The two dimensional Mazo limit

Rusek, Fredrik; Anderson, John B

*Published in:* Proc., 2005 IEEE International Symposium on Information Theory, Adelaide

*DOI:* 10.1109/ISIT.2005.1523482

*2005*

**Link to publication**


**General rights**

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

**Take down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.
The Two Dimensional Mazo Limit

Fredrik Rusek and John B. Anderson
 Dept. of Information Technology
 Lund University, Lund, Sweden
 Email: {fredrikr, anderson}@it.lth.se

Abstract—Faster Than Nyquist (FTN) signaling is extended. We send FTN pulse trains that overlap in both time and frequency; this is called two dimensional Mazo signaling. The minimum time and frequency separation that achieves $d_{\min}^2 = 2$ for root raised cosine pulses is found. Two dimensional signaling is more bandwidth efficient than one dimensional. A simple decoder is tested and it verifies the distance results.

I. INTRODUCTION

The concept of Faster Than Nyquist (FTN) signaling is well established. If a PAM system is based on orthogonal pulses, the pulses can be packed closer than the Nyquist rate $1/T$ without suffering any distance loss. In a bandpass system a QAM signal constellation is used. The rate where the square distance falls below 2. This is called the two dimensional Mazo limit. Let

$$\bar{\alpha}(\Delta) = \alpha$$

at the Nyquist rate $1/T$. Mazo showed [1] that for ideal sinc pulses the bandwidth can be reduced to .401 Hz/bit/s without any Euclidean distance loss. This corresponds to setting $T_\Delta = .802T$ instead of $T$. This value is referred to as the Mazo limit. More recently the limits for root raised cosine pulses with nonzero excess bandwidth were derived in [2]. Efficient receivers for FTN signaling were also presented in this paper for the first time. Methods of computing the minimum distance of FTN signaling can be found in [3] and [4]. Mazo-type limits can also be derived for other pulse shapes, [5]. Mazo limit phenomena turn up in other places as well. For example, it occurs in CPM, see [6] and references therein.

Instead of packing signals in time they can be stacked in frequency. The data can be divided into $K$ users; then we form $K$ signals of the type (1) and modulate each with a carrier, the carriers being separated by $f_\Delta = 2\pi f_0$. This will guarantee that the $K$ different transmissions can be decoded independently. Note that if an offset QAM multicarrier modulation system (OQAM-MCM) is used $f_\Delta$ can be set to $1/T$ without loss of orthogonality even for $\alpha > 0$, [7]. However, the possible benefits of OQAM-MCM are not investigated in this paper.

Just as FTN gave up orthogonality in time we can also give up orthogonality in frequency; instead of decreasing the symbol time separation we can decrease the symbol frequency separation $f_\Delta$. The most general strategy is to decrease both. For some combinations of time and frequency separation the minimum distance will fall below 2. These points in a two dimensional space are all referred to as two dimensional Mazo limits. Even if we only stack signals tighter in frequency we still refer to this as two dimensional to distinguish from the original Mazo limit.

The symbols can be thought of as located at points in a lattice separated by $T_\Delta$ and $f_\Delta$. This is illustrated in figure 1.

II. UPPER BOUNDS TO THE TWO DIMENSIONAL MAZO LIMIT

In this section we search for upper bounds to the two dimensional Mazo limit. Let $\bar{a}$ denote $K$ sequences of input symbols $a$, i.e. $\bar{a}$ is the set $\{a_0, \ldots, a_{K-1}\}$. These will generate $K$ signals $s_{a_0}(t), \ldots, s_{a_{K-1}}(t)$ according to (1). The baseband representation of the transmitted signal is now
formed as
\[ s_a(t) = \sum_{k=0}^{K-1} s_{ak}(t)e^{i2\pi f_k t} \]
where frequency \( f_k \) is chosen as \( f_k = kf_\Delta \). If \( f_\Delta = 2W \) the \( K \) transmissions are non overlapping in frequency.

