SOME THEORETICAL LIMITS OF EFFICIENCY OF LINEAR AND NONLINEAR EQUALIZERS

Jugurta Montalvão[1,2], Bernadette Dorizzi[1], and João Cesar M. Mota[3]

[1]"Institut National des Télécommunications (INT), " EPH Department, 9, rue Charles Fourier, 91011 Evry, France.
Bernadette.Dorizzi@int-evry.fr

[2]Tiradentes University, rua Lagarto, 264, 49010-390, Aracaju, Brazil. E-mail : jugurta.montalvao@int-evry.fr.

[3]Federal University of Ceará (UFC)/CT/DEE, C.P.6001, Fortaleza, Brazil. mota@dee.ufc.br

Abstract - In this paper, we investigate some advantages and limitations of nonlinear architectures on adaptive equalization of dispersive and linear channels. The channel equalization problem is seen as one of classification, and we present some theoretical results which relate the zeros of the channel impulsive response and the geometrical dispersion of channel states. These theoretical results and some related consequences are illustrated by means of simple examples.

Resumo - Uma discussão é apresentada sobre vantagens e limites da aplicação de estruturas de equalização lineares e não-lineares sobre canais lineares e dispersivos. Nesta discussão, a equalização é vista como uma tarefa de classificação, e alguns resultados teóricos são apresentados de forma a explorar a relação existente entre os zeros do polinômio que modela o canal e a dispersão geométrica dos estados deste canal. Estes resultados teóricos e também algumas de suas consequências práticas são ilustradas por meio de exemplos.

Keywords: Nonlinear equalization, spectral nulls revisited.

1. INTRODUCTION

Linear equalizers have long been used in digital communication systems and their importance is associated to their low complexity and theoretical tractability. However, nonlinear structures like the decision feedback equalizer (DFE) or more complex feedforward devices (e.g. equalizers proposed in [1, 2, 3, 4, 5]) can outperform linear equalizers. Furthermore, Gibson et al have shown [3] that the optimum equalizer is nonlinear in all realistic cases where noise is present and the channel is non-minimum phase.

Nevertheless, despite its interesting performance, the nonlinear approach is in general strongly limited by the inherent equalizer complexity and/or the high convergence time needed by most common nonlinear filter adaptation techniques.

In [4], we have proposed a new approach to reduce nonlinear equalizer complexity using clustering methods over the channel output and a natural consequence of such work was the study of the relationship between the spectral characteristics of the channel impulsive response and the possibility of correct channel states classification (see channel state definition in Section 2). Indeed, we have found some useful theoretical results regarding the possibility of finding a classification boundary which correctly classifies the channel states.

In Section 2, we present the problem formulation and the two new theorems followed by their respective proofs. The first one states that there is no linear separation between state classes when there is an in-band spectral null. The last theorem states that when the spectral nulls are in the edge of the band, states of different classes are superposed regardless of the dimension of the classification projection space. In such a situation, even a nonlinear classification boundary may not provide a correct state classification.

In fact, despite the well known impossibility of obtaining a zero forcing equalizer when the channel presents an in-band spectral null, we emphasize that, strictly speaking, obtaining linear separation between classes is not equivalent to inverting the channel.

Section 3 presents some noticeable consequences of the theorems, while in Section 4, some illustration examples are given. Conclusions and remarks are presented in the last part of this paper.

2. EQUALIZATION AND CLASSIFICATION

In digital communication systems, where a message source emits one symbol $a(n)$ every T seconds, with the symbol belonging to a finite alphabet of S complex symbols, it is possible to associate a received signal to a corresponding symbol $a(n - d)$, where d is the decision delay. This reasoning can be extended to a sequence of synchronously sampled received signals and, in such a way, these sequences can be grouped according to their respective "labels" $a(n - d)$.

Then, in an S-ary modulation scheme, we have S classes of received sequences and, in an equalization context, it is sufficient to find the label of each received sequence to estimate the emitted symbol $a(n - d)$.

In this paper, we consider a communication scheme where digital data $a(n)$ is drawn with equal probability from a symmetrical finite alphabet $\{a_s : 1 \leq s \leq S\}$, forming an i.i.d. sequence $\{a(n)\}$, with variance σ_a^2. Furthermore, the sampled noise, $b(n)$, is additive white Gaussian with zero mean and variance σ_b^2, and $\hat{a}(n - d)$ is a decided symbol with decision delay $d \geq 0$. This means that, at each time n, the received signal $x(n)$ corresponds to an emitted symbol $a(n - d) = a_s$ possibly attenuated and corrupted by an additive white noise and by an intersymbol colored

interference[1].

Figure 1 shows a schematic representation of such a model.
Furthermore, a shorthand notation is provided by defining some column vectors:

- channel vector: $\mathbf{f} = [\ f_0\ f_1\ \cdots\ f_{N-1}\ ]^T$, where f_i, $i = 0, N-1$, are the coefficients of the FIR channel;

- noise vector: $\mathbf{b}(n) = [\ b(n)\ \cdots\ b(n-M+1)\ ]^T$, where M is the order of the equalizer;

- input channel vector:
  $\mathbf{a}(n) = [\ a(n)\ \cdots\ a(n-d)\ \cdots\ a(n-N-M+2)\ ]^T$;

- input equalizer vector:
  $\mathbf{x}(n) = [\ x(n)\ \cdots\ x(n-M+1)\ ]^T$.

