



LUND UNIVERSITY

Improving OFDM: Multistream faster than Nyquist signaling

Anderson, John B; Rusek, Fredrik

Published in:
ITG-Fachbericht 191

2006

[Link to publication](#)

Citation for published version (APA):

Anderson, J. B., & Rusek, F. (2006). Improving OFDM: Multistream faster than Nyquist signaling. In *ITG-Fachbericht 191* VDE Verlag gmbh Berlin.

Total number of authors:

2

General rights

Unless other specific re-use rights are stated the following general rights apply:

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

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

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.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Improving OFDM: Multistream Faster-than-Nyquist Signaling

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

Abstract—Mazo’s concept of Faster Than Nyquist signaling is extended to pulse trains that modulate adjacent subcarriers, in a manner similar to orthogonal frequency division multiplex (OFDM) transmission. Despite pulses that are faster than the Nyquist limit and subcarriers that significantly overlap, the transmission system achieves the isolated pulse error performance. Systems with at least twice the spectral efficiency of OFDM can be achieved at the same error probability. Receiver design is challenging, and we report tests of several options.

Key Words: *OFDM, Coded modulation, Mazo limit, Faster than Nyquist, Bandwidth efficient coding*

I. INTRODUCTION

The subject of this paper is the ultimate limits to signaling with linear modulation. Consider baseband linear signals of the form

$$s(t) = \sqrt{E_s} \sum_{n=1}^N a[n]h(t - nT), \quad (1)$$

in which the $a[n]$ are data values over an M -ary alphabet, $h(t)$ is a unit-energy baseband pulse and E_s is the signal’s average symbol energy. This simple form underlies QAM, TCM, OFDM, and many other transmission systems. Most often, $h(t)$ is an orthogonal pulse, meaning that the correlation $\int h(t - nT)h^*(t - mT)dt$ is zero, $m \neq n$. Many signals of type (1) can be stacked in frequency through modulation by a set of subcarriers $\{f_k\}$ to form the inphase and quadrature (IQ) signal

$$s(t) = \sqrt{2E_s/T} \sum_{k=1}^K \left[\sum_{n=1}^N a_k^I[n]h(t - nT) \cos 2\pi(f_0 + f_k) - \sum_{n=1}^N a_k^Q[n]h(t - nT) \sin 2\pi(f_0 + f_k) \right] \quad (2)$$

This is a generalization to a superposition of $2K$ linear modulations, and it carries $2NK$ data values. If $f_k = kf_\Delta$, $k = 1, 2, \dots$, and f_Δ is equal twice the single-sided bandwidth B of $h(t)$, the $2K$ signals are mutually orthogonal. In OFDM-like signals both conditions hold, at least approximately: $h(t)$ is orthogonal to its own T -shifts and f_Δ is twice the bandwidth of $h(t)$.

The signal design here is based on *orthogonality*. According to classical results, there exist about $2W\tau$ orthogonal

signals in W positive Hertz and τ seconds. By means of filters matched to each one, data values that modulate the amplitude of each one can be maximum-likelihood (ML) detected independently, and therefore about $2W\tau$ symbols can be transmitted. If h in (2) is set to $\sqrt{1/T}\text{sinc}(t/T)$ and $f_\Delta = 1/T$, the product $2W\tau$ is $2(K/T)(NT) = 2KN$; this shows that Eq. (2) carries as many data values as any scheme based on orthogonality can carry. For a given number of symbols carried by (2), T may be varied, in effect trading off N and K , that is, W and τ . Only the time–bandwidth product matters, and (2) always carries twice $W\tau$ symbols. In fact there is no need for subcarriers since (1) alone achieves $2W\tau$ by taking $T = 1/2W$, $K = 1$ and $N \approx \tau/T$ symbols.

If the aim is to achieve the *error rate* of a stacked orthogonal-signal system (2), without necessarily using orthogonal signals, the story is more interesting, and that is our subject in this paper.

