# Adaptive Equalization For Partially Bandwidth-Occupied ADSL Transceivers

X. F. Wang, W.-S. Lu, and A. Antoniou
Department of Electrical and Computer Engineering
University of Victoria
Victoria, B.C., CANADA, V8W 3P6
Email: {xfwang, wslu, andreas}@ece.uvic.ca

**Abstract**— A new optimization criterion for the design of a time-domain equalizer in asymmetric digital subscriber line systems is studied. It is shown that the new criterion maximizes the signal-to-noise ratio regardless of the profile of bandwidth occupancy. Based on the fact that the solution of the optimization problem is unique, two adaptation algorithms are developed in the time and frequency domains, respectively. The computational complexities of these two algorithms are comparable to that of the conventional least-mean-square algorithm. Simulation results show that the two algorithms converge at a satisfactory speed but the frequency-domain algorithm offers a slightly faster convergence and a smaller excess mean-square error.

## I. Introduction

Discrete multitone (DMT) was selected as the standard line-coding method for asymmetric digital subscriber line (ADSL) systems [1]. For a DMT-based ADSL transceiver over a highly dispersive channel such as a typical telephone line, the major signal impairments are intersymbol interference (ISI) and inter-carrier interference (ICI) [2]. These impairments can be avoided by appending a cyclic prefix whose duration is the same as that of the channel impulse response (CIR) to each symbol. Since the CIR duration of a typical telephone line is comparable to the symbol length that is allowed by the permissible transmission delay, the channel efficiency would be significantly degraded by adding such a long cyclic prefix. To avoid this problem, a time-domain equalizer (TEQ) is often employed to shorten the CIR such that a cyclic prefix of much reduced duration can be used [2].

The most popular criterion for designing a TEQ is the mean-square error (MSE), which was originally proposed in [3] for systems where the entire bandwidth is occupied by transmitted signals and is now widely applied to ADSL systems [4][5][6]. One of the important advantages of this criterion is that it allows a simple adaptive implementation [4]-[6]. However, in many ADSL systems only a portion of the Nyquist bandwidth is used and, for such systems, the MSE criterion does not maximize the signal-to-noise ratio (SNR). To overcome this problem, we propose a new criterion that maximizes the SNR regardless of bandwidth occupancy. The corresponding TEQ can be obtained by minimizing the MSE subject to the constraint that the norm of the target impulse response (TIR) at the occupied frequency bands be a constant.

To compute the coefficients of the proposed TEQ, the CIR is assumed to be known. However, in practice the CIR is often unknown to the receiver and, more importantly, it may vary slowly. Hence, an adaptive implementation of the proposed TEQ, which does not require knowledge of the CIR, is desirable. In this paper, we study the adaptive implementation of the proposed TEQ. Note that the conventional least-mean-square (LMS) algorithm does not apply to the proposed TEQ owing to the nonlinear constraint added in the optimization problem although it is applicable to the MSE-based TEQ. Based on the fact that the constrained optimization problem has a unique global minimum, we first propose a steepest-descent algorithm for computing the coefficients of the new TEQ. The convergence of the algorithm is guaranteed for an appropriately selected step size. Based on the steepest-descent algorithm, we then develop a stochastic gradient algorithm in the time domain for TEQ adaptation. To avoid the slow convergence and large excess MSE of the time domain implementation, we further develop a frequency-domain adaptation algorithm. This algorithm fully utilizes the fast Fourier transform (FFT) and inverse FFT computing resources in an ADSL modem and, as a result, offers a computational complexity comparable to that of an adaptive implementation of the MSE-based TEQ [5]. Simulations are carried out to compare the convergence of both the time-domain and frequency-domain algorithms.

## II. A New TEQ

The objective of the TEQ, whose block diagram is illustrated in Fig. 1, is to shorten the CIR such that the equalized CIR is equal to a TIR of length $v$ or less, where $v - 1$ is the length of the cyclic prefix. If this objective is achieved, there will be no ISI and ICI. In practice, however, this objective cannot in general be achieved by a TEQ of short length and residual ICI and ISI exist. A popular criterion for the design of a TEQ is the MSE criterion. The corresponding solution is a pair of $\mathbf{w}$ and $\mathbf{b}$ that minimizes the MSE, $E[e_k^2]$, under the constraint that $||\mathbf{b}||$ be a constant. Evidently, if the signal spectrum is flat over the entire Nyquist bandwidth, the MSE criterion leads to the maximum SNR. However, in ADSL systems often the lower portion of the spectrum is reserved for other sevices and portions of the bandwidth are not usable due to channel nulls. As a result, an ADSL transceiver usually does not occupy the entire bandwidth. In this case, the MSE criterion does not maximize the SNR. In the light of this, we propose a new TEQ that maximizes the SNR

$$\mathrm{SNR} = \frac{E[z_k^2]}{E[e_k^2]} \tag{1}$$

regardless of the profile of bandwidth occupancy.