Define the error sequence \( \overline{e} = \overline{a} - \overline{b} \) and assume that the data symbols are normalized by \( 1/\sqrt{2} \). Then the normalized Euclidean distance between the signals generated by sets \( \overline{a} \) and \( \overline{b} \) is
\[
d^2(\overline{a}, \overline{b}) = \int_{-\infty}^{\infty} |s_a(t) - s_b(t)|^2 dt
= \int_{-\infty}^{\infty} \left| \sum_{k=0}^{K-1} s_{ak-bk}(t)e^{i2\pi f_k t} \right|^2 dt
= \int_{-\infty}^{\infty} \left| \sum_{k=0}^{K-1} s_{ak}(t)e^{i2\pi f_k t} \right|^2 dt
= \int_{0}^{\infty} \left| \sum_{k=0}^{K-1} S_{ek}(f - f_k) \right|^2 df,
\]
where we have used Parseval’s identity in the last equality. It can be seen that this is not time invariant if \( f_\Delta < 2W \). Since the Euclidean distance between the signals generated by \( \overline{a} \) and \( \overline{b} \) only depends on \( \overline{a} - \overline{b} \) we write
\[
d^2(\overline{a}, \overline{b}) = d^2(\overline{a} - \overline{b}) = d^2(\overline{e}).
\]
The minimum Euclidean distance is defined as
\[
d^2_{\text{min}} = \min_{\overline{e}} d^2(\overline{e}).
\]

Our goal is now to find the smallest possible product \( T_\Delta f_\Delta \) that maintains \( d^2_{\text{min}} = 2 \). We want to find the Euclidean distance of the error event \( \overline{e} = \{e_0, \ldots, e_{K-1}\} \). Furthermore, without loss of generality, assume that both \( e_0[n] \neq 0 \) and \( e_{K-1}[n] \neq 0 \) for some \( n \). Since the signals are non overlapping in frequency in the intervals \( F_0 = [f_0 - W, f_0 - W + f_\Delta] \) and \( F_{K-1} = [f_{K-1} + W - f_\Delta, f_{K-1} + W] \) the Euclidean distance generated by \( e \) can be lower bounded as
\[
d^2(\overline{e}) = \int_{0}^{\infty} \left| \sum_{k=0}^{K-1} S_{ek}(f - f_k) \right|^2 df
\geq \int_{0}^{f_\Delta} |S_{e0}(f - f_0)|^2 df + \int_{f_0-W}^{f_0-W+f_\Delta} |S_{eK-1}(f - f_{K-1})|^2 df.
\]
\[
\geq d^2(e_0)|_{f_0} + d^2(e_{K-1})|_{f_{K-1}}.
\]
For some product of \( T_\Delta f_\Delta \) the lower bound (7) will equal 2, and \( T_\Delta f_\Delta \) is an upper bound to the two dimensional Mazo limit. By searching for the minimum \( f_\Delta \) for a given \( T_\Delta \) that fulfills
\[
\int_{f_0-W}^{f_0-W+f_\Delta} |S_{e0}(f - f_0)|^2 df \geq 1, \quad \forall e_0,
\]
we can find an upper bound to the two dimensional Mazo limit by only exhausting error events for a single transmission, i.e. \( e_k = 0, \ k > 0 \). This is true since the two terms on the right hand side of (7) have identical minimum values. Note that (8) is time invariant since there is no overlap in frequency in this frequency interval.

We collect the results of this search in table 1 for different values of \( T_\Delta \). The normalized bandwidth is calculated according to (2) and equals \( T_\Delta f_\Delta/2 \); this is for infinitely many frequency carriers. All events out to length nine have been exhausted. The value .704 is the one–dimensional Mazo limit for 30% root raised cosine pulses, resulting in nbw .4577 Hz/bit/s. This value can not be violated without having \( d^2_{\text{min}} < 2 \). It is seen that significant bandwidth reductions are possible by time and frequency stacking; the bottom row of table 1 gives a coding gain of 3.98 dB compared to 16 QAM. These, however, are only upper bounds.