We also define the channel convolution matrix:

$$
\mathbf{F} = \begin{bmatrix}
f_0 & 0 & \cdots & 0 \\
f_1 & f_0 & \cdots & \vdots \\
f_2 & f_1 & \cdots & 0 \\
\vdots & f_2 & \cdots & f_0 \\
f_{N-1} & \vdots & \cdots & f_1 \\
0 & f_{N-1} & \cdots & f_2 \\
\vdots & \vdots & \cdots & \vdots \\
0 & 0 & \cdots & f_{N-1}
\end{bmatrix}_{(N+M-1)\times(M)}
$$

Thus, we can write:

$$
\mathbf{x}^T(n) = \bar{\mathbf{x}}^T(n) + \mathbf{b}^T(n) \tag{1}
$$

where $\bar{\mathbf{x}}^T(n) = \mathbf{a}^T(n).\mathbf{F}$ is the channel state vector.
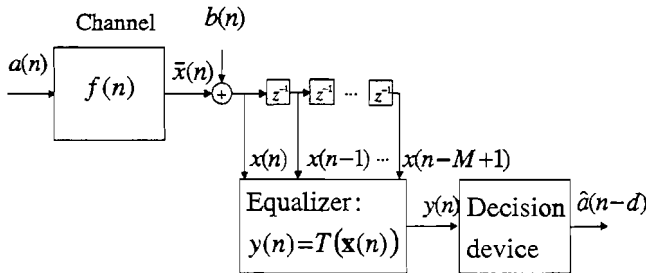

Figure 1. Simple digital communication system model with equalizer.

A classical linear equalizer is a linear projection: $y(n) = T_l(\mathbf{x}(n)) = \mathbf{x}^T(n) \cdot \mathbf{h}$, where $T_l(\cdot)$ is a linear transformation and $\mathbf{h}$ is the vector $\mathbf{h} = [\ h_0\ h_1\ \cdots\ h_{M-1}\ ]^T$, where , h_i, $i = 0, \ldots, M-1$, are coefficients of a linear transversal equalizer (LTE). In such a case, we look for the "best" projection of the samples $\mathbf{x}(n)$ onto a line in the direction of $\mathbf{h}$. Fig. 2 illustrates this search for the best direction of $\mathbf{h}$ for a BPSK modulation scheme ($a(n) = \pm 1$) and a two-taps LTE. In this illustration, the vector on the left is clearly better than the other.

[1]Actually, in Theorems 1 and 2, presented in this section, only the channel model and the symmetry of the modulation alphabet are taken into account.

In the classification sense, this corresponds to finding Fisher's linear discriminant $\mathbf{h}$ [6, 4]. So, the linear filter and the slicer, play a part of a classification device that tries to perform a linear separation between the classes corresponding to different labels $a(n-d)$ of $x(n)$.
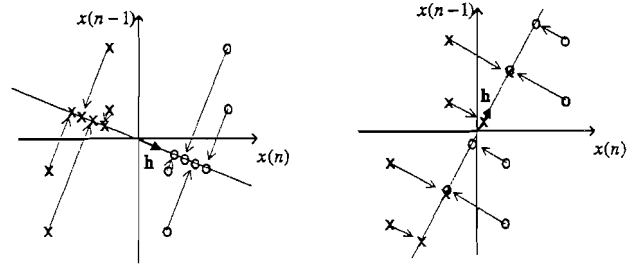

Figure 2. Projection of samples (states) on 2 different lines (two-dimensional illustration).

Clearly, for a given M and delay d, if a linear separation between classes does not exist, the linear equalizer does not have a good performance. Nevertheless, it is possible to increase the number of equalizer inputs and/or to change the decision delay in order to try to obtain the linear separation between classes. In a classification sense, changing the number of equalizer inputs corresponds to varying the dimensions of the classification space. On the other hand, changing the decision delay corresponds to changing the label of each sample. Figure 3 illustrates this by means of simple examples. In fact, Figures 3a and 3b show that linear separation may be obtained by changing M (from 1 to 2, in this case) under a fixed delay, and Figures 3c and 3d show that changes on d, under a fixed M, affect the state labels and, consequently, may also provide a linear separation. In Figures 3b, 3c and 3d the lines between the states represent one possible decision boundary.

Unfortunately, it has been observed that for some specific channels the states classes remain linearly non-separable, whatever the delay d and the equalizer order M.

In such cases, we should use a nonlinear equalizer which can provide a much more complex decision boundary. In fact, the main feature of nonlinear equalization is that, if there is no superposition of states belonging to different classification sets, it is possible to perfectly equalize such channel in a noiseless environment even if we utilize a memoryless (i.e. $M = 1$) nonlinear equalizer. Figure 4 illustrates that by showing a possible decision function in a one-dimensional state projection space for a bipolar modulation case with channel $\mathbf{f} = [\ 1\ 0.8\ 0.4\ ]^T$ and a decision delay $d = 0$.