Before continuing, let us be more precise about the measurement of error and bandwidth. For (1), the ML error probability depends on h : Asymptotically in the signal-to-noise ratio E_b/N_0 , the probability of incorrect detection of an $a[n]$ in additive white Gaussian noise (AWGN) with density $N_0/2$ is $P_e \sim Q(\sqrt{d_{\min}^2 E_b/N_0})$, where $d_{\min} \leq d_{\text{MF}}$. If the K signal pairs in (2) do not overlap in frequency, the same applies there. Here E_b is the bit energy $E_s/\log_2 M$ and d_{MF} is the matched-filter bound. The last measures the performance of orthogonal-pulse signaling with the same alphabet; for binary transmission $d_{\text{MF}}^2 = 2$. The paper will concentrate on the binary case, so the error rate to be achieved is $Q(\sqrt{2E_b/N_0})$. As for signal bandwidth, it is well known that for uncorrelated data symbols the power spectral density $S_k(f)$ of the k th subcarrier is proportional to $|H(f - kf_\Delta - f_0)|^2 + |H(f + kf_\Delta + f_0)|^2$. The normalized bandwidth is measured by

$$\text{NBW} \triangleq \frac{W}{R} \quad \text{Hz/bit/s}, \quad (3)$$

where W is some measure of the positive frequency bandwidth of the entire transmission (2) (such as 99% power bandwidth) and $R = 2K/T_\Delta$ is the data rate of the entire signal in bit/s. For eq. (1), (3) gives a consistent measure if W is the positive baseband bandwidth and $K = 1$.

The idea that more can be achieved with (1) at the same error probability was proposed by Mazo [1] in 1975. In

a technique he called faster-than-Nyquist (FTN) signaling, binary-modulated $\text{sinc}(t/T)$ pulses with bandwidth $1/2T$ Hz appear once each T_Δ , where $T_\Delta < T$; this is faster than $1/T$, the Nyquist limit to orthogonal pulse trains with that bandwidth. A full ML sequence detection is now required, which in principle compares all N -symbol signals to the noisy received signal. By finding the minimum distance d_{\min} of this signal set (see Section II), one can estimate the asymptotic symbol P_e as $\sim Q(\sqrt{d_{\min}^2 E_b/N_0})$. Mazo and later papers showed the surprising result that d_{\min}^2 is in fact $d_{\text{MF}}^2 = 2$ for $T_\Delta/T > .802$; that is, nothing is lost asymptotically by increasing the symbol rate 24.7%.

The fundamental reason for this is that as the pulse rate grows another signal difference (or “error event”) eventually has a distance less than the square-distance 2 antipodal event. A similar phenomenon occurs with other orthogonal $h(t)$ than the sinc pulse; see [2] for the root RC pulse. Moreover, it often appears with both linear and nonlinear coded modulations; see [3] for the case when $h(t)$ is a Butterworth filter response and [5], Chapter 6, when the coded modulation is CPM. In all these cases, the wideband error performance is unchanged under filtering until a surprisingly narrow bandwidth, after which it suddenly drops. We will call this threshold bandwidth the *Mazo limit*. Its significance is that there is no point transmitting in a wider bandwidth in a linear channel with AWGN, if sufficient receiver processing is available.

Mazo signaled too fast in time, but in a subcarrier system one can also signal too widely in frequency. Now the signal is (2) but the subcarriers overlap in frequency and cannot be separated by filtering. Yet one can hope, as Mazo did, that P_e remains $\sim Q(\sqrt{2E_b/N_0})$. We have introduced this idea in [4]. It was called two-dimensional Mazo signaling there because the symbols can be associated with points in a lattice spaced every f_Δ and T_Δ . This is illustrated in Fig. 1. Ref. [4] shows

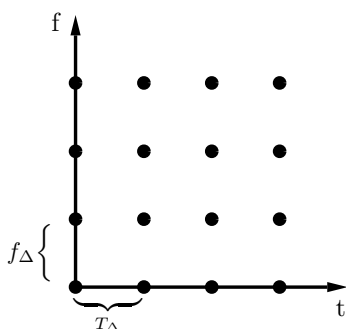


Fig. 1. Two dimensional Mazo signaling, in time and frequency. Dots represent symbols separated by f_Δ and T_Δ .

that simultaneous frequency and time squeezing can indeed increase the symbols transmitted in a given time–bandwidth at the same P_e , in a way that neither compression alone can accomplish.

