Index of Critical Points in Rational L^2 Approximation

Laurent Baratchart[1], *Martine Olivi*[1]

Abstract: This article is concerned with the l^2-approximation of a given transfer-function f by a rational one whose order is prescribed. We show that the sum of the indices of the critical points of the criterion is generically equal to 1, and in particular does not depend on f.

1 Introduction.

This paper deals with a rational approximation problem, which is related to system theory in the following way. If f is the transfer function of a l^2-stable linear constant discrete-time system driven by a white noise δ, the output $y = f\delta$ is a stationary process whose spectrum allows one, in principle, to recover f. If the latter is to be modelled by a rational function h of order at most n, and if we put $\hat{y} = h\delta$, the minimization of the covariance of $y - \hat{y}$ is achieved when the l^2-norm of $f - h$ is itself minimal. The continuous-time analogue reduces to the above, upon performing the substitution $z \to (z+1)/(z-1)$.

Rational approximation is a new trend in system identification, but the criterion under consideration here has remained relatively untouched in the literature. To our knowledge, this question is unsolved, both from computational and theoretical viewpoint. This article introduces an invariant quantity in the single-input single-output version of this problem, namely the sum of the indices of the critical points of the criterion. This result shows, for instance, that the number of critical points is generically odd. Our tools are essentially borrowed from differential topology, mixed up with a bit of classical function theory, but we could not include arguments depending on transversality theory, and still get a reasonably sized paper. The next paragraph describes the problem and sets some notations, before embarking into the proof itself.

2 The L^2 approximation problem.

Throughout, the word *system* means "single-input single-output linear constant causal discrete dynamical system", and the words *transfer function* mean the transfer function of such a system. We define the *order* of a system to be the dimension of the state-space in its minimal realizations [7]. When the order is finite, the transfer function is a proper rational function,

[1]Institut National de Recherche en Informatique et Automatique, Avenue E. Hughes Sophia-Antipolis 06560 Valbonne (France)

and the order is equal to the degree of the denominator when the fraction is in reduced form.

Let $f = \sum_{k=0}^{\infty} f_k z^{-k}$, where $f_k \in \mathbf{R}$, be the (possibly non-rational) transfer-function of a system. Assume that f is l^2-stable, that is $\sum_k f_k^2 < \infty$.

The discrete-time version of the l^2 approximation problem, in the single-input single-output case, can be stated as follows.

For any integer $n \geq 1$, find a finite-dimensional stable system, whose order is at most n, and whose transfer-function $h = \sum_{k=0}^{\infty} h_k z^{-k}$ is such that $\sum_k (f_k - h_k)^2$ is as small as possible.

If h is a best approximant of f, we must have $h_0 = f_0$ since the constant term does not affect the order. Consequently, we can always assume $f_0 = h_0 = 0$, and thus, from now on, every transfer function will be strictly proper.

It is convenient to settle this in the classical framework of real Hardy spaces (e.g. [11]). Let T be the unit circle. Let U (resp. $\bar{U}$) be the open (resp. closed) unit disk. We define H_2^- to be the *real* Hilbert space of those functions g holomorphic in the complement of $\bar{U}$, vanishing at infinity, that can be written

$$g(z) = \sum_{k=1}^{\infty} g_k z^{-k} \text{ with } g_k \in \mathbf{R} \text{ and } \|g\|^2 = \sum_k g_k^2 < \infty.$$

Now, the transfer function h of a finite dimensional system is a rational fraction which lies in H_2^- if and only if it is stable, that is if and only if its poles are in U. Hence, the set S_n^- of stable transfer-functions of order at most n is naturally included in H_2^-. Therefore, the question amounts to ask the following:

Given $f \in H_2^-$, and $n \geq 1$, find $\hat{h} \in S_n^-$ such that: $\|f - \hat{h}\| = \inf_{h \in S_n^-} \|f - h\|$

It can be proved (see e. g. [1]) that the problem stated above admits a solution. This solution is not always unique, but generically is [2]. Moreover, if we denote by Σ_n^- the set of transfer-functions of order precisely n, it is shown in [10] that any solution is in Σ_n^-, unless f itself lies in S_{n-1}^-. This fact lies at the root of a differential approach to the subject, since Σ_n^- is a differential manifold while S_n^- is not.

For the moment being, we summarize some facts concerning Hardy spaces ([11]) that will be of constant use in the sequel.