<table>
<thead>
<tr>
<th>( T_\Delta )</th>
<th>( f_\Delta )</th>
<th>nbw (Hz/bit/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.78</td>
<td>.39</td>
</tr>
<tr>
<td>9</td>
<td>.8</td>
<td>.36</td>
</tr>
<tr>
<td>16</td>
<td>.86</td>
<td>.34</td>
</tr>
<tr>
<td>104</td>
<td>.97</td>
<td>.27</td>
</tr>
</tbody>
</table>

**TABLE I**

Upper bounds on the minimum frequency separations possible for different time separations. Normalized bandwidth applies for infinitely many frequency carriers.

III. Finding \( d^2_{\text{min}} \)

Finding the exact value of the minimum distance for arbitrary values of \( f_\Delta \) and \( T_\Delta \) is a hard task. If we consider error events having support \( K \) in frequency and want to exhaust all error events out to \( N \) error symbols in time we are facing \( 9^{NK} \) events in total. The trivial symmetry property
\[
d^2(\overline{e}) = d^2(-\overline{e}) = d^2(j\overline{e}) = d^2(-j\overline{e})
\]
reduces this number four fold.
In the previous section we found upper bounds to the two-dimensional Mazo limit which only required exhausting $9^N$ error events. Here we propose an algorithm that finds the minimum distance in an efficient way for an arbitrary number of frequency carriers.

The algorithm is based on the following simple lemma:

Lemma 1: Assume $f_\Delta$ and $T_\Delta$. If $d_{\text{min}}^2 < 2$ then at least one of the right-hand terms in equation (7) is smaller than 1.

Proof Assume the contrary, then $d_{\text{min}}^2 \geq 1 + 1 = 2$. ■

We can exploit Lemma 1 to find $d_{\text{min}}^2$ without exhausting all events. We propose the following algorithm which is given for $K = 3$ carriers but can be extended to any $K$:

1. Let $M_{e_0} = \{e_0 : d^2(e_0)_{|s_0} \leq 2, \text{ where } s_0 = [f_0 - W, f_0 - W + f_\Delta]\}$.

2. Let $M_{e_2} = \{e_2 : d^2(e_2)_{|s_2} \leq 2, \text{ where } s_2 = [f_2 + W, f_2 + W]\}$. This is essentially the same search as in step 1.

3. For each $e_0 \in M_{e_0}$ and $e_2 \in M_{e_2}$ such that $d^2(e_0)_{|s_0} + d^2(e_2)_{|s_2} \leq 2$ find the error event $e_1$ that minimizes $d^2(\{e_0, e_1, e_2\}) = d^2(\hat{e})$.

This algorithm exhausts less than $|M_{e_0}||M_{e_2}| 
9^N$ events, which is usually much less than $9^3 N$. The extension of this algorithm to $K > 3$ and stricter bounding criteria as well as detailed explanations of how to efficiently perform the involved steps will be reported in a future paper. Different strategies should be applied in different parts of the $(T_\Delta, f_\Delta)$ plane.

For each computation of $d^2(\hat{e})$ in step 3 of the algorithm care must be taken concerning the time variation of the Euclidean distance function. It is possible to avoid the time variation by generating an alternate signal according to

$$s_n(t) = \sum_{k=0}^{K-1} \sum_{n=-\infty}^{\infty} a_k[n]h(nT_\Delta)e^{2\pi/j\Delta t-nf_\Delta}$$

(10)

But it can be shown that this introduces an identical frequency variation of the Euclidean distance instead. Therefore we stick to our original signal generation (3) and deal with time variation.