Nevertheless, whatever equalizer structure applied (i.e. linear or nonlinear), it is worthwhile to increase its number of inputs because it provides an enlargement of distances between states of different classes and thus their separation becomes "easier". This is illustrated in Figures 4 and 5, where the minimal distance between two states of different classes increases from 0.4 to 0.6325 when M grows from 1 to 2. Figure 5 also shows possible linear and nonlinear classification boundaries.
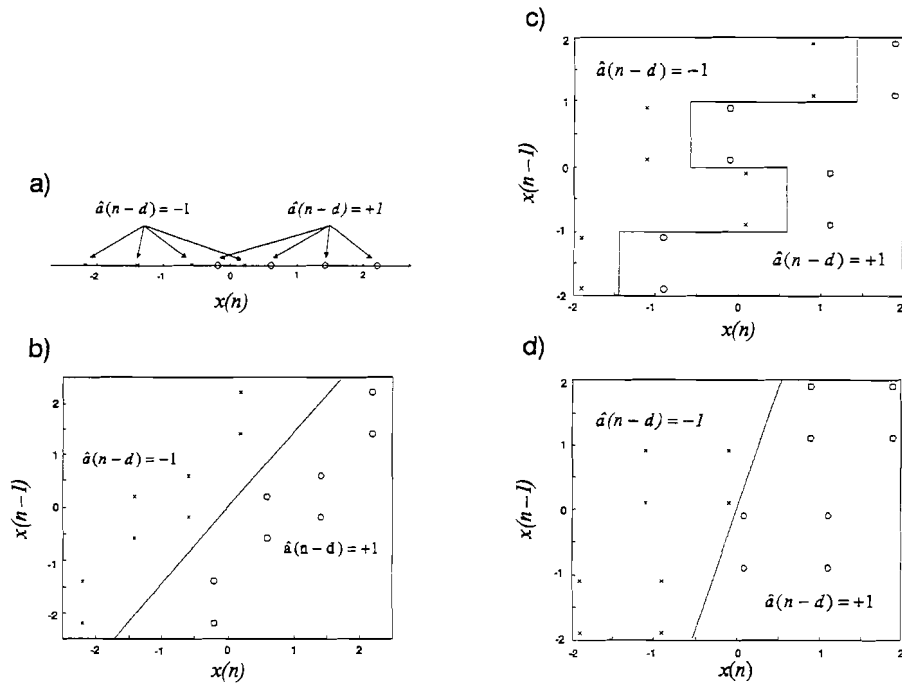
In fact, these distances between state projections also

Figure 3. Illustrations of the effect of changing M or d. Figures a and b show respectively the channel states when $f = [1\ 0.8\ 0.4]$, $d = 0$ and a) $M = 1$, b) $M = 2$. Equivalently, Figures c and d show the channel states when $f = [0.5\ 1\ 0.4]$, $M = 2$ and c) $d = 0$, d) $d = 1$.
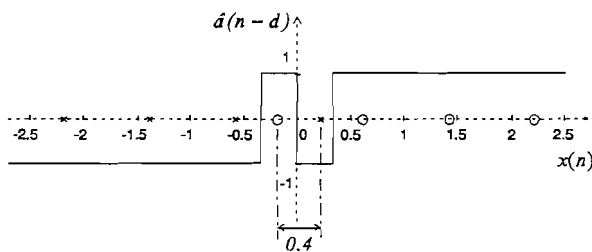

Figure 4. Decision function of an one-dimensional ($M = 1$) equalizer.

correspond to a measure of performance in the presence of noise. For instance, we can easily see that the memoryless equalizer illustrated in Figure 4 has good performance only if the inequality $P(|b(n)| > 0.2) \ll P(|b(n)| < 0.2)$ holds. That is, when the noise dispersion is confined to the between state gap. Equivalently, the condition of good performance of the two-dimensional equalizer for the same channel (see fig. 5) is $P(|b(n)| > 0.3162) \ll P(|b(n)| < 0.3162)$. Comparing the two inequalities, we can conclude that the two-dimensional (two-taps) equalizer has more "space to accommodate the noise".

Figure 6 shows a comparison of simulation results from an optimal one-dimensional Bayes equalizer, a single slicer decision device and a two-dimensional linear equalizer. This comparison illustrates that if the noise dispersion is high, the memoryless equalizer and the single slicer decision device have equivalent poor performance. Furthermore, the linear Wiener equalizer [7] with two inputs (two-dimensional state projection space) outperforms the one-

dimensional optimal Bayes equalizer in this example. This result is to be expected since the two-dimensional Wiener equalizer takes into account the correlation between every two consecutive channel outputs. Indeed, we can generalize these considerations to an M-dimensional equalizer.
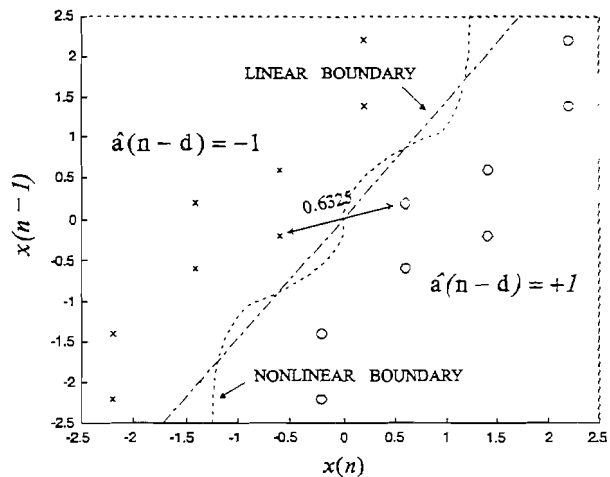

Figure 5. Linear and nonlinear decision boundaries in a two-dimensional projection state space.