We first introduce another Hardy space H_2^+, which is symmetric in some sense to H_2^-, and consists of those functions that are holomorphic in U and can be written

$$g(z) = \sum_{k=0}^{\infty} g_k z^k \text{ with } g_k \in \mathbf{R}, \text{ and } \|g\|^2 = \sum_k g_k^2 < \infty$$

To each $g \in H_2^-$ (resp. $g \in H_2^+$), we can associate $g^* \in L^2(T)$ by putting:

$$g^* = \sum_{k=1}^{\infty} g_k e^{-ik\theta} \text{ (resp. } g^* = \sum_{k=0}^{\infty} g_k e^{ik\theta})$$

This establishes an isometry between H_2^- (resp. H_2^+) and the real subspace of $L^2(T)$ consisting of functions whose Fourier coefficients of positive (resp. strictly negative) order are zero. Hence the orthogonal sum

$$H_2 = H_2^- \oplus H_2^+$$

is isometric to the subspace of $L^2(T)$ consisting of functions whose Fourier coefficients are real, and this allows us to express the scalar product in H_2 as the one of $L^2(T)$, which in turn can be converted into a line integral

$$< f, g >= \frac{1}{2i\pi} \int_T f(z)\, g(\frac{1}{z})\, \frac{dz}{z}. \tag{1}$$

It will be convenient to introduce some more notations. For $u \in H_2$, we define

$$\check{u} = u(1/z) \quad \text{and} \quad u^\sigma = \frac{\check{u}}{z}.$$

A mechanical consequence of (1) is that

$$< uv, w >=< v, \check{u}w >$$

whenever both sides do make sense. Moreover, $u \to u^\sigma$ maps H_2^- onto H_2^+ and conversely.
We are now ready to introduce the function which will be our main object of study.

3 The function Ψ_f^n.

Let P_n be the set of real polynomials of degree at most n, and $\mathcal{P}_n^1$ the subset of monic polynomials of degree n whose roots are in U. Taking the coefficients as coordinates, P_{n-1} can be naturally identified with $\mathbf{R}^n$. Similarly, the subset of P_n consisting of monic polynomials of degree n can be identified with $\mathbf{R}^n$ as an affine space, taking as coordinates all coefficients except the leading one. This allows to consider $\mathcal{P}_n^1$ as an open subset of $\mathbf{R}^n$. Now, Σ_n^- consists of all rational fractions $h = p/q$, where

$$p = p_{n-1}z^{n-1} + p_{n-2}z^{n-2} + ... + p_0 \quad \in P_{n-1}$$

$$q = z^n + q_{n-1}z^{n-1} + ... + q_0 \quad \in \mathcal{P}_n^1$$

are coprime polynomials. Hence Σ_n^- is isomorphic to an open set of $\mathbf{R}^{2n}$, and the natural inclusion $\Sigma_n^- \to H_2^-$ is an embedding [1]. This makes Σ_n^- into a smooth submanifold of H_2^-, and the p_i's and q_j's as above are coordinates.

Let us consider the smooth function $\Gamma_f^n : \Sigma_n^- \to \mathbf{R}$, defined by $\Gamma_f^n(h) = \|f-h\|^2$. The best approximants we are looking for are among the critical points (*i.e.* those at which the derivative is 0) of Γ_f^n.
If p/q is a critical point of Γ_f^n, differentiating with respect to the p_i's yields

$$< f - \frac{p}{q}, \frac{P_{n-1}}{q} > = 0 \tag{2}$$

and with respect to the q_j's

$$< f - \frac{p}{q}, \frac{pP_{n-1}}{q^2} > = 0 \tag{3}$$

Introducing the n-dimensional linear subspace of H_2^- defined by $V_q = P_{n-1}/q$ we see that (2) means precisely that p/q is the orthogonal projection $\pi_q(f)$ of f onto V_q. On another hand, for *any* $q \in \mathcal{P}_n^1$, we may define a polynomial $L_f^n(q) \in P_{n-1}$ by the formula

$$L_f^n(q) = q\pi_q(f)$$

Because critical points of Γ_f^n are of the form $L_f^n(q)/q$, we are led to consider the map

$$\Psi_f^n \ : \ \mathcal{P}_n^1 \to \mathbf{R} \ \text{ defined by } \ \Psi_f^n(q) = \left\| f - \frac{L_f^n(q)}{q} \right\|^2,$$

and the l^2-approximation problem consists in minimizing this function.