In order to find the worst case time offset for a given error event $e$ we must compute

$$\min_{\tau_p \in \mathbb{S}} \{d^2(\tau_p(e))\},$$

(11)

where $\tau_p$ is the $p$-step time delay operator, and $s$ is given as

$$s(t) \doteq \arg \min_{\tau_p} : d^2(\tau_p(e)) = d^2(\hat{e}).$$

(12)

Note that this must be done also for a completely limited search. From eq. (4) it can be seen that $s$ also must satisfy

$$s(nT_\Delta) = u, u \in \mathbb{Z}.$$ (13)

Normally $s$ is a large number and the computational effort is high. We can choose “nice” numbers for $f_\Delta$ and $T_\Delta$ to get a small $s$ but this limits the possible choices for $f_\Delta$ and $T_\Delta$.

Instead we try to solve for the worst case analytically. The Euclidean distance of the $r$ step delayed error event $\hat{e}$ equals

$$d^2(\tau_r(e)) = \int_{-\infty}^{\infty} s_{\tau_r(e)}(t)s_{\tau_r(e[n])}(t)dt$$

$$= \int_{-\infty}^{\infty} \sum_{k,l=0}^{K-1} \sum_{m,n=0}^{N-1} e_k[m]e_l^n[n]h(t-(r+m)T_\Delta) \times h^*(t-(r+n)T_\Delta)e^{2\pi/j\Delta t(l-k)}dt$$

$$= \int_{-\infty}^{\infty} \sum_{k,l=0}^{K-1} \sum_{m,n=0}^{N-1} e_k[m]e_l^n[n]h(t-mT_\Delta) \times h^*(t-nT_\Delta)e^{2\pi/j\Delta t(l-k+r)}dt$$

$$= \sum_{k,l=0}^{K-1} \sum_{m,n=0}^{N-1} e_k[m]e_l^n[n]e^{2\pi/j\Delta t(l-k)rT_\Delta}$$

$$= \sum_{k,l=0}^{K-1} \sum_{m,n=0}^{N-1} e_k[m]e_l^n[n]e^{2\pi/j\Delta t(l-k)r\Delta T}$$

$$= \sum_{k,l=0}^{K-1} \sum_{m,n=0}^{N-1} e_k[m]e_l^n[n]e^{2\pi/j\Delta t(l-k)\Delta T}$$

where

$$\gamma_{k,l}[m, n] = \int_{-\infty}^{\infty} h(t-nT_\Delta)h^*(t-nT_\Delta)e^{2\pi/j\Delta t(l-k)}dt$$

(15)

and

$$\beta_{k,l}(\hat{e}) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} e_k[m]e_l^n[n]\gamma_{k,l}[m, n].$$

(16)

The last expression of (14) is the Euclidean distance of $e$ at time–frequency offset $\phi$; this we write as

$$d^2(\hat{e}; \phi) = \sum_{k=0}^{K-1} \sum_{l=0}^{K-1} e^{2\pi/j\Delta (l-k)}\beta_{k,l}(\hat{e}).$$

(17)

To find the worst case $\phi$ we take the derivative of $d^2(\hat{e}; \phi)$

$$\frac{\partial}{\partial \phi} d^2(\hat{e}; \phi) = \frac{K-1}{K-1} \sum_{k=0}^{K-1} \sum_{l=0}^{K-1} j2\pi(l-k)e^{2\pi/j\Delta (l-k)}\beta_{k,l}(\hat{e}).$$

(18)

For $f_\Delta \geq W$ we have $\gamma_{k,l}[m, n] = 0$ for $|l - k| > 1$; this implies that $\beta_{k,l}(\hat{e}) = 0$ for $|l - k| > 1$ as well. Setting the derivative equal to zero we get

$$\frac{\partial}{\partial \phi} d^2(\hat{e}; \phi) = j2\pi \sum_{k=1}^{K-1} e^{2\pi j\phi}\beta_{k,k-1}(\hat{e}) - j2\pi \sum_{k=0}^{K-1} e^{-2\pi j\phi}\beta_{k,k+1}(\hat{e})$$

$$= 0$$

(19)
Equation (19) is the second order second order equation

\[ (e^{\pm 2\pi \phi})^2 \sum_{k=1}^{K-1} \beta_{k,k-1}(\epsilon) = \sum_{k=0}^{K-1} \beta_{k,k+1}(\epsilon) \]  

