ already. Thus, the transversality theorem [11, Ch. 2, Sec. 3] implies that, for almost every $X \in B_\rho^N$, the partial map

$\mathfrak{T}_X : [0,1] \to \mathbf{PM}_N$ given by $\mathfrak{T}_X(t) := \mathfrak{T}(t,X)$ is transversal to $A_k \cap U$. However, since $\operatorname{Im} D\mathfrak{T}_X$ has real dimension 1 and TA_k has real dimension $2N-2$, their sum can never generate $\mathbb{R}^{2N}$. This means that for almost every $X \in B_\rho^N$ the path $\mathfrak{T}_X : [0,1] \to U$ does not meet A_k and therefore connects P_0 to P_1 in $U \setminus A_k$, as desired. □

Remark: the transversality argument used above to prove that $U \cap \mathbf{PM}_N(r)$ is connected shows, more generally, that a finite union of smooth manifolds of codimension at least 2 cannot disconnect a Euclidean open set. Because a proper complex algebraic variety decomposes into a union of this type [14, Ch. 1, Sec. 1A], we get if $Z \subset \mathbb{C}^N \sim \mathbb{R}^{2N}$ is such a variety and if $U \subset \mathbb{R}^{2N}$ is a connected open set that $U \setminus Z$ is connected. This geometrically "obvious" fact is contained in [17, Thm. 12.4.2] when U is the complement of an algebraic variety (a Zariski open set), but it is not so easy to find a reference in the present case where $\mathbf{PM}_N(r)$ is indeed proper algebraic but U may not be Zariski open.

Notations

$\mathbb{C}$	field of complex number
$\mathbb{C}^+$	open upper half-plane
$\mathbb{C}^-$	open lower half-plane
$\mathbb{T}$	unit circle
$\mathbb{D}$	open unit disk
$P(Z,\beta)$	Pick matrix associated with the sequence of interpolation data (z_k,β_k)
$\mathbb{P}_Z^+$	the set of interpolation values $\beta \in \mathbb{C}^l$ such that $P(Z,\beta) > 0$
$\mathbf{P}_N$	complex polynomials of degree at most N
$\mathbf{PE}_N$	complex polynomials of exact degree N
$\mathbf{PM}_N$	monic complex polynomials of degree N
$\mathbf{P}_{2N}^+$	non negative real polynomials of degree at most $2N$
$\mathbf{PE}_{2N}^+$	non negative real polynomials of exact degree $2N$
$\mathbf{PM}_{2N}^+$	non negative real monic polynomials of degree $2N$
$\mathbf{S}_N$	stable (no roots in $\overline{\mathbb{C}^-}$) complex polynomials of degree at most N
$\mathbf{SE}_N$	stable complex polynomials of exact degree N
$\mathbf{SM}_N$	stable monic complex polynomials of degree N
$\mathbf{SB}_N$	polynomials of degree at most N stable in the broad sense
$\mathbf{SBM}_N$	monic polynomials of degree N stable in the broad sense
$H^2(\mathbb{C}^-)$	the Hardy space of the lower half-plane
$L^2(\mathbb{R})$	the space of square integrable functions on the real line
$F^*(s) = F(\bar{s})^*$	the para-Hermitian conjugate of a rational matrix function $F(s)$
$\mathring{V}$	denotes the interior of a set V in a topological space

References

BallHorst [1] A. Ball, Joseph and ter Horst, Sanna. *Robust Control, Multidimensional Systems and Multivariable Nevanlinna-Pick Interpolation*, volume 203 of *Operator Theory: Advances and Applications*, chapter 2, pages 13–88. Birkhaüser, 2010.

BCQ15 [2] L. Baratchart, S. Chevillard, and T. Qian. Minimax principle and lower bounds in h^2-rational approximation. *Journal of Approximation Theory*, 2015. DOI: 10.1016/j.jat.2015.03.004.

BO98 [3] L. Baratchart and M. Olivi. Critical points and error rank in best H^2 matrix rational approximation. *Constructive Approximation*, 14:273–300, 1998.

BOS14 [4] L. Baratchart, M. Olivi, and F. Seyfert. Generalized Nevanlinna-Pick interpolation on the boundary. Application to impedance matching. In *Proceedings of the MTNS (Groninguen, Netherlands)*, 2014.

Belevitch56 [5] V. Belevitch. Elementary applications of scattering formalism to network design. *IRE Trans. Circuit Theory*, 1956.

BGL2001 [6] I. Byrnes, T. Georgiou, and A. Lindquist. A generalized entropy criterion for nevanlinna-pick interpolation with degree constraint. *IEEE trans. on Autom. Control*, 46(5):822–839, 2001.

Carlin-BB [7] H.J. Carlin and P.P. Civalleri. *Wideband Circuit Design*. CRC Press, 1997.

Fuhrmann [8] P. Fuhrmann. *Linear systems and operators in Hilbert spaces*. Mc Graw Hill, 1981.

Garnett [9] J. Garnett. *Bounded Analytic Functions*. Academic Press, 1981.

Georgiou1987 [10] T. Georgiou. A topological approach to Nevanlinna-Pick interpolation. *SIAM J. MATH ANAL.*, 18(5):1248–1260, 1987.

GuiPol [11] Victor Guillemin and Alan Pollack. *Differential Topology*. Prentice-Hall, 1974.

onMatchingdir [12] J.W. Helton. Broadbanding:gain equalization directly from data. *Circuits and Systems, IEEE Transactions on*, 28(12):1125–1137, 1981.

KFA69 [13] R. Kalman, P. Falb, and A. Arbib. *Topics in Mathematical System Theory*. McGraw-Hill, New-York, 1969.

Mumford [14] David Mumford. *Algebraic Geometry I: Complex Projective Varieties*. Springer, 1995.

Munkres [15] James Munkres. *Elements of algebraic topology*. Addison-Wesley, 1930.

RN [16] Frigyes Riesz and Béla Szőkefalvi-Nagy. *Functional Analysis*. Dover, 1990.

SW [17] John Sommese, Andrew and Weldon Wampler, Charles. *The Numerical Solution of Systems of Polynomials Arising in Engineering and Science*. World Scientific, 2005.