3.1 The nature of L_f^n.

We state first a division lemma in sets of functions which are holomorphic in an open disk. It will be of constant use in the sequel.
Let U_λ denote the open disk centered at 0 of radius λ, T_λ its boundary circle. We shall denote by $H(U_\lambda)$ the space of functions holomorphic in U_λ, and by $\mathcal{P}_n^\lambda$ the set of real monic polynomials of degree n whose roots lie in U_λ.

Lemma 1 *Let $\lambda > 0$ be a real number. Let further $u \in H(U_\lambda)$ and $q \in \mathcal{P}_n^\lambda$. There exists a unique function $v \in H(U_\lambda)$, and a unique $w \in P_{n-1}$ such that*

$$u = qv + w.$$

If u and q take on real values for real arguments, the same holds true for v and w. If $\lambda = 1$ and $u \in H_2^+$, then $v \in H_2^+$.

Proof: The existence and uniqueness of v and w are a special case of Weierstrass's extended division theorem [5]. The proof, in this reference, establishes integral representation formulas for v and w as follows. Let $\mu < \lambda$ be such that $q \in \mathcal{P}_n^\mu$. Then

$$v(z) = \frac{1}{2i\pi} \int_{T_\mu} \frac{u(\xi)}{q(\xi)} \frac{d\xi}{\xi - z} \tag{4}$$

$$\text{and}\quad w(z) = \frac{1}{2i\pi}\int_{T_\mu}\frac{u(\xi)}{q(\xi)}\left[\frac{q(\xi)-q(z)}{\xi - z}\right]d\xi.$$

Taking the conjugates of the Taylor coefficients, the second assertion is an easy consequence of uniqueness. If $\lambda = 1$ and $u \in H_2^+$, then $qv \in H_2^+$. If B is the Blaschke product made from the roots of q, the factorisation theorem ([11],th.17.17) implies $qv = Bf_1$ with $f_1 \in H_2^+$. As B/q is a rational fraction with no poles in $\bar{U}$, we conclude $v \in H_2^+$, Q.E.D.

When v and w are defined as in the lemma, we call them respectively the *quotient* and the *remainder* of the division of u by q, and we put

$$v = \mathcal{Q}_q(u) \quad \text{and} \quad w = \mathcal{R}_q(u)$$

Now, let us introduce a new notation. If $r \in P_k$, we put $\quad \widetilde{r} = z^k \check{r}(z)$ It is immediate that $r \to \tilde{r}$ is an involution of P_k and that the roots (possibly infinite) of $\tilde{r}$ are the inverses of those of r. We give a word of warning about this: if r is now considered as an element of P_{k+1} whose leading coefficient is zero, the two definitions of $\tilde{r}$ do not agree. For that reason, we shall always specify which P_k r is supposed to belong to. In particular, the remainder of the division by a polynomial of degree k will be considered as a member of P_{k-1}.

We are now able to describe the nature of L_f^n.

Proposition 1 *For $f \in H_2^-$ and $q \in \mathcal{P}_n^1$, we have*

$$L_f^n(q) = \mathcal{R}_q(\widetilde{\tilde{q}f^\sigma}). \tag{5}$$

.

Proof: Let $s \in P_{n-1}$. The following sequence of equalities holds

$$< f - \frac{\mathcal{R}_q(\widetilde{\tilde{q}f^\sigma})}{q}, \frac{s}{q} > = < \frac{fq - \mathcal{R}_q(\widetilde{\tilde{q}f^\sigma})}{q}, \frac{s}{q} > = < \frac{1}{q}, \left(\overbrace{fq - \mathcal{R}_q(\widetilde{\tilde{q}f^\sigma})}^{\vee}\right)\frac{s}{q} >$$

Since multiplying by z is an isometry of H_2, the above is also equal to

$$< \frac{z^{n-1}}{q}, z^{n-1}\left(\overbrace{fq - \mathcal{R}_q(\widetilde{\tilde{q}f^\sigma})}^{\vee}\right)\frac{s}{q} >$$

$$= < \frac{z^{n-1}}{q}, (\tilde{q}f^\sigma - \mathcal{R}_q(\tilde{q}f^\sigma))\frac{s}{q} > = < \frac{z^{n-1}}{q}, \mathcal{Q}_q(\tilde{q}f^\sigma)s > .$$

By lemma 1, $\mathcal{Q}_q(\tilde{q}f^\sigma)$ is in H_2^+, and so is $\mathcal{Q}_q(\tilde{q}f^\sigma)s$, whereas z^{n-1}/q is in H_2^-. Hence, the last expression is zero. This shows that (2) is verified if we let $p = \mathcal{R}_q(\widetilde{\tilde{q}f^\sigma})$. Q.E.D.