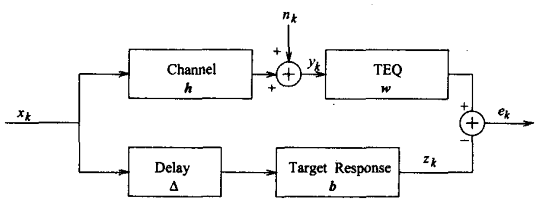


Fig. 1. Block diagram of the TEQ.

Let us assume that the TEQ is of length $l$ and that the decision delay is $\Delta$. From Fig. 1, we have

$$e_k = \mathbf{w}^T\mathbf{y}_k - \mathbf{b}^T\mathbf{x}_{\Delta,k} \tag{2}$$

where $\mathbf{x}_{\Delta,k} = [x_{k-\Delta}\ x_{k-1-\Delta}\ \cdots\ x_{k-l-1-\Delta}]^T$ and $\mathbf{y}_k = [y_k\ y_{k-1}\ \cdots\ y_{k-v-1}]^T$. It follows that

$$E[e_k^2] = \mathbf{w}^T\mathbf{R}_{yy}\mathbf{w} + \mathbf{b}^T\mathbf{R}_{xx}\mathbf{b} - 2\mathbf{b}^T\mathbf{R}_{xy,\Delta}\mathbf{w} \tag{3}$$

where $\mathbf{R}_{xx} = E[\mathbf{x}_{\Delta,k}\mathbf{x}_{\Delta,k}]^T$, $\mathbf{R}_{yy} = E[\mathbf{y}_k\mathbf{y}_k^T]$, and $\mathbf{R}_{xy,\Delta} = E[\mathbf{x}_{\Delta,k}\mathbf{y}_k^T]$. Since the number of subchannels is large relative to the number of taps in the TIR, we can assume that the amplitude response of the TIR is flat within each subchannel. Consequently, with the fact that the signal spectrum tends to be flat [7], the signal energy at the output of the TEQ can be expressed as

$$E[z_k^2] = \sigma_x^2 \sum_{i\in I} |\tilde{b}_i|^2 \tag{4}$$

where $\sigma_x^2$ is the signal spectrum density, $\tilde{b}_i$ is the $i$th N-point FFT of $\mathbf{b}$, $N$ is the symbol length, and $I$ is the index set for the occupied subchannels. Let $\mathbf{Q}$ be the N-point FFT matrix with

$$q_{m,n} = \frac{1}{\sqrt{N}} \exp\left[-j\frac{2\pi(m-1)(n-1)}{N}\right]$$

Eq. (4) can be rewritten as

$$E[z_k^2] = \sigma_x^2\mathbf{b}^T\mathbf{B}\mathbf{b} \tag{5}$$

where

$$\mathbf{B} = \mathbf{Q}_v^H\mathbf{I}_d\mathbf{Q}_v \tag{6}$$

$\mathbf{Q}_v$ is an $N \times v$ matrix consisting of the $(\Delta+1)$th to the $(\Delta+v)$th columns of $\mathbf{Q}$, and $\mathbf{I}_d$ is an $N \times N$ diagonal matrix with the $i$th diagonal entry equal to 1 if $i \in I$ or 0 otherwise. Hence, maximizing the SNR is equivalent to

$$\begin{aligned} &\text{minimize} && F_{\mathbf{c}} = \mathbf{c}^T\mathbf{A}\mathbf{c} && (7)\\ &\text{subject to} && \mathbf{c}^T\hat{\mathbf{B}}\mathbf{c} = 1 && (8)\end{aligned}$$

where $\mathbf{c} = [\mathbf{w}^T\ \mathbf{b}^T]^T$, $\hat{\mathbf{B}}$ is a $(v+l)\times(v+l)$ matrix with $\mathbf{B}$ as its lower right block and zeros elsewhere, and

$$\mathbf{A} = \begin{bmatrix} \mathbf{R}_{yy} & -\mathbf{R}_{xy,\Delta}^T \\ -\mathbf{R}_{xy,\Delta} & \mathbf{R}_{xx}\end{bmatrix} \tag{9}$$