(20)

which can be analytically solved; the case \( f_\Delta < W \) gives a fourth order equation and is omitted here. The two solutions \( \phi_1, \phi_2 \) to (20) are

\[ \{2\pi \phi_1, 2\pi \phi_2\} = \text{arg} \left\{ \pm \sqrt{\frac{\sum_{k=0}^{K-1} \beta_{k,k+1}(\epsilon)}{\sum_{k=1}^{K-1} \beta_{k,k-1}(\epsilon)}} \right\}, \]  

(21)

where \text{arg} \{ \} denotes the angle in radians of a complex number. Whether \( \phi_1 \) is the minimum or the maximum can be determined from the sign of \( \sum_{k=0}^{K-1} \beta_{k,k+1}(\epsilon) \). This method of determining the worst time offset of the error event is only valid for strictly irrational products of \( f_\Delta T_\Delta \). If rational numbers are used not all values for \( \phi \) are obtainable; we can then find the worst point simply by finding the two closest allowed points to the minimizing \( \phi_{\text{min}} \). This is illustrated in figure 2. The asterix shows \( \phi_{\text{min}} \) and the two points closest obtainable points to \( \phi_{\text{min}} \). The left point gives the smallest Euclidean distance.

\[ \begin{align*}
\text{Fig. 2}. & \quad \text{Distance of an error event \( \epsilon \) as a function of \( \phi \). The asterix indicates } \phi_{\text{min}} \text{ and the dots indicates the two possible locations for the worst case Euclidean distance.}
\end{align*} \]

In all our searches for the two dimensional Mazo limit below we have searched over error events up to length 7 in time and length 3 in frequency. An exhaustive search would involve \( 9^{21} \approx 10^{20} \) events, which is far beyond our computing capability. Our proposed algorithm, combined with the stricter bounding technique mentioned in section III, only requires testing roughly 500000 events; for the leftmost point in figure 3 it was as low as \( \approx 60000 \). The two dimensional Mazo limit is shown in figure 3. At the points marked at the solid curve the minimum distance is 2. Significant gains are obtained compared to the one dimensional case which has nbw .4577 Hz/bit/s as Mazo limit.

\[ \begin{align*}
\text{Fig. 3}. & \quad \text{Results for the 2 dimensional Mazo limit. The level curves indicates constant normalized bandwidth of (from top) .325, .3125 and .30 Hz/bit/s.}
\end{align*} \]

IV. DECODING

These normalized bandwidth gains cannot be exploited unless we are able to derive a decoder for this type of coded modulation. Here we show that decoding is indeed possible. Unfortunately the proposed decoder thus far only works for \( K = 2 - 4 \) frequency carriers; this increases the normalized bandwidths suggested in figure 3.

The decoder is based on the minimum Euclidean distance receiver. The receiver should choose as its output the signal \( s_n(t) \) that minimizes

\[ \int_{-\infty}^{\infty} |r(t) - s_n(t)|^2 dt, \]  

(22)

where \( r(t) \) is the received signal. It is well known that minimizing

\[ \min_n \int_{-\infty}^{\infty} |r(t) - s_n(t)|^2 dt \]  

(23)

is equivalent to maximizing

\[ \max_n \int_{-\infty}^{\infty} r^*(t)s_n(t) + r(t)s_n^*(t) - |s_n(t)|^2 dt. \]  

(24)