Note that (5) and the integral representation of the remainder given by (4) show that L_f^n is a smooth function. Thereby, we get as an immediate corollary:

Corollary 1 *The map $\Psi_f^n : \mathcal{P}_n^1 \to \mathbf{R}$ is a smooth function.*

3.2 Critical points of Ψ_f^n.

Now, the partial derivatives of Ψ_f^n are computed as

$$\frac{\partial \Psi_f^n}{\partial q_i}(q) = -2 < f - \frac{L_f^n(q)}{q}, \frac{\frac{\partial}{\partial q_i}(L_f^n(q))}{q} > +2 < f - \frac{L_f^n(q)}{q}, \frac{z^i L_f^n(q)}{q^2} >$$

and since $\partial L_f^n(q)/\partial q_i$ is in P_{n-1}, it follows from the very definition of the projection that the first term on the right hand-side is zero. Hence, we get

$$\frac{\partial \Psi_f^n}{\partial q_i}(q) = 2 < f - \frac{L_f^n(q)}{q}, \frac{z^i L_f^n(q)}{q^2} > . \tag{6}$$

By (3), it is clear that the denominator of any critical point of Γ_f^n is a critical point of Ψ_f^n. Conversely, let q be a critical point of the latter. If we put $p = L_f^n(q)$, it follows from (6) and from the definition of the projection that (2) and (3) are satisfied. However, p/q will then be a critical point of Γ_f^n only if it belongs to Σ_n^-, *i.e.* only if $L_f^n(q)$ and q are coprime. *This is not always the case*, so that Ψ_f^n might have "more" critical points that Γ_f^n. In spite of the fact that these additional points have no chance to be best approximants unless $f \in S_{n-1}^-$ (though they are critical points of a lower order problem as we shall see), they will be needed in this paper to construct an invariant of the problem, namely the index.
If q is a critical point of Ψ_f^n as above, and d is the g.c.d. of p and q whose degree is k, combining linearly (2) and (3), and taking into account the fact that

$$q\,P_{n-1} + p\,P_{n-1} \;=\; d\,P_{2n-1-k}$$

by Bezout's identity, yields readily

$$< f - \frac{p}{q}, \frac{dP_{2n-1-k}}{q^2} >= \; 0.$$

If $q = dq_1$ and $p = dp_1$, this can be rewritten as

$$< f - \frac{p_1}{q_1}, \frac{P_{2n-1-k}}{dq_1^2} >= \; 0.$$

We first observe from this equation that p_1/q_1 is a critical point of Γ_f^{n-k}. On another hand, for any $s \in P_{n-1}$, we have

$$< \frac{p_1}{q_1}, \frac{\tilde{q}_1 s}{dq_1^2} >=< \frac{p_1 q_1}{q_1}, \frac{z^{n-k}s}{dq_1^2} >=< p_1, \frac{z^{n-k}s}{dq_1^2} >= \; 0$$

where the last equality holds because $p_1 \in H_2^+$ and $z^{n-k}s/dq_1^2 \in H_2^-$. We deduce from this

$$\forall i \in \{0, ..., n-1\} \quad < f, \frac{\tilde{q}_1 z^i}{dq_1^2} >= \; 0. \tag{7}$$

3.3 Extension of the domain of Ψ_f^n.

So far, Ψ_f^n has been defined only on $\mathcal{P}_n^1$. In fact, one of the advantages of considering Ψ_f^n instead of Γ_f^n is that the former can be, under mild assumptions on f, extended to a smooth function defined on a neighborhood of the closure, in $\mathbf{R}^n$, of $\mathcal{P}_n^1$. This last set, denoted by Δ_n in the sequel, consists obviously of all real monic polynomials of degree n whose roots are in $\bar{U}$. The important fact about Δ_n, which we shall prove later on, is that it is a compact manifold, thereby allowing us to use classical tools from topology. To proceed with the above-mentioned extension, it would be sufficient to assume that f is holomorphic in a neighborhood of T. However, in order to ensure further existence properties, it is convenient to work with the following class of functions: for $0 < r < 1$, define $\mathcal{H}_r \subset H_2^-$ to be the space of functions which are holomorphic for $|z| > r$ and continuous for $|z| \geq r$. If $f \in \mathcal{H}_r$, it should be noted that $f^\sigma \in H(U_{1/r})$.