At first glance it may seem that two-dimensional squeezing can add nothing to FTN time compression, since an ordinary OFDM system easily trades symbols per subcarrier against the number of subcarriers. But the subcarriers in that case are acting as independent linear modulations. With f_Δ less than

the bandwidth of a subcarrier, Eq. (2) is not linear modulation. The signal interrelations that produce d_{\min} work in new ways when both f_Δ and T_Δ can be varied independently. The plan of the rest of the paper is to justify this by measuring distance in Section II. It will turn out that with many subcarriers the spectral efficiency of orthogonal-pulse OFDM can be nearly doubled. Section III explores a receiver possibility.

II. DISTANCE UNDER TIME AND FREQUENCY COMPRESSION

In this section we will review how to calculate the parameter d_{\min} and thus locate the Mazo limit; then we will look at what happens when N and K are large. First we need to normalize the parameters f_Δ and T_Δ that affect d_{\min} and show how they set the time–bandwidth product.

Let the normalized pulse time compression be $T'_\Delta \triangleq T_\Delta/T \leq 1$. Let the normalized frequency compression be $f'_\Delta \triangleq f_\Delta/(1/T) = f'_\Delta T$; this can be viewed as the compression relative to sinc-modulated orthogonal subcarriers. For a system (2) with a practical pulse $h(t)$, the signal occupies time $NT_\Delta + \epsilon$ and bandwidth $Kf_\Delta + \delta$, where ϵ and δ are the extra time and bandwidth margin that a practical pulse needs at times 1 and N and subcarriers 1 and K . There are $2NK$ data symbols, and on a per-symbol basis, the time–bandwidth consumption is

$$W\tau \Big|_{\text{symbol}} \approx \frac{1}{2} f_\Delta T_\Delta = \frac{1}{2} (f_\Delta/T) (T_\Delta T) = \frac{1}{2} f'_\Delta T'_\Delta \quad (4)$$

as both N and K become large. The product $f'_\Delta T'_\Delta$ (or $f_\Delta T_\Delta$) is the fractional reduction in time–bandwidth consumption compared to $1/2$, the least possible value for strictly orthogonal signaling.

The normalized distance between two signals $s^{(1)}(t)$ and $s^{(2)}(t)$ of form (2) is

$$d^2 = \frac{\int [s^{(1)}(t) - s^{(2)}(t)]^2 dt}{2E_b} \quad (5)$$

where E_b is the average energy per data bit. Because of the linear form of (2), only the difference between the symbol streams matters. We can define a matrix of I-signal differences as

$$\Delta \mathbf{A}^I = \begin{bmatrix} a_1^I[1] & a_1^I[2] & \dots & a_1^I[N] \\ \dots & \dots & \dots & \dots \\ a_K^I[1] & a_K^I[2] & \dots & a_K^I[N] \end{bmatrix}$$

and similarly for the Q-signal differences $\Delta \mathbf{A}^Q$. The square distance is then the integral of a squared signal (2) with $\Delta \mathbf{A}^I$ and $\Delta \mathbf{A}^Q$ in place of the respective I and Q data values. For binary signaling, the matrix components $\in \{-2, 0, +2\}$. An *error event* is a region of nonzero components that begins at some position (n, k) . The minimum distance is the minimum of (5) over all such events. Finding d_{\min} is difficult but possible. Typically, searching over regions of size 5×3 is sufficient; the distance for a given event pattern depends on the time start n , but not on the starting subcarrier k . Distance may be computed by direct integration of the difference signal, but

a much more efficient method is based on the autocorrelations of $h(t)$. All of these matters, together with useful bounds on d_{\min} , are discussed in ref. [4].

One can compute d_{\min} and the precise time and frequency occupancy for N pulses and a small number of subcarriers. This is explored in [4]. Here we concentrate on the per-symbol time–bandwidth product (5) when N and K are large. We search for the combination of f'_{Δ} and T'_{Δ} that gives the smallest product, while attaining $d_{\min}^2 = 2$. This gives an ultimate time–bandwidth limit for the OFDM-like FTN system.