Using the Lagrange multipliers [8], the optimization problem given in (7) and (8) can be transformed into

$$\text{minimize}\quad L(\mathbf{c},\alpha) = \mathbf{c}^T\mathbf{A}\mathbf{c} - \alpha(\mathbf{c}^T\hat{\mathbf{B}}\mathbf{c} - 1) \tag{10}$$

The gradient and the Hessian of the Langrange function $L(\mathbf{c},\alpha)$ with respect to $\mathbf{c}$ can then be found as

$$\nabla_{\mathbf{c}} L(\mathbf{c},\alpha) = 2(\mathbf{A} - \alpha\hat{\mathbf{B}})\mathbf{c} \tag{11}$$

$$\nabla_{\mathbf{c}}^2 L(\mathbf{c},\alpha) = 2(\mathbf{A} - \alpha\hat{\mathbf{B}}) \tag{12}$$

Solving $\nabla_{\mathbf{c}} L(\mathbf{c},\alpha) = 0$ and (8) for $\mathbf{c}$, we have

$$\alpha = \mathbf{c}^T\mathbf{A}\mathbf{c} \tag{13}$$

It follows that only those $\mathbf{c}$ that are the eigenvectors of the matrix pencil $\mathbf{A} - \alpha\hat{\mathbf{B}}$ are the stationary points of function $L(\mathbf{c},\alpha)$. Denote the eigenvector associated with the smallest eigenvalue $\alpha_{min}$ of the matrix pencil by $\mathbf{c}^*$. From (12), it can be shown that the Hessian is positive semidefinite only at $\mathbf{c}^*$. Furthermore, it can be verified that

$$\mathbf{d}^T\ \nabla_{\mathbf{c}}^2 L(\mathbf{c},\alpha)\,|_{\mathbf{c}=\mathbf{c}^*}\ \mathbf{d} = 0$$

if and only if $\mathbf{d}$ and $\mathbf{c}^*$ have the same direction. Hence, the optimization problem has a unique global minimizer.

## III. Adaptive Implementation

### A. A Steepest-Descent Algorithm

The uniqueness of the solution to the optimization problem given by (7) and (8) suggests that a steepest-descent method can be used to compute the coefficients of the TEQ. At the $k$th iteration, $\mathbf{c}_{k+1}$ can be computed as

$$\mathbf{c}_{k+1} = \mathbf{c}_k - \mu(\mathbf{A} - \alpha_k\hat{\mathbf{B}})\mathbf{c}_k \tag{14}$$

$$\mathbf{c}_{k+1} \leftarrow \frac{\mathbf{c}_{k+1}}{\mathbf{c}_{k+1}^T\hat{\mathbf{B}}\mathbf{c}_{k+1}} \tag{15}$$

where

$$\alpha_k = \mathbf{c}_k^T\mathbf{A}\mathbf{c}_k \tag{16}$$

A sufficient condition for the convergence of the above steepest-descent algorithm is given in Proposition 1 below.

*Proposition 1:* For an arbitrary starting point, the iterative procedure given by (14) and (15) converges to $\mathbf{c}^*$ if

$$\mu < \frac{2}{\lambda_{\max}(\mathbf{A} - \alpha_{\min}\hat{\mathbf{B}})} \tag{17}$$

where $\lambda_{\max}$ stands for the maximum eigenvalue and $\alpha_{min}$ is the minimum eigenvalue of the matrix pencil $\mathbf{A} - \alpha\hat{\mathbf{B}}$.