Proposition 2 *If $f \in \mathcal{H}_r$, Ψ_f^n extends to a smooth function $\Psi_f^n : \mathcal{P}_n^{1/r} \to \mathbf{R}$.*

Proof: If $q \in \mathcal{P}_n^1$, the properties of the orthogonal projection show that

$$\Psi_f^n(q) = \left\| f - \frac{L_f^n(q)}{q} \right\|^2 = < f, f > - < f, \frac{L_f^n(q)}{q} > . \tag{8}$$

From lemma 1 and proposition 1, we first obtain a smooth extension of L_f^n to a map $\mathcal{P}_n^{1/r} \to P_{n-1}$ by setting

$$\widetilde{L_f^n(q)} = \frac{1}{2i\pi} \int_{T_\alpha} \frac{\tilde{q} f^\sigma(\xi)}{q(\xi)} \left[\frac{q(\xi) - q(z)}{\xi - z} \right] d\xi$$

where $\alpha < 1/r$ is an upper bound for the moduli of the roots of q. Having this at our disposal, it is now sufficient to extend smoothly $< f, z^k/q >$ to $\mathcal{P}_n^{1/r}$. This is achieved by putting

$$< f, \frac{z^k}{q} >= \frac{1}{2i\pi} \int_{T_\alpha} f^\sigma(\xi) \frac{z^k}{q(\xi)} d\xi, \ Q.E.D.$$

The next lemma shows a recursive property of Ψ_f^n, which will enable us proceed inductively in a forthcoming proof.

Lemma 2 *Let $f \in \mathcal{H}_r$ and $q \in \mathcal{P}_n^{1/r}$. Suppose $q = q_1 q_2$ where q_2 is monic of degree k, and has all its roots of modulus 1. Then $\Psi_f^n(q) = \Psi_f^{n-k}(q_1)$.*

Proof: From (8), it is sufficient to prove that

$$L_f^n(q) = q_2 L_f^{n-k}(q_1). \tag{9}$$

But, since inverse and conjugate agree on T, we have $\tilde{q}_2 = \pm q_2$ and (9) follows immediately from (5). Q.E.D.

In order to study the behaviour of the derivative of Ψ_f^n on the boundary $\partial\Delta_n$ of Δ_n, we shall need the following lemma:

Lemma 3 *Let $f \in \mathcal{H}_r$ and $q \in \mathcal{P}_n^{1/r}$ be such that $q = q_1 q_2$, where q_2 is irreducible over* $\mathbf{R}$. *Denote by v the quotient of the division of $f^\sigma \tilde{q}_1$ by q_1. If $q_2 = z - a$, then*

$$\Psi_f^n(q) = (a^2 - 1)v^2(a) + \Psi_f^{n-1}(q_1).$$

If $q_2 = (z - \xi_1)(z - \bar{\xi}_1)$, then

$$\Psi_f^n(q) = \frac{-v^2(\xi_1)(1-\xi_1^2)(1-\xi_1\bar{\xi}_1)^2 - v^2(\bar{\xi}_1)(1-\xi_1\bar{\xi}_1)^2(1-\bar{\xi}_1^2)}{(\xi_1 - \bar{\xi}_1)^2}$$

$$+\frac{2v(\xi_1)v(\bar{\xi}_1)(1-\xi_1^2)(1-\bar{\xi}_1^2)(1-\xi_1\bar{\xi}_1)}{(\xi_1-\bar{\xi}_1)^2} + \Psi_f^{n-2}(q_1).$$

Proof: The proof consists of a straightforward computation using (5), and is left to the reader.

Corollary 2 *With the same notations as in lemma 3, we have*

$$\frac{\partial}{\partial a}\Psi_f^n((z-a)q_1)_{|a=1} = 2v^2(1), \qquad \frac{\partial}{\partial a}\Psi_f^n((z+a)q_1)_{|a=1} = 2v^2(-1),$$

and if $z^2 - 2\alpha z + 1 = (z-\xi_1)(z-\bar{\xi}_1)$, $\quad \frac{\partial}{\partial \beta}\Psi_f^n((z^2-2\alpha z+\beta)q_1)_{|\beta=1} = 2|v(\xi_1)|^2.$