Figure 2 assumes the root RC pulse with excess bandwidth $\alpha = .3$ and shows the location of the best allowed $f'_{\Delta}T'_{\Delta}$ product at the Mazo limit, as a function of the normalized time compression T'_{Δ} . Explanation of the figure helps to show

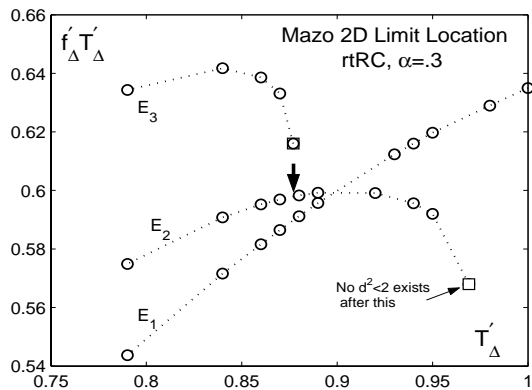


Fig. 2. Trajectories of $f'_{\Delta}T'_{\Delta}$ vs. T'_{Δ} over the range $[.79, 1]$ for three families of critical error events; root RC pulse, $\alpha = .3$. The maximum of the three curves is the least possible time–bandwidth product.

how the limit comes about. The different error events form event families whose members, due to I/Q symmetries, lead to identical distance behavior. There are three families that affect the binary root RC $\alpha = .3$ Mazo limit, denoted in the figure E_1, E_2, E_3 . One of the members of E_1 is for example

$$\Delta \mathbf{A}^I = \begin{bmatrix} -2 & 0 \\ 2 & -2 \\ 0 & -2 \end{bmatrix}; \quad \Delta \mathbf{A}^Q = \begin{bmatrix} 0 & -2 \\ 2 & -2 \\ -2 & 0 \end{bmatrix}$$

As T'_{Δ} increases, there is a least f'_{Δ} that allows $d_{\min}^2 = 2$ for family E_1 ; the product of this with T'_{Δ} is plotted in Fig. 2. Two other families affect the Mazo limit. The *maximum* of these curves for each T'_{Δ} is the minimum allowed time–bandwidth $f'_{\Delta}T'_{\Delta}$ at the Mazo limit for this root RC case. In the figure, the least time–bandwidth for any $T'_{\Delta} \in [.79, 1]$ is shown by the heavy arrow: It lies at $\approx .598$ Hz-s, with time compression $T'_{\Delta} \approx .875$ and frequency compression $f'_{\Delta} \approx .683$.

From the foregoing it can be seen that there is only one free parameter, namely T'_{Δ} . The Mazo constraint then sets the time–bandwidth. Almost all event families have $d^2 > 2$ over the T'_{Δ} range and do not play a role. For $\alpha = .3$ it is interesting to observe that Mazo’s original FTN limit for time compression alone occurs at $T'_{\Delta} = .704$. For root RC pulses

with $\alpha = .1, .2$, the best $f'_{\Delta}T'_{\Delta}$ product in the range $[.79, 1]$ is $\approx .555$ and $\approx .587$, respectively. RC (not root) pulses perform similarly. Gaussian pulse behave poorly. Sinc pulses have time–bandwidth no higher than $.5$, since $.5$ is attained for frequency compression alone, with $T'_{\Delta} = 1$.

III. DECODING

Decoding of this type of coded modulation is complex. Full sequence estimation grows exponentially with the number of subcarriers K , so MLSE decoding is ruled out. We have two desires for the decoder. The first is to obtain close to MLSE error performance; the second is to heavily reduce complexity. A decoding algorithm with major complexity reduction but only minor error loss is the so called M -algorithm, [5]. This method was tested in [4], but it only worked well for 2–4 subcarriers, since otherwise it was not clear in what order the symbols should be decoded.

In this paper we test an iterative method based on one dimensional BCJR-algorithms and soft interference cancellation (SIC). A model is shown in Fig. 3.

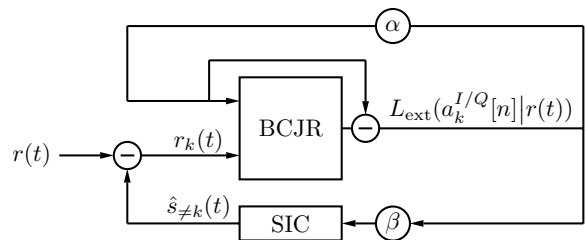


Fig. 3. System model for the decoder. The block SIC produces a signal based on estimates for all symbols except for symbols on row k . The block BCJR is a normal BCJR algorithm.

