Title Sheet

COVER INFORMATION FOR SAE TECHNICAL PAPERS

In order to ensure that the correct title, author names, and company affiliations appear on the cover and title page of your paper, please provide the information requested below.

PAPER TITLE
(upper and lower case):

Use of an Implicit Filtering Algorithm for Mechanical System Parameter Identification

AUTHORS (upper and lower case)	COMPANY (upper and lower case)
1. J. W. David	1. North Carolina State University
2. C.T. Kelley	2. North Carolina State University
3. C.Y. Cheng	3. North Carolina State University

ABSTRACT

Optimal design of high-speed valve trains requires the use of an accurate analytical model. While the governing differential equations are important, the coefficients (or parameters) used in these equations are equally as important. Since many of the parameters used in valve train models are difficult to measure directly, parameter identification based on experimental data is required to assure model accuracy.

This paper addresses the parameter identification problem for a valve train model, formulating a scalar cost function which represents the difference in measured and predicted system response. Minimizaton of this cost function yields the 10 unknown system parameters. As the cost function has many local minima, a global optimization scheme must be employed. An implicit filtering algorithm is implemented which applies a scale reduction scheme in conjunction with a gradient projection algorithm to avoid becomming trapped in local minima and thus produces near global minima of the cost function. The implicit filtering algorithm has several tuning parameters which allows its adaptation to many problems. For this problem, implicit filtering proved to be 2 to 4 times more efficient than the adaptive random search method previously employed.

INTRODUCTION

Figure 1 shows a schematic of a typical push-rod type valve train which is the focus of this study. In the design and optimization of such mechanisms for high-speed internal combustion engines, an accurate analytical model is essential. Whether the model is simple [1,2] or complex [3,4], accurate parameters are essential for the accuracy of the model. For linear systems, parameter estimation is a rather mature art while for nonlinear systems the problem of determining system parameters is much more difficult. In that a valve train is a transient mechanism with some components exhibiting steady-state response [5], the problem of parameter estimation is increasingly difficult.

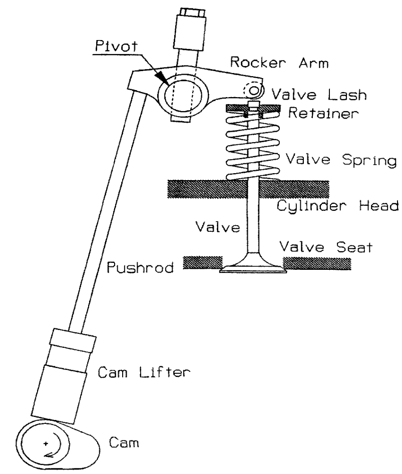

Figure 1: A Typical Pushrod Valve Train [4]

THE MODEL

Figure 2 shows a schematic of a model developed by Kim and David [4] for use in high-speed valve train analysis. It has the advantage of being computationally efficient while having good accuracy for engine speeds of 9,000 rpm and above. This model has several parameters which physically describe the mechanism. The mass parameters can be directly measured and are thus not a problem in this case. The valve spring reactive force, as well parameters k_5 and c_5, are also easily determined. However, the stiffness parameters k_1 through k_4 are not easily measured and the damping parameters c_1 through c_4, along with the two friction coefficients (rocker arm pivot and valve stem) are extremely difficult to measure. The use of a system identification procedure allows these unknows to be determined in a manner which obtains the best correlation of the model with experimental data.

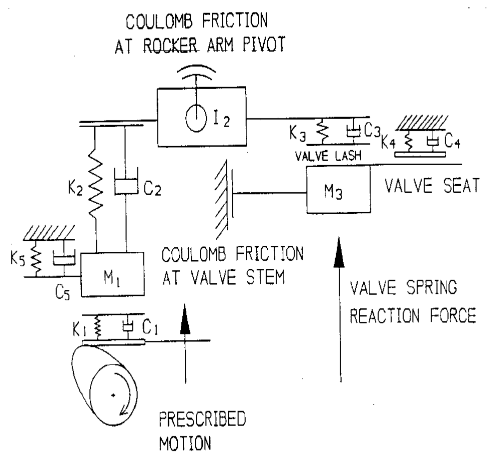

Figure 2: Schematic of the Valve Train Model [4]

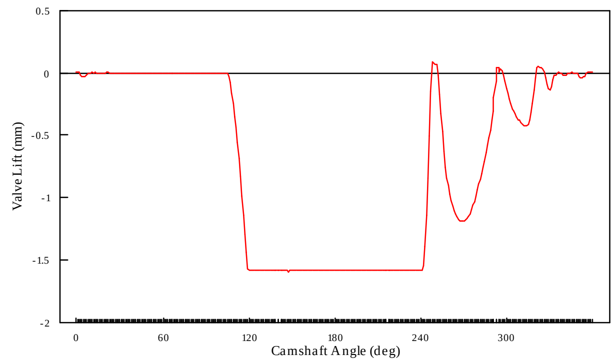

Figure 3: Short Range Proximeter Measurement of Valve Motion [5]

EXPERIMENTAL APPARATUS AND TEST RESULTS

The test apparatus consists of an engine block less the connecting rods and pistons and with a straight shaft replacing the crankshaft. It has a complete valve train including the cam drive mechanism. A 30 hp direct current motor is used to drive the assembly. Various sensors are used to measure the dynamic response of the mechanism, such as strain-gages, accelerometers, and proximeters. The camshaft rotation angle is measured with an optical encoder which triggers a computer to acquire data.