*Proof:* Note that $\alpha_k$ is the output noise-to-signal ratio corresponding to $\mathbf{c}_k$. Hence, in order to prove the proposition, we need only to show the convergence of $\alpha_k$. From (14), (15), and (16), we have

$$\alpha_{k+1} = \alpha_k + \frac{T_k}{\mathbf{b}_{k+1}^T\mathbf{B}\mathbf{b}_{k+1}} \tag{18}$$

where

$$T_k = \mu\mathbf{b}_k^T(\mathbf{A} - \alpha_k\mathbf{B})[\mu(\mathbf{A} - \alpha\mathbf{B}) - 2\mathbf{I}](\mathbf{A} - \alpha\mathbf{B})\mathbf{b}_k \tag{19}$$

It follows that if (17) is satisfied, matrix $\mu(\mathbf{A} - \alpha\mathbf{B}) - 2\mathbf{I}$ will be negative definite. Consequently, we have

$$T_k \leq 0 \tag{20}$$

Since the denominator of the second term at the right-hand side of (18) is always positive, we have

$$\alpha_{k+1} - \alpha_k \leq 0 \tag{21}$$

where the equality holds only if $(\mathbf{A} - \alpha_k\mathbf{B})\mathbf{b}_k = 0$, i.e., $\mathbf{b}$ is a stationary point of the optimization problem. However, at stationary points other than the minimizer, the algorithm will not stop due to the rounding errors. This proves the proposition. ■

In practice, however, the gradient is unknown. In the rest of the section, two stochastic gradient methods for the adaptation of the TEQ will be developed in the time and frequency domains, respectively.

### B. A Time-Domain Adaptation Algorithm

In the steepest-descent method given by (14) and (15), the unknown matrix $\mathbf{A}$ can be approximated instantaneously as

$$\mathbf{A} \approx [\mathbf{y}_k^T\ \mathbf{x}_{\Delta,k}^T]^T[\mathbf{y}_k^T\ \mathbf{x}_{\Delta,k}^T] \tag{22}$$

In doing so, we obtain a stochastic gradient method in the time domain. From (3), (14), (15), and (22), at the $k$th iteration, we compute

$$\begin{aligned} \mathbf{b}_{k+1} &= \mathbf{b}_k + \mu e_k\mathbf{x}_{\Delta,k} + \mu e_k^2\mathbf{B}\mathbf{b}_k && (23)\\ \mathbf{w}_{k+1} &= \mathbf{w}_k - \mu e_k\mathbf{y} && (24)\\ t &= \mathbf{b}_{k+1}^T\mathbf{B}\mathbf{b}_{k+1} && (25)\\ \mathbf{b}_{k+1} &\leftarrow \frac{\mathbf{b}_{k+1}}{t} && (26)\\ \mathbf{w}_{k+1} &\leftarrow \frac{\mathbf{w}_{k+1}}{t} && (27) \end{aligned}$$

Since $\alpha_k$ is updated through a nolinear recursion, the necessary and sufficient condition for the convergence of the algorithm with respect to the step size $\mu$ is difficult to obtain. However, the convergence of the algorithm in the neighbourhood of $\mathbf{c}^*$ can be obtained by letting $\alpha_k = \alpha_{\min}$. In addition, the condition stated in (17) was found useful as a guideline for the selection of $\mu$.

### C. A Frequency-Domain Adaptation Algorithm

Since an ADSL modem has on-board FFT and IFFT computing resources, an efficient frequency-domain adaptation algorithm can also be derived as an alternative to the time-domain adaptation algorithm. Furthermore, a frequency-domain adaptation enjoys a faster convergence and yields a smaller excess MSE due to its block processing mechanism.

Let $\tilde{\mathbf{e}}_k$, $\tilde{\mathbf{b}}$, and $\tilde{\mathbf{w}}$ denote the $N$-point FFT of $[e_k\ e_{k-1}\ \cdots\ e_{k-N+1}]^T$, $[\mathbf{0}_{1,\Delta}\ \mathbf{b}^T\ \mathbf{0}_{1,N-v-\Delta}]^T$, and $[\mathbf{w}^T\ \mathbf{0}_{1,N-L}]^T$, respectively, where $\mathbf{0}_{m,n}$ stands for a zero matrix of size $m \times n$. In the frequency domian, the optimization problem can be formulated as

$$\begin{aligned} &\text{minimize} && E[\tilde{\mathbf{e}}_k^H\tilde{\mathbf{e}}_k] && (28)\\ &\text{subject to} && \tilde{\mathbf{b}}^H\mathbf{I}_d\tilde{\mathbf{b}} = 1 && (29) \end{aligned}$$