The term \( \int_{-\infty}^{\infty} r^*(t)s_n(t) + r(t)s_n^*(t)dt \) of (24) can be expressed as

\[ \begin{align*}
\int_{-\infty}^{\infty} r^*(t)s_n(t) + r(t)s_n^*(t)dt &= \\
2\Re \left\{ \int_{-\infty}^{\infty} r^*(t) \sum_{k=0}^{K-1} \sum_{n=-\infty}^{\infty} a_k[n] h(t - nT_\Delta)e^{2\pi f_k t} \right\} = \\
2\Re \left\{ \sum_{k=0}^{K-1} \sum_{n=0}^{\infty} a_k[n] \int_{-\infty}^{\infty} r^*(t) h(t - nT_\Delta)e^{2\pi f_k t} \right\} = \\
2\Re \left\{ \sum_{n=0}^{\infty} \sum_{k=0}^{K-1} a_k[n] r_k[n] \right\},
\end{align*} \]  

(25)
where
\[ r_k[n] = \int_{-\infty}^{\infty} r^*(t) h(t - n T) e^{j 2\pi f t} \, dt, \] (26)
and \( \Re \{ \cdot \} \) is the real value of a complex number. The term \( \int_{-\infty}^{\infty} |s_k(t)|^2 \, dt \) of (24) can similarly be expressed as
\[ \int_{-\infty}^{\infty} |s_k(t)|^2 \, dt = \sum_{l,k=0}^{K-1} \sum_{m,n=0}^{\infty} a_l[m] a_k^*[n] \lambda_{l,k}[m,n], \] (27)
where
\[ \lambda_{l,k}[m,n] = \int_{-\infty}^{\infty} h(t - m T) h^*(t - n T) \times e^{j 2\pi f t} e^{-j 2\pi f t} \, dt. \] (28)
Since \( \lambda_{l,k}[m,n] = \lambda_{k,l}[n,m] \) the summations in eq. (28) can be organized as
\[ \int_{-\infty}^{\infty} |s_k(t)|^2 \, dt = 29\Re \left\{ \sum_{l=1}^{K-1} \sum_{m=0}^{\infty} a_l[m] \times \sum_{k=0}^{m} \sum_{n=0}^{\infty} a_k^*[n] \lambda_{l,k}[m,n] \right\}. \] (29)

Eq. (29) together with the final expression of eq. (25) can be seen as the trellis description of the receiver; at each time \( m \) the receiver uses the current symbol of data stream \( l, a_l[m] \) and previous symbols, \( a_k[n], n \leq m, k \leq l \). This trellis is in some sense two dimensional (over both \( n \) and \( k \)). Since the number of states in this trellis is an astronomical number, MLSE is out of the question. Instead we do reduced sequence estimation (RSE) using the well known \( M \)-algorithm [6]. We traverse the trellis column by column in figure 1.

From simulations we can see that the \( M \) (list size) needed is often very large, e.g. 1000 or even more. This can be avoided by introducing a Residual Noise Canceler (RNC). This is a simple device that is cascaded on the output of the decoder which tries to improve the decoded symbol sequence. For all \( n \) and \( 0 \leq k \leq K - 2 \) it tests all possible combinations of the symbols \( a_k[n], a_k[n+1], a_{k+1}[n], a_{k+1}[n+1] \), and gives output the combination that minimizes eq. (22). This often improves the BER while decreasing the needed \( M \).

Some simulations are shown in figures 4–5. The simulations verify the obtained minimum distance values well. As previously mentioned the decoder only works for a small number of frequency carriers. In further research we are seeking a good decoder for an arbitrary number of frequency carriers to fully exploit the large gains suggested in table 1. Possible structures could be an iterative decoder as proposed in [8]. It appears that large complexity reductions are possible.

V. CONCLUSIONS

The concept of Faster Than Nyquist signaling has been generalized. We have shown that it is possible to reduce the frequency separation between the different data transmissions to less than twice the baseband bandwidth of an individual transmission. By also introducing time FTN signaling in each transmission the normalized bandwidth can be further reduced.

REFERENCES