In the previous examples it was sufficient to take $M = 2$ in order to obtain a linear separation between the two classes. In fact, typical communication channels impose a minimal value of M for which there is linear separation between classes. However, we have observed that there is a pathological class of channels for which there is no finite
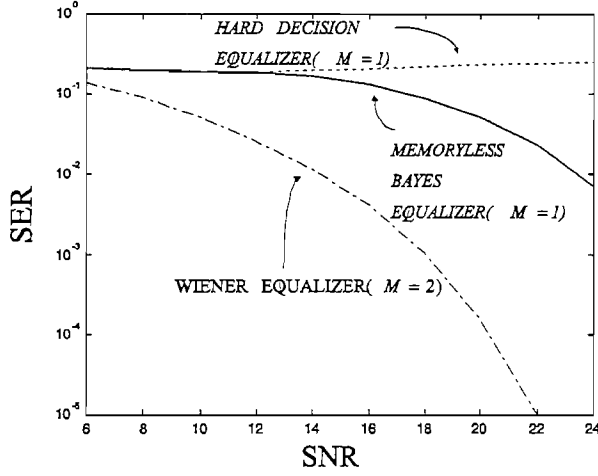
Figure 6. SER comparisons between a slicer equalizer, a nonlinear memoryless equalizer and a two-dimensional linear equalizer.

value of M or d which makes a linear separation between classes possible. Furthermore, a second class of pathological channels causes the superposition of states from different classes regardless of the values of M and d.

Since the linear equalizer provides only a linear classification boundary, the first class of pathological channels can not be equalized by a linear device. On the other hand, it is evident that even a nonlinear equalizer cannot provide a classification boundary when two states from different classes are superposed.

In order to analytically investigate such limitations concerning the linear and nonlinear equalizers, we have studied those pathological channels. That is, we have studied some particular relationships between zeros of the polynomial channel representation and their corresponding state projection features in an M-dimensional space.

Therefore, we present two new theorems:

Theorem 1: If the polynomial of a finite length channel $F(z) = \sum_0^{N-1} f_i z^{-i}$ has, at least, one zero z_u such as $|z_u| = 1$, then the state projection classes are not linearly separable for all possible delays d and for all M.

Proof: Let us consider the polynomial representation of a finite length channel:

$$F(z) = \sum_0^{N-1} f_i z^{-i} = P(z)\left(1 - z_u z^{-1}\right)$$

where we have an auxiliary polynomial:

$$P(z) = f_0 \prod_{\substack{i=0 \\ i \neq u}}^{N-2} (1 - z_i z^{-1}) = \sum_0^{N-2} p_i z^{-i}$$

So, the channel $F(z)$ can be considered as a summation of two auxiliary channels: $F(z) = P(z) - z_u P(z)z^{-1}$. Then

$$\mathbf{f} = [p_0\ p_1\ \cdots\ p_{N-2}\ 0]^T - z_u\,[0\ p_0\ p_1\ \cdots\ p_{N-2}]^T \quad (2)$$

where

$$f_0 = p_0;$$
$$f_i = p_i - z_u p_{i-1}, i = 1, 2, \ldots, N - 2$$
$$f_{N-1} = -z_u p_{N-2}.$$

Since a transversal linear filter followed by a hard decision device performs a classification by means of a hyperplan, we utilize a linear equalizer in order to try a linear separation between classes of states. Therefore, the global impulse response before the hard decision device can be calculated as a convolution:

$$\mathbf{g} = \mathbf{F} \cdot \mathbf{h} \quad (3)$$

where $\mathbf{g} = [\ g_0\ g_1\ \cdots\ g_{M+N-2}\ ]^T$ is the global impulse response vector.

Considering Eq. (2), the convolution matrix $\mathbf{F}$ can be rewritten as

$$\mathbf{F} = \underbrace{\begin{bmatrix} p_0 & & \mathbf{0} \\ \vdots & \ddots & p_0 \\ p_{N-2} & \ddots & \vdots \\ & & p_{N-2} \\ \mathbf{0} & & \mathbf{0} \end{bmatrix}}_{\mathbf{P0}} - z_u \underbrace{\begin{bmatrix} \mathbf{0} & & \mathbf{0} \\ p_0 & & \\ \vdots & \ddots & p_0 \\ & & \vdots \\ p_{N-2} & \ddots & \\ \mathbf{0} & & p_{N-2} \end{bmatrix}}_{\mathbf{P1}}$$

or, more concisely,

$$\mathbf{F} = \mathbf{P0} - z_u \mathbf{P1} \quad (4)$$

Defining $\mathbf{p0}_i^T$ as row i of matrix $\mathbf{P0}$ and $\mathbf{p1}_i^T$ as row i of matrix $\mathbf{P1}$, we have the following properties:

- Prop. 1: $\mathbf{p0}_i = \mathbf{p1}_{i+1}$ for $i = 0, 1, \ldots, M + N - 3$ or $\mathbf{p0}_{i-1} = \mathbf{p1}_i$ for $i = 1, 3, \ldots, M + N - 2$;

- Prop. 2: $\mathbf{p0}_{M+N-2}^T = [\ 0\ 0\ \ldots\ 0\ ]_{(1 \times M)}$;

- Prop. 3: $\mathbf{p1}_0^T = [\ 0\ 0\ \ldots\ 0\ ]_{(1 \times M)}$;

Now, applying (4) in (3), we have $\mathbf{g} = \mathbf{P0} - z_u\mathbf{P1h}$. Then, we can calculate each coefficient of $\mathbf{g}$ as

$$g_i = (\mathbf{p0}_i^T - z_u\mathbf{p1}_i^T)\mathbf{h} \quad (5)$$

and considering a certain decision delay d, we have

$$|g_d| = |(\mathbf{p0}_d^T - z_u\mathbf{p1}_d^T)\mathbf{h}| \quad (6)$$

In order to evaluate the maximum intersymbol interference, we define the sum $\sum_d' |g_i| = \sum_{i \neq d} |g_i|$. It is evident that if $\sum_d' |g_i| \geq |g_d|$ whatever the adopted modulation scheme, then there are at least two states from different classes that are incorrectly classified $\left(\sum_d' |g_i| > |g_d|\right)$ or superposed on the

classification boundary $\left(\sum_d' |g_i| = |g_d|\right)$. In other words, if this inequality is verified for a fixed N, M and d, it is not possible to separate the classes by means of a hyperplan in the state space. Keeping this in mind, we can investigate the case when z_u is complex and $|z_u| = 1$, regarding the coefficients of vector g. Then, from Eq. (5) and for $|z_u| = 1$ or, equivalently, $z_u = e^{j\phi}$, we have

$$|g_i| = [|\mathbf{p0}_i^T\mathbf{h}|^2 + |\mathbf{p1}_i^T\mathbf{h}|^2$$
$$-e^{j\phi}(\mathbf{p0}_i^T)^*\mathbf{h}^*\mathbf{p1}_i^T\mathbf{h} - e^{-j\phi}\mathbf{p0}_i^T\mathbf{h}(\mathbf{p1}_i^T)^*\mathbf{h}^*]^{1/2}$$

One possible way of investigation at this point consists of analyzing the maximum and minimum values of $|g_i|$. In order to do this, we calculate the partial derivative of $|g_i|$ with respect to ϕ and we make this derivative equal to zero. The maximum and minimum values of $|g_i|$ are given by:

$$|g_i|_{\max} = \sqrt{\left(|\mathbf{p0}_i^T\mathbf{h}| + |\mathbf{p1}_i^T\mathbf{h}|\right)^2} = |\mathbf{p0}_i^T\mathbf{h}| + |\mathbf{p1}_i^T\mathbf{h}|$$

$$|g_i|_{\min} = \sqrt{\left(|\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right)^2} = \left||\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right|$$

Applying these results in (6) and (5), we have $|g_d|_{\max} = |\mathbf{p0}_d^T\mathbf{h}| + |\mathbf{p1}_d^T\mathbf{h}|$. Furthermore, we can assert that $\sum_d' |g_i| \geq \sum_d' |g_i|_{\min}$. Expanding this last inequality, we have $\sum_d' |g_i| \geq \sum_d' \left||\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right|$ and, applying the triangle inequality, it follows that $\sum_d' |g_i| \geq \left|\sum_d' |\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right|$. At this point, a particular algebraic manipulation is necessary:

$$\sum_d' |g_i| \geq$$
$$\left|\sum_{i=0}^{d-1} |\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}| + \sum_{i=d+1}^{M+N-2} |\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right| \geq$$

$$\left|\left|\sum_{i=0}^{d-1} |\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right| - \left|\sum_{i=d+1}^{M+N-2} |\mathbf{p0}_i^T\mathbf{h}| - |\mathbf{p1}_i^T\mathbf{h}|\right|\right|$$

Finally, applying Prop.1-3, almost all terms are canceled and we can simplify the last inequality to:

$$\sum_d' |g_i| \geq \left||\mathbf{p1}_d^T\mathbf{h}| + |\mathbf{p0}_d^T\mathbf{h}|\right| \Rightarrow \sum_d' |g_i| \geq |g_d|_{\max}$$

Therefore, we can conclude that if the channel polynomial has at least one zero on the unitary circle, the inequality $\sum_d' |g_i| \geq |g_d|$ holds for all d or M, and the classes of states are not linearly separable □