where $\mathbf{I}_d$ is an $N \times N$ diagonal matrix with its $i$th diagonal entry equal to 1 if the $i$th subchanel is occupied or 0 otherwise. Since during the training period, a pseudo-random sequence of length $N$ is transmitted periodically, we have

$$\tilde{\mathbf{e}}_k = \tilde{\mathbf{Y}}_k\tilde{\mathbf{w}} - \tilde{\mathbf{X}}_k\tilde{\mathbf{b}} \tag{30}$$

where $\tilde{\mathbf{Y}}_k = \text{diag}[\tilde{\mathbf{y}}_k]$, $\tilde{\mathbf{X}}_k = \text{diag}[\tilde{\mathbf{x}}_k]$, $\tilde{\mathbf{y}}_k = [y_k\ y_{k-1}\ \cdots\ y_{k-N+1}]^T$, and $\tilde{\mathbf{x}}_k = [x_k\ x_{k-1}\ \cdots\ x_{k-N+1}]^T$. Let

$$\tilde{\mathbf{A}} = E\left\{\begin{bmatrix} \mathbf{Y}_k^H\mathbf{Y}_k & -\mathbf{Y}_k^H\mathbf{X}_k \\ -\mathbf{X}_k^H\mathbf{Y}_k & \mathbf{X}_k^H\mathbf{X}_k \end{bmatrix}\right\} \tag{31}$$

Denoting $\tilde{\mathbf{c}}$ as $[\tilde{\mathbf{w}}^H\ \tilde{\mathbf{b}}^H]^H$, the optimization problem given in (28) and (29) can then be rewritten as

$$\begin{aligned} &\text{minimize} && \tilde{\mathbf{c}}^H\tilde{\mathbf{A}}\tilde{\mathbf{c}} && (32)\\ &\text{subject to} && \mathbf{b}^H\mathbf{P}_v^H\mathbf{I}_d\mathbf{P}_v\mathbf{b} = 1 && (33) \end{aligned}$$

where $\mathbf{P}_v = \mathbf{Q}_v\mathbf{Q}_v^H$. Following similiar steps as those in the time-domain development, we obtain a stochastic algorithm in the frequency domain, in which the $k$th iteration consists of computing

$$\begin{aligned} \tilde{\mathbf{b}}_{k+1} &= \tilde{\mathbf{b}}_k + \mu\mathbf{P}_v(\mathbf{X}^H\tilde{\mathbf{e}} + \tilde{\mathbf{e}}^H\tilde{\mathbf{e}}\mathbf{I}_d\tilde{\mathbf{b}}_k) && (34)\\ \tilde{\mathbf{w}}_{k+1} &= \tilde{\mathbf{w}} - \mu\mathbf{P}_l\mathbf{Y}^H\tilde{\mathbf{e}} && (35)\\ t &= \tilde{\mathbf{b}}_k^H\mathbf{I}_d\tilde{\mathbf{b}}_k && (36)\\ \tilde{\mathbf{b}}_{k+1} &\leftarrow \frac{\tilde{\mathbf{b}}_{k+1}}{t} && (37)\\ \tilde{\mathbf{w}}_{k+1} &\leftarrow \frac{\tilde{\mathbf{w}}_{k+1}}{t} && (38) \end{aligned}$$

where $\mathbf{P}_l = \mathbf{Q}_l \mathbf{Q}_l^H$ and $\mathbf{Q}_l$ is an $N \times l$ matrix consisting of the first $l$th columns of $\mathbf{Q}$. Note that the two projection operations $\mathbf{P}_v$ and $\mathbf{P}_l$ can be efficiently computed by using the FFT and IFFT.

## IV. Simulation Results

Extensive simulations have been carried out to examine the performance of the proposed adaptation algorithm. As an example, we considered ANSI recommended carrier-serving area loop 1 [1]. The sampling frequency was 1024 kHz and the length of the IFFT was 256 samples. The SNR at the input of the receiver was assummed to be 30 dB. The ideal SNR at the output of the TEQ was found to be 24.47 dB. In Figs. 2 and 3, ensemble-averaged learning curves of the proposed time-domain and frequencey-domain adaptation algorithms are illustrated, respectively, in terms of the SNR . Same step size was used for the two algorithms. It can be observed from the figures that an SNR of about 21 dB is achieved after about 3500 and 2000 iterations by the time-domain algorithm and the frequency-domain algorithm, respectively. For the time-domain algorithm, the SNR converges to about 22.5 dB after 5000 iterations. On the other hand, the in the frequency-domain algorithm, the SNR converges to about 23 dB after about 4000 iterations. Considering that 5000 iterations correspond to only 20 symbols, the convergence speeds of the two algorithm can be regarded as satisfactory in practice. As expected, the results also show that the frequency-domain algorithm has a faster convergence and yields a smaller excess MSE.