4 The topology of Δ_n.

Remind that the set Δ_n consists of monic polynomials of P_n whose roots are of modulus at most 1. As usual, we identify such a polynomial

$$q = z^n + z^{n-1}q_{n-1} + ... + q_0 = \prod_{i=1}^{n}(z - \xi_i)$$

with the point $(q_{n-1}, \cdots, q_1, q_0)$ of $\mathbf{R}^n$. We denote by $\|q\|_e$ the euclidean norm of this vector and by $M(q)$ the sup of the $|\xi_i|$'s. Finally, if $t \in \mathbf{R}$, we define a new monic polynomial $q * t$ of P_n by putting

$$q * t = z^n + tq_{n-1}z^{n-1} + ... + t^n q_0 = \prod_{i=1}^{n}(z - t\xi_i).$$

It is plain that Δ_0 reduces to the point 0, and Δ_1 is the segment $[-1, 1]$. It is also easy to show that Δ_2 is the triangle with vertices (-2,1),(2,1) and (0,-1). In all three cases, we see that Δ_n is homeomorphic to a ball. In fact, this result holds for any n.
We denote by B_n (resp. $\bar{B}_n$) the open (resp. closed) unit ball of $\mathbf{R}^n$, and by S_{n-1} the unit sphere.

Proposition 3 *There exist an homeomorphism $\varphi : \mathbf{R}^n \to \mathbf{R}^n$ which maps B_n onto $\mathcal{P}_n^1$, S_{n-1} onto $\partial\Delta_n$, and thus $\bar{B}_n$ onto Δ_n.*

Proof: Consider the continuous function $\varphi : \mathbf{R}^n - \{0\} \to \mathbf{R}^n - \{0\}$, which maps the poynomial q to

$$\varphi(q) = q * \lambda \quad \text{with} \quad \lambda = \frac{\|q\|_e^2}{M(q)}.$$

Now define $\mu_q : \mathbf{R}^+ \to \mathbf{R}^+$ by putting

$$\mu_q(t) = \|q * t\|_e^2 = \sum_{i=1}^{n} t^{2i} q_{n-i}^2$$

As this map is strictly increasing, it is an homeomorphism, and one can easily prove that

$$q = \varphi(q) * \nu \quad \text{with} \quad \nu = \mu_{\varphi(q)}^{-1}\left(M(\varphi(q))\right)$$

which ensures that φ is one-to-one.

We shall now show that putting $\varphi(0) = 0$ (we remind the reader that 0 is to be identified with the polynomial z^n) defines a continuous bijection $\varphi : \mathbf{R}^n \to \mathbf{R}^n$. To this effect, it is sufficient to prove that

$$\|q\|_e^2 / M(q) \to 0 \text{ when } q \to 0. \tag{10}$$

Let ξ_1 be a root of maximum modulus of q. Since the coefficients are symmetric functions of the roots,

$$\forall i \in \{1, ...n\}, \quad |q_{n-i}| \leq |\xi_1|^i \binom{n}{i}$$

and since ([8], chap.11) $|\xi_1| \leq 1 + \sum_{j=0}^{n-1} |q_j| \leq 1 + n^{\frac{1}{2}} \|q\|_e$, we have

$$\forall i \in \{1, ...n\}, \quad \frac{|q_{n-i}|}{|\xi_1|} \leq (1 + n^{\frac{1}{2}} \|q\|_e)^{i-1} \binom{n}{i}$$

which proves that $\|q\|_e / M(q)$ remains bounded when $q \to 0$, thereby implying (10). By invariance of the domain ([9], th.36.5), φ is now an homeomorphism. Q.E.D.

Though Δ_n is a topological manifold with boundary by proposition 3, it is however not smooth, i.e. its boundary $\partial\Delta_n$ has corners. The smooth part of $\partial\Delta_n$, which will play an important role in the sequel, consists of those polynomials having exactly one irreducible factor over $\mathbf{R}$ whose roots are of modulus 1. It will be denoted by F_n^1.

5 The index theorem.

Since $\partial\Delta_n$ is topologically a sphere, any continuous map $u : \partial\Delta_n \to \mathbf{R}^n - \{0\}$ induces a continuous map

$$\frac{u}{\|u\|_e} : \partial\Delta_n \to S_{n-1}$$

and we can consider its *Brouwer degree* [9], [6].

Now, if $f \in \mathcal{H}_r$ has no critical point on $\partial\Delta_n$, we define the *index* of f as the Brouwer degree of $\nabla^n_f/\|\nabla^n_f\|_e$, where ∇^n_f is the gradient vector of Ψ^n_f. So far, it is not clear that the index exists for a good deal of functions f. If we topologize $\mathcal{H}_r$ with the sup norm, it can be proved by a transversality argument [3] that it is defined on an open dense subset of $\mathcal{H}_r$.