Theorem 2: Since the modulation alphabet $\{a_s : 1 \leq s \leq S\}$ is symmetrical, if the polynomial $F(z) = \sum_0^{N-1} f_i z^{-i}$ has at least one zero $z_u \in \Re$ and $|z_u| = 1$, then there are at least S states, belonging to S different state classes, superposed at the origin of the state projection space for any d and M.
Proof: Considering again the polynomial representation of a finite length channel: $F(z) = \sum_0^{N-1} f_i z^{-i}$, we can use the same formulation as Theorem 1. Thus, the channel states are

given by $\bar{\mathbf{x}}^T(n) = \mathbf{a}^T(n)\mathbf{F}$, where each vector component is given by:

$$\bar{x}(n-j) = \left[\ a(n-j)\ \ \ldots\ \ a(n-j-N+2)\ \right]\cdot\mathbf{p}$$
$$-z_u\left[\ a(n-j-1)\ \ \ldots\ \ a(n-j-N+1)\ \right]\cdot\mathbf{p}$$

where $j = 0, 1, \ldots, M-1$ and $\mathbf{p} = \left[\ p_0\ p_1\ \ldots\ p_{N-2}\ \right]^T$. Then, when $z_u = \pm 1$, a symbol vector $\mathbf{a}(n)$ satisfying the property:

$$[\ a(n-k)\ \ldots\ a(n-k-N+2)\ ] =$$
$$\pm[\ a(n-k-1)\ \ldots a(n-k-N+1)\ ], \forall k \in \{0, \ldots, M-1\}$$

is projected on the origin of the M-dimensional state projection space. Indeed, there are at least $4S$ symbol vectors having such a property, namely:

$$\mathbf{a}_{s\pm}^T = \left[+a_s\ \ -a_s\ \ +a_s\ \ -a_s\ \ldots(-1)^{N+M-2}a_s\right];$$

$$\mathbf{a}_{s\mp}^T = \left[-a_s\ \ +a_s\ \ -a_s\ \ +a_s\ \ldots(-1)^{N+M-1}a_s\right];$$

$$\mathbf{a}_{s+}^T = [+a_s\ \ +a_s\ \ +a_s\ \ +a_s\ \ \ldots\ \ +a_s]_{(1 \times M+N-1)};$$

$$\mathbf{a}_{s-}^T = [-a_s\ \ -a_s\ \ -a_s\ \ -a_s\ \ \ldots\ \ -a_s]_{(1 \times M+N-1)};$$

where $\mathbf{a}_{s\pm} = -\mathbf{a}_{s\mp}$ and then these vectors are mapped on two overlapping states at the origin when $z_u = +1$. Equivalently, we have also $\mathbf{a}_{s+} = -\mathbf{a}_{s-}$ and these vectors are mapped on the origin when $z_u = -1$. Furthermore, for each vector pair, we can observe that $a_{s\pm} \neq a_{s\mp}$ where $a_{s\pm} = a(n-d)|\mathbf{a}(n) = \mathbf{a}_{s\pm}$ and $a_{s\mp} = a(n-d)|\mathbf{a}(n) = \mathbf{a}_{s\mp}$; and that $a_{s+} \neq a_{s-}$ where $a_{s+} = a(n-d)|\mathbf{a}(n) = \mathbf{a}_{s+}$ and $a_{s-} = a(n-d)|\mathbf{a}(n) = \mathbf{a}_{s-}$, for all possible delay $d \geq 0$. Therefore, for each symbol sequence pair, the superposed states have different labels □

3. DISCUSSIONS AND CONSEQUENCES

Theorem 1 states that a linear equalizer cannot provide a linear separation between classes if the channel has, at least, one zero on the unitary circle. In turn, Theorem 2 shows that even by use of a nonlinear equalizer, a zero error probability is not attainable when the channel has, at least, one zero equal to $+1$ or -1.

In conventional communication jargon, "the eye is open" when the classes are linearly separated. Evidently, even a channel without in-band spectral nulls may cause a closed eye but, in such cases, it is enough to change M and/or d in order to open the eye.

What is more, it is well known that a channel having spectral nulls cannot be perfectly equalized since $H(z) = 1/F(z)$ is undetermined. However, obtaining linear separation between classes is not equivalent to obtaining the zero forcing equalizer. Actually, since zero forcing equalizers recover correct symbols by means of a linear transformation, this corresponds to linearly mapping each class of states into one single point. Clearly, it is more restrictive than mapping the classes into linearly separable regions.

An interesting consequence of the results presented concerns the use of a mix strategy. That is, to split the

equalization task into two parts: a linear projection which groups the states belonging to a same class, and a nonlinear classification, which is simpler as long as the linear part does its work well.

This mix strategy is well known in the classification context but, despite its advantages concerning task simplification, our results show that it is not always a good idea to place a linear Wiener equalizer before a nonlinear one. Indeed, taking into account theorems 1 and 2, if the channel has zeros on the unitary circle, the Wiener equalizer tends to map close to the origin some states of different classes. Consequently, this reduces the mean distance between states and even a nonlinear equalizer placed after this linear equalizer will not be robust regarding the noise level.

This consequence can also be understood by analyzing the resulting polynomial transfer function: since an FIR equalizer tends to compensate for a channel zero on the unitary circle by placing several other zeros around this circle, even if $|z_u| = \pm 1$ and z_u is not real, the resulting polynomial $G(z) = F(z)H(z)$, where $H(z) = \sum_{i=0}^{M-1} h_i z^{-i}$, "tends" to have one or several zeros equal to $+1$ or -1. From Theorem 2, this can be understood as a tendency of the nonlinearly separable states to go towards the origin as long as the linear equalizer order increases.

4. ILLUSTRATION EXAMPLES

In this section, we present some examples to illustrate the former results. For simplicity, these examples are carried out with a BPSK ($a_s = \pm 1$) modulation scheme.