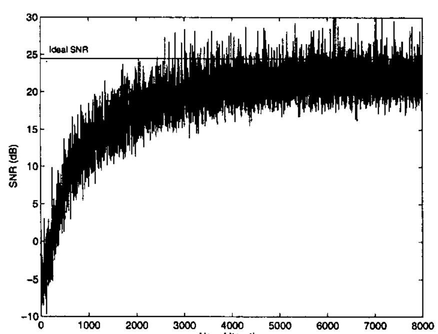


Fig. 2. Learning curve of the proposed time-domain adaptation algorithm.

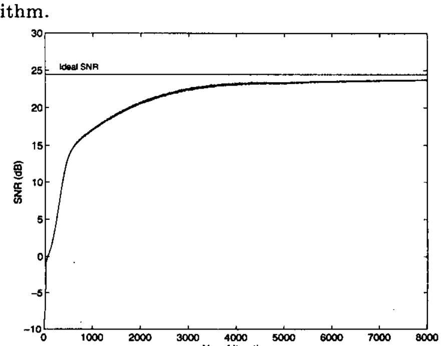


Fig. 3. Learning curve of the proposed frequency-domain adaptation algorithm.

## V. Conclusions and Future Work

A new criterion for the design of a TEQ for ADSL systems has been studied. It has been shown that the new criterion leads to a maximum SNR regardless of the profile of bandwidth occupancy. Moreover, it has been shown that the solution to the optimization problem based on the new criterion is unique. This fact allowed us to develop two stochastic gradient algorithms for the adaptive implementation of the new TEQ. Simulation results have demonstrated that two algorithms converge at a satisfactory speed but the frequency-domain algorithm offers a faster convergence and yields a smaller excess MSE.

A sufficient condition for the convergence of the algorithms with respect to the step size has been derived for the case where the gradient estimation is not affected by noise. The convergence of the algorithms for more realsitic situations where the gradient estimation is affected by noise is currently under investigation.

### Acknowledgement

The authors would like to acknowledge Micronet, Networks of Centers of Excellence Program, and PMC-Sierra Inc. for supporting this research.

## References

[1] *ANSI T1.413-1995, Asymmetric Digital Subscriber Line (ADSL) Metallic Interface.*

[2] J. S. Chow, J. Tu, and J. M. Cioffi, "A discrete multitone transceiver system for HDSL applications," *IEEE J. Selct. Areas Commun.*, vol. 9, no. 6, pp. 859-908, Aug. 1991.

[3] D. Falconer and F. Magee, "Adaptive channel memory truncation for maximum likelihood sequence estimation," *Bell System Tech. J.*, vol. 52, no. 9, pp. 1541-1562, Nov. 1973.

[4] J. V. Kerckhove and P. Spruyt, "Adaptive optimization criterion for FDM-based DMT-ADSL equalization," *IEEE Int. Conf. Commun.*, vol. 3, pp. 1328-1334, 1996.

[5] J. S. Chow, J. M. Cioffi, and J. A. C. Bingham "Equalizer training algorithms for multicarrier modulation systems," *IEEE Int. Conf. Commun.*, vol. 1, pp. 761-765, 1993.

[6] I. Lee, J. S. Chow, and J. M. Cioffi, "Performance evaluation of a fast computation algorithm for the DMT in high-speed subscriber loop," *IEEE J. Selct. Areas Commun.*, vol. 13, no. 9, pp. 1564-1570, Dec. 1995.

[7] P. Chow, J. M. Cioffi, and J. Bingham, "A practical discrete multitone transceiver loading algorithm for data transmission over spectrally shaped channels," *IEEE Trans. Commun.*, vol. 43, pp. 773-775, Feb./Mar./Apr. 1995.

[8] R. Fletcher, *Practical Methods of Optimization*, 2nd ed., Wiley, New York, 1987.