The task of the decoder is to put out estimates of the transmitted bits such that the *a posteriori* probability (APP) of an individual bit is maximized, i.e.

$$\hat{a}_k^{I/Q}[n] = \arg \max_{a \in \{-1, 1\}} P\{\hat{a}_k^{I/Q}[n] = a | r(t)\}, \quad (6)$$

where $r(t)$ is the received signal. Here and throughout, superscript I/Q means “I respectively Q”. Instead of working with probabilities it is convenient to work with log likelihood ratios (LLRs) :

$$L(a_k^{I/Q}[n]) = \frac{P\{a_k^{I/Q}[n] = 1\}}{P\{a_k^{I/Q}[n] = -1\}} \quad (7)$$

Since the data symbols are independent we can as usual express the conditional LLR $L(a_k^{I/Q}[n]|r(t))$ as

$$L(a_k^{I/Q}[n]|r(t)) = L_{\text{ext}}(a_k^{I/Q}[n]|r(t)) + L(a_k^{I/Q}[n]) \quad (8)$$

where $L_{\text{ext}}(a_k^{I/Q}[n]|r(t))$ denotes the extrinsic information about $a_k^{I/Q}[n]$ contained in $r(t)$.

The true APPs of the data bits can be found by a multidimensional BCJR algorithm, but as with MLSE the complexity grows exponentially with K , and the APPs have to be approximated with reduced complexity. One way to perform this is by

an iterative method. In each iteration of the decoding process an estimate of the signal based on all symbols *except* subcarrier k is formed. This estimate, denoted $\hat{s}_{\neq k}(t)$, is

$$\sqrt{\frac{2E_s}{T_\Delta}} \sum_{l=1, l \neq k}^K \left[\sum_{n=1}^N b_l^I[n] h(t - nT_\Delta) \cos 2\pi(f_0 + f_l) - \sum_{n=1}^N b_l^Q[n] h(t - nT_\Delta) \sin 2\pi(f_0 + f_l) \right]$$

where $b_l^{I/Q}[n]$ are the soft estimates of $a_l^{I/Q}[n]$, i.e.

$$b_l^{I/Q}[n] = P\{a_l^{I/Q}[n] = 1\} - P\{a_l^{I/Q}[n] = -1\}$$

The tentative received signal for subcarrier k is formed as

$$\hat{r}_k(t) = r(t) - \hat{s}_{\neq k}(t)$$

Together with with the extrinsic information about the bits on subcarrier k , $L_{\text{ext}}(a_k^{I/Q}[n]|r(t))$, the signal $\hat{r}_k(t)$ is fed to a standard BCJR algorithm.

Note that the extrinsic information fed to the SIC and to the BCJR are first attenuated by coefficients α and β . What values to use for α and β is however not clear and they have to be determined by simulation. We have used slowly increasing functions over the iterations for both α and β . Furthermore we have found it beneficial to set $\alpha = 0$ in the last iteration, i.e. only account for the extrinsic information via the SIC. We believe that this is due to suboptimal functions for α and β .

We do not assume the Forney observation model, i.e. a whitening filter, and instead work directly on the outputs of the pulse-matched filter. The samples can be written as

$$y_k^{I/Q} = \sum_{m=-\infty}^{\infty} g_m a_k^{I/Q}[m-l] + \eta_m + \nu_k$$

where

$$g_m = \int_{-\infty}^{\infty} h(t)h(t - mT_\Delta)dt$$

and η_k are noise samples of a filtered AWGN process having autocorrelation $R_\eta(n) = 2N_0g_n$. The term ν_k is noise emanating from the fact that $\hat{s}_{\neq k}(t)$ is not a perfect estimate. It is usually assumed to be Gaussian. Since the BCJR requires the noise variance as input, $\text{var}(\nu_k)$ is estimated as

$$\sigma_\nu^2 = \frac{1}{N-1} \sum_{n=1}^N (y_n^{I/Q})^2 - E\{a_k^2[m]\} \sum_{n=-\infty}^{\infty} g_n - 2N_0$$

Since σ_ν^2 can be below zero, one takes $\max(0, \sigma_\nu^2)$ as estimate of $\text{var}(\nu_k)$.