First, we consider the channel model $f_1 = [\ 1.000\ -1.9708\ 1.3308\ -\ 0.3600\ ]^T$, which has one zero on the unitary circle. Figure 7 shows the location of these zeros.
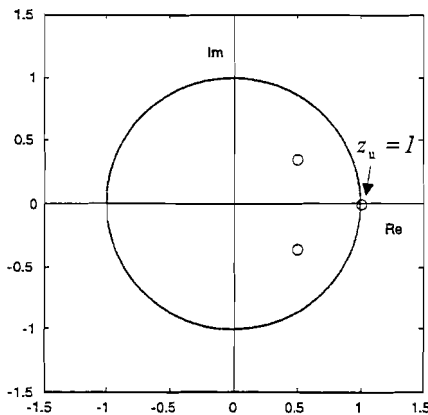

Figure 7. Complex zeros of the first channel f_1.

From Theorem 2, we have two symbol sequences: $a_{s+} = [\ 1\ 1\ 1\ 1\ ]^T$ and $a_{s-} = [\ -1\ -1\ -1\ -1\ ]^T$ such as:

$$a_{s+}^T f_1 = a_{s-}^T f_1 = 0$$

Figure 8 shows these two one-dimensional superposed states. Such states cannot be correctly classified even by an one-dimensional Bayes equalizer.

Deliberately forgetting Theorem 2, we may try to separate these states by increasing the projection space dimension M. But, according to Theorem 2, this procedure will not work well for this particular channel.
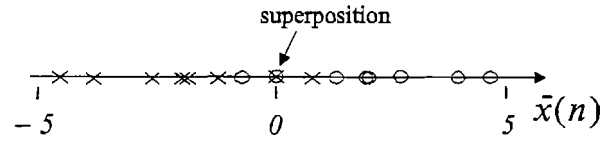

Figure 8. Superposition of two 1-dimensional states.

Just to illustrate that superposition remains even if we change the equalizer order, we show the superposed states in a two-dimensional projection by means of the convolution matrix $\mathbf{F}$:

$$\mathbf{F} = \begin{bmatrix} 1.0000 & 0.0000 \\ -1.9708 & 1.0000 \\ 1.3308 & -1.9708 \\ -0.3600 & 1.3308 \\ 0.0000 & -0.3600 \end{bmatrix}$$

and the same two symbol sequences a_{s-} and a_{s-}: $a_{s+} = [\ 1\ 1\ 1\ 1\ 1\ ]^T$, $a_{s-} = [\ -1\ -1\ -1\ -1\ -1\ ]$ and $a_{s+}^T \cdot \mathbf{F} = a_{s-}^T \cdot \mathbf{F} = [\ 0\ 0\ ]$.

Figure 9 shows the resulting two-dimensional state projection. In such a figure, we can also observe the superposition, as expected.
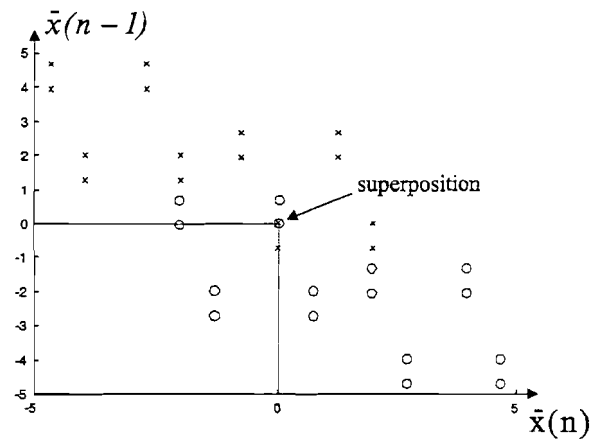

Figure 9. Superposition of two 2-dimensional states.

As a second example, we illustrate now the last consequence presented in Section 3. In this second example, we have chosen the channel model $f_2 = [\ 1.0000\ -1.6180\ 1.0000\ ]^T$, which has two complex zeros on the unitary circle. Figure 10 shows the location of these zeros.

According to Theorem 1, we know that no linear hyperplan can separate all the states with different labels. Nevertheless, just as an illustration, we "force" such a separation by means of the hyperplan corresponding to the transversal
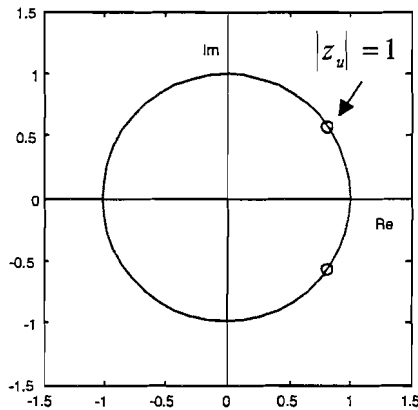
Figure 10. Complex zeros of f_2.

linear Wiener equalizer with high order. We use a 101-tap equalizer (101 has been chosen for clarity of figures presentation). Figure 11 shows the zeros of the resultant polynomial: $G_{op}(z) = F(z)H_W(z)$, where $H_W(z)$ is the Wiener equalizer. We notice that as the number of taps (M-order) increases, the zeros tend to fall on the unitary circle. Furthermore, it also implies that at least one real zero equal to $+1$ or -1 will tend to appear.