If, moreover, the critical points of Ψ^n_f which lie in $\mathcal{P}^1_n$ are non degenerate (as before, this is the case for an open dense set of $\mathcal{H}_r$ [3]), it is a classical result that the index of f is also equal to the sum

$$\sum_{x_i}(-1)^{\epsilon_i} \tag{11}$$

where x_i ranges over the (necessarily finite) set of critical points as above and ϵ_i denote the Morse index of x_i (*i.e.* the number of negative eigenvalues of the Hessian matrix of Ψ^n_f) of x_i [4]. In particular, a critical point which is a minimum contributes by 1 to the sum.
This last property shows the relevance of the index to our study.
Now we shall establish the index theorem, which constitutes the main result of this paper.

Theorem 1 *Whenever it is defined, the index is equal to 1.*

Proof: Choose $f \in \mathcal{H}_r$ for which the index is defined. We first prove that the index does not depend on the function f. In order to do so, choose $m \geq n$, and define a vector field W^n_m by

$$W^n_m = \sum_{j=1}^{m} \nabla^n_{z^{-j}}$$

We shall prove by double induction on n and the number ν of irreducible factors of q whose roots are of modulus 1, that there is no $q \in \partial\Delta_n$ such that

$$\nabla^n_f(q) = \mu\, W^n_m(q) \quad \text{with} \quad \mu < 0. \tag{12}$$

It is easily seen that this condition is sufficient for ∇^n_f and W^n_m to be homotopic, and thus for the degrees of $\nabla^n_f/\|\nabla^n_f\|$ and $W^n_m/\|W^n_m\|$ to be equal, which will prove our contention.

Consider the smooth maps

$$\phi_1 \;:\; \mathbf{R} \times \mathcal{P}^1_{n-1} \;\to\; P_n \;\text{ given by } \phi_1(a, q_1) = (z-a)q_1,$$

$$\phi_2 \;:\; \mathbf{R} \times \mathcal{P}^1_{n-1} \;\to\; P_n \;\text{ given by } \phi_2(a, q_1) = (z+a)q_1.$$

$$\phi_3 \;:\; \mathbf{R} \times]-1,1[\times \mathcal{P}^1_{n-2} \;\to\; P_n \;\text{ given by } \phi_3(\beta, \alpha, q_1) = (z^2 - 2\alpha z + \beta)q_1.$$

Then for $q \in F^1_n$, we define $V^n(q)$ as follows:

-If $q = (z-1)q_1$, $\qquad V^n(q) = \dfrac{\partial \phi_1}{\partial a}(1, q_1)$

-If $q = (z+1)q_1$, $\qquad V^n(q) = \dfrac{\partial \phi_2}{\partial a}(1, q_1)$

-If $q = (z^2 - 2\alpha z + 1)q_1$, $\qquad V^n(q) = \dfrac{\partial \phi_3}{\partial \beta}(1, \alpha, q_1)$

Lemma 4 *If "." denotes the scalar product in $\mathbf{R}^n$, for any $f \in \mathcal{H}_r$, and $q \in F^1_n$,*

$$\nabla^n_f(q).V^n(q) \geq 0 \quad \textit{and} \quad W^n_m(q).V^n(q) > 0$$

Proof: The first inequality follows easily from the corollary to lemma 3. Moreover, if $q = q_1 q_2$ where q_2 is irreducible over $\mathbf{R}$ and has all its roots of modulus 1, the equality holds if and only if $\mathcal{Q}_{q_1}(\widetilde{q}_1 f^\sigma)$ does not vanish on the roots of q_2. Now, to see the second inequality is true, and since

$$W^n_m(q).V^n(q) = \sum_{j=1}^{m} \nabla^n_{z^{-j}}(q).V^n(q),$$

it suffices to prove that one term at least in the sum is strictly positive. But, if l is the multiplicity of the root 0 in q_1, $\mathcal{Q}_{q_1}(z^l \widetilde{q}_1)$ is a nonzero constant. Q.E.D.

Now, we come back to the proof of the theorem and assume (12) is true for some $q \in \partial\Delta_n$.

If $\nu = 1$, $q \in F^1_n$, and we see that the two inequalities of lemma 4 cannot be simultaneously satisfied. This also completes the case $n = 1$.

If $\nu > 1$, we put $q = q_1 q_2$ where q_2 is irreducible over $\mathbf{R}$ of degree k and has all its roots of modulus 1.