A BCJR algorithm for the matched filter model is derived in [6] and has the same structure as the one for the Forney model. We set the ISI length to 5, i.e. $g_m = 0$, $m > 6$, and treat the rest as Gaussian noise included in η_k .

Such a simple receiver seems to work well for $f'_\Delta T'_\Delta \geq .7$, especially if f'_Δ is large. We tested $f'_\Delta = 1$ and $T'_\Delta = .7$, a product of .7, with 20 subcarriers. The receiver test is shown in figure 4. After 10 iterations, the receiver obtained BER 10^{-5}

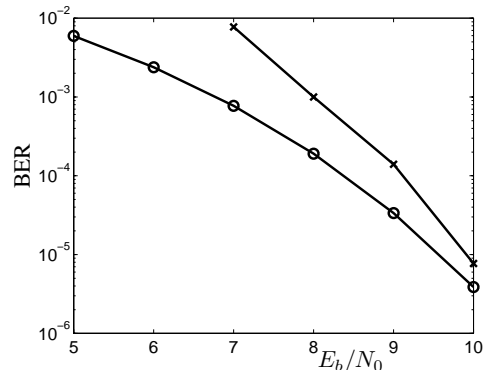


Fig. 4. Receiver test. \circ = uncoded BPSK; \times = performance of the iterative decoder after 10 iterations.

at roughly 10 dB, .5 dB from $Q(\sqrt{2E_b/N_0})$.

For smaller products than .7 the receiver fails to work well. In the literature decoding methods have been proposed for two dimensional ISI problems, e.g. the multistrip method from [7]. However, we are dealing with higher order modulations and much worse ISI. Future research will study multidimensional reduced-complexity BCJR which hopefully will reduce σ_ν^2 in the first iteration, and thereby push down the decodable $f'_\Delta T'_\Delta$ products to the theoretical limit.

IV. CONCLUSIONS

The concept of faster than Nyquist signaling has been generalized to subcarrier OFDM-like systems. Much lower frequency separation between subcarriers than the subcarrier bandwidth can be used without loss of error probability. By also introducing time-FTN signaling in each subcarrier the bandwidth per data symbol can be further reduced. The time and frequency compressions where the signal set minimum distance first falls below 2 are called the two dimensional Mazo limit. It turns out that even for a low number of frequency carriers this limit is lower than the time-only Mazo limit. With many carriers, spectral savings of at least 50% are possible in principle. A sample decoder was tested to verify the method.

REFERENCES

- [1] J.E. Mazo, "Faster-than-Nyquist Signaling," *Bell Syst. Tech. J.*, vol. 54, pp. 1451–1462, Oct. 1975.
- [2] A.D. Liveris and C.N. Georghiadis, "Exploiting Faster-Than-Nyquist Signaling," *IEEE Trans. Commun.*, vol. 51, pp. 1502–1511, Sep. 2003.
- [3] F. Rusek and J.B. Anderson, "M-ary Coded Modulation by Butterworth Filtering," *Proc.*, 2003 Int. Symp. Information Theory, Yokohama, p. 184, June 2003.
- [4] F. Rusek and J.B. Anderson, "The Two Dimensional Mazo Limit," *Proc.*, 2005 Int. Symp. Information Theory, Adelaide, pp. 970–974, Sept. 2005.
- [5] J.B. Anderson and A. Svensson, *Coded Modulation Systems*, Plenum/Kluwer, New York, 2003.
- [6] G. ColaVolpe and A. Baerbieri, "On MAP Symbol Detection for ISI Channels Using the Ungerboeck Observation Model", *IEEE Commun. Letters*, vol. 9, Aug. 2005.
- [7] M. Marrow and J.K. Wolf, "Iterative Detection of 2-Dimensional ISI Channels," in *Proc.*, 2003 IEEE Intern. Workshop Inform. Theory, Paris, pp. 131–134, April 2003.