Consequently, we can conclude that if we try to apply a Wiener equalizer on a channel which comes into Theorem 1, the resulting global impulse response tends to be included in the case treated in Theorem 2, which illustrates the ideas presented in Section 3.
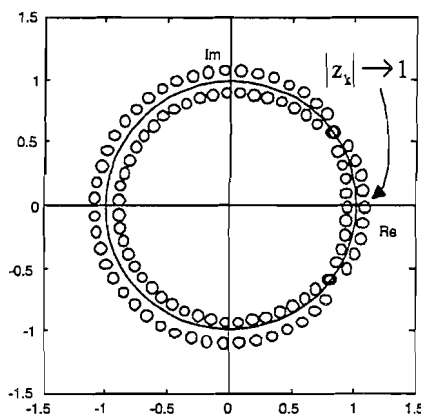

Figure 11. $G_{op}(z)$ zeros.

5. CONCLUSIONS AND REMARKS

In this paper, we have investigated some aspects of equalization regarding the polynomial zeros analysis of the corresponding channel by a classification approach.

We have introduced two theorems which relate the location of the zeros on the unitary circle to the possibility of equalization. These theorems show one theoretical limit of both linear and nonlinear equalizers.

Moreover, we have also investigated a specific

consequence of these results regarding the validity of equalizing using mixed linear and nonlinear devices.

Actually, we have used these two results in order to evaluate the performance limit of a clustering based channel identification strategy (according to [4]).

6. ACKNOWLEDGMENTS

REFERENCES

[1] S. Cheng, B. Mulgrew, and P. M. Grant, "A clustering technique for digital communications channel equalization using radial basis function networks", *IEEE Trans. on Neural Network*, vol. 4, no. 4, pp. 571–579, 1993.

[2] E. Chng, H. Yang, and W. Skarbek, "Reduced complexity implementation of bayesian equalizer using local RBF networks for channel equalization problem", *Electronics Letters*, vol. 32, no. 1, pp. 17–19, 1996.

[3] G. J. Gibson, S. Siu, S. Chen, C. F. N. Cowan, and P. M. Grant, "The application of nonlinear architectures to adaptive channel equalization", *Proc. IEEE*, pp. 0649–0653, 1990.

[4] J. R. Montalvão, J. C. M. Mota, B. Dorizzi, and C. C. Cavalcante, "Reducing bayes equalizer complexity: A new approach for clusters determination", in *ITS'98/IEEE*, 1998.

[5] J. C. Patra and R. N. Pal, "A functional link artificial neural network for adaptive channel equalization", *Signal Processing*, vol. 43, pp. 181–195, 1995.

[6] R. O. Duda and P. E. Hart, *Pattern Classification and Scene Analysis*, Wiley, New York, USA, 1973.

[7] S. Haykin, *Adaptive Filter Theory*, Prentice-Hall, Englewood Cliffs, NJ, USA, 3 edition, 1996.

[8] J. Brossier, *Signal et Communication Numérique : Egalisation et Synchronisation*, Hermes, Paris, France, 1997.

[9] J. G. Proakis, *Digital Communications*, McGraw-Hill, USA, 1984.

[10] B. Mulgrew, "Applying radial basis functions", *IEEE Sig. Proces. Magazine*, 1996.

[11] S. Theodoridis, C. M. S. See, and C. F. N. Cowan, "Nonlinear channel equalization using clustering technique", in *ICC'92/IEEE*, 1992, pp. 1277–1279.

[12] B. Dorizzi, J. C. Mota, and F. Albu, "A step towards equalization for radiomobile channels: Neural networks and variable selection", in *COST 254/IEEE*, 1997, pp. 1277–1279.

Jugurta R. Montalvão Filho graduated in Electrical Engineering at the Federal University of Paraíba (UFPB), Brazil. He also received his M.Sc. degree in Electronic and Communication from UNICAMP (Universidade Estadual de Campinas) in 1995. In 1996, he joined Tiradentes University in Aracaju, Brazil. Since 1997, he has been working as a doctoral candidate at the University of Orsay (Paris XI), and performing research at the Institut National des Télécommunications (INT), France. His research interests include neural networks, signal processing and room acoustics.

Bernadette Dorizzi was born in 1955 and got her Ph.D. (Thèse d'état) in Theoretical Physics at the University of Orsay (Paris XI-France) in 1983, on the study of integrability of dynamical systems. She is Professor at Institut National des Télécommunications since September 1989, and the head of the Electronics and Physics department where she leads the "Neural Network and Pattern Recognition" research team.

João Cesar M. Mota was born in Rio de Janeiro, Brazil, on November 17th, 1954. He received the B.Sc. degree in Physics from the Universidade Federal do Ceará (UFC), Brazil, in 1978, the M.Sc. degree from Pontifícia Universidade Católica (RJ), Brazil, in 1984, and the Ph.D. degree from Universidade Estadual de Campinas (UNICAMP), Brazil, in 1992, all in Electrical Engineering. Since August 1979, he has been with the Department of Electrical Engineering of UFC. His research interests include digital communications, adaptive filter theory and practice and signal processing. He is a member of the Sociedade Brasileira de Telecomunicações (SBrT), IEEE Comunications Society and IEEE Signal Processing Society.