Consider the linear map $\theta \;:\; \mathcal{P}^{1/r}_{n-k} \to \mathcal{P}^{1/r}_n$ defined by $\theta(s) = q_2 s$. By lemma 2

$$\Psi^n_f \circ \theta \;=\; \Psi^{n-k}_f.$$

Differentiating and identifying θ with its matrix in the canonical bases of $\mathbf{R}^{n-k}$ and $\mathbf{R}^n$ yields

$$\nabla^n_f(q).\theta = \nabla^{n-k}_f(q_1).$$

Since this last equality holds for all $f \in \mathcal{H}_r$, we have $W^n_m(q).\theta \;=\; W^{n-k}_m(q_1)$, and (12) would imply

$$\nabla^{n-k}_f(q_1) \;=\; \mu \, W^{n-k}_m(q_1).$$

But q_1 has one less irreducible factor whose roots are of modulus 1 than q, which leads by induction to a contradiction and achieves the proof of the invariance of the index.

To end with, we must compute the index for some funtion $f \in \mathcal{H}_r$. Though this has been impossible so far for any precise function, we have

Lemma 5 *For every f in an open dense subset of Σ_n^-, the index is well defined and is equal to 1.*

Proof: The well definedness of the index is established in [3].
As to its value, suppose r_0/q_0 is a rational fraction in Σ_n^- for which the index is well defined. We shall prove that Ψ_{r_0/q_0} has one and only one critical point, namely r_0/q_0, which is obviously a minimum and is easily checked to be non degenerate.
Let $q = q_1 d$ be a critical point of Ψ_{r_0/q_0} in $\mathcal{P}_n^1$, where d is the g.c.d. of q and $L^n_{r_0/q_o}(q)$. Thus (7) is true and implies

$$\frac{p_0}{q_0}\frac{q_1}{\tilde{q}_1\tilde{q}} = h^+ + z^{-n}h^-$$

where $h^+ \in H_2^+$ and $h^- \in H_2^-$. Upon multiplying by q_o, we see that $q_0 z^{-n}h^-$ is at the same time in H_2^- and in H_2^+, and thus h^- is 0. As a consequence, we have

$$\frac{p_0}{q_0}\frac{q_1}{\tilde{q}_1\tilde{q}} \in H_2^+$$

which is possible if and only if $q_1 = q_0$. Q.E.D.

This achieves the proof of the theorem. Consequently, we see that if the critical points of Ψ_f^n which lie in $\mathcal{P}_n^1$ are, in addition, non degenerate, the sum in (11) is equal to 1, and the number of critical points is thereby odd. This can be proved to hold generically in $\mathcal{H}_r$ [3].

6 Concluding remarks.

The extension of the index theorem to the multi-input multi-output case needs an *ad hoc* definition for Ψ_f^n, and is currently under consideration.
The index introduced in this paper is the first invariant of the problem we know of. Using lemma 2 and the stratified structure of Δ_n, one can derive a numerical continuation method to find recursively a *local* optimum of the criterion, but we shall not discuss this here.

Bibliographie.

[1] BARATCHART L. *Existence and generic properties of L^2 approximants for linear systems*, IMA J. Math. Cont. & Information, **vol. 3**, pp. 89-101, 1986.

[2] BARATCHART L. *Recent and new results in rational L^2 approximation*, to appear in N.A.T.O. A.S.I. series.

[3] BARATCHART L. *Sur l'Approximation Rationnelle L^2 pour les Systèmes Dynamiques Linéaires*, Thèse de doctorat d'état, Universite' de Nice, septembre 1987.

[4] BORISOVICH Yu., BLIZNYAKOV N., IZRAILEVICH Ya., FOMENKO T. *Introduction to Toplology*, MIR Publishers, Moscow 1985.

[5] GUNNING R. C., ROSSI H. *Analytic Functions of Several Complex Variables* Prentice-Hall Englewood Cliffs N.J., 1965.

[6] HIRSCH M.W. *Differential Topology*, Graduate Texts in Math. Springer-Verlag New-York, 1976.

[7] KALMAN R. E., FALB P. L., ARBIB M. A. *Topics in mathematical system theory*, Mc. Graw-Hill, New York.

[8] LANG S. *Algebra* Addison-Wesley, Reading Ma., 1980.

[9] MUNKRES J.R. *Elements of Algebraic Topology* Addison-Wesley, Menlo Park Ca., 1984.

[10] RUCKEBUSH G. *Sur l'approximation rationnelle des filtres*, **Rapport** n^0 35, C.M.A. Ecole Polytechnique, 1978.

[11] RUDIN W. *Real and complex analysis*, Mc. Graw-Hill, New York, 1966.