Figure 3 shows a typical valve motion trace obtained from a short-range proximeter. The valve is initially closed and as it begins to open, the proximeter measures the displacement. Next, the valve moves outside the range of the sensor and thus, the displacement curve is flat. Then, the proximeter senses the valve on the closing event and measures the subsequent displacement. Notice that after closing, the valve bounces off the seat, often more than once. The amplitude of this first bounce is an excellent indicator of the dynamic performance of the mechanism [6] and is used for future comparisons. The range of the proximeter used in this test is about 1.5 mm. Other sensors, such as long range proximeters, are capable of measuring the entire displacement curve of the valve. However, results indicate that the precision of the short range data shown in Figure 3 produces better correlation of the analytical model due to the increased resolution in the valve seating area and subsequent definition of the valve bounce phenomena. Data is taken at a family of engine speeds, usually 7,000 rpm to limiting speed (between 9,000 rpm and 12,000 rpm for the type of valve train presented here) in increments of 100 rpm. Thus, an experiment produces a family of 20 or more displacement curves like Figure 3 that are used in the parameter identification.

COST FUNCTION

Because the system is transient and highly nonlinear, the parameter identification is cast as a least squares error between the predicted response of the model and the experimental data. This error is summed over the number of data points and number of test engine speeds:

$$f = \sum_{i=1}^{nspeed} \sum_{j=1}^{npoints} w_j (y_i(j) - y_i^*(j))^2 \tag{1}$$

where y_j is any predicted system response and y_j^* is the corresponding measured response. For this study, the response is the valve motion. The weighting function w_j is included to exclude the data for which the valve is out of range of the sensor. It can also be used to weight other data points to increase/decrease correlation with a particular portion of the data. The experimental data can be any measurable quantity of the valve train, or even multiple sensor readings (such as proximeter and strain-gage measurements). As the predicted response of the model is a function of the unknown parameters, the goal is to minimize the error between the measured and predicted responses. In an ideal situation with no modeling misconceptions or numerical errors, the cost function would be zero.

THE OPTIMIZATION PROBLEM

Figures 4 and 5 show three dimensional sections of the optimization landscape. Figure 4 is a plot of the total cost function against two of the unknown parameters and figure 5 is a similar plot with the cost function computed using only a single engine speed.

The figures illustrate two significant features of the problem. First, there are many local minima that correspond to objective functions far from optimal. These local minima will entrap traditional optimization algorithms and lead to premature termination of the iteration. Secondly, even when the objective function value is small, the local minima limit the accuracy to which optimal parameter values can be determined. In both figures we see a simple objective with local minima caused by low-amplitude, high-frequency noise.

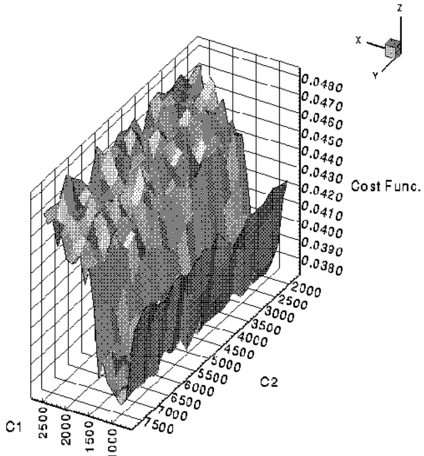

Figure 4: Cost Function Variation with Two Parameters

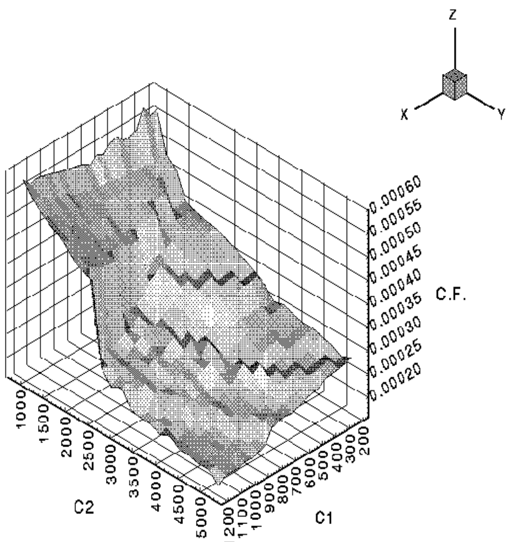

Figure 5: Cost Function Variation for Two Parameters, Single RPM

Non-standard optimizers must be used to drive the cost function to a small value. However, one does not need and should not expect to find the global minimum. A sensible goal of an optimization of a function of this kind is simply to find a value of the parameters that make the objective function small. There may be several such sets of values of the parameters and the valve train models based on those sets of parameters will have similar characteristics.

In this paper, implicit filtering [7] performs a sequence of optimizations with varying resolution, thereby seeking to capture the low-frequency behavior of the objective and find a low function value without becoming trapped in a local minimum before reaching an acceptable parameter set.

THE OPTIMIZER: IMPLICIT FILTERING

Implicit filtering, originally proposed in the context of computer aided design of semiconductors [8,9,10] is a generalization of the gradient projection algorithm of [11] in which derivatives are computed with difference quotients. The step sizes (called scales) in the difference quotients are changed as the iteration progresses with the goal of avoiding local minima that are caused by high-frequency, low amplitude oscillations such as those seen in Figs. 4 and 5. Real filtering could be performed, but this requires sampling and filtering the entire solution space and thus, is computationally quite expensive. Implicit filtering is very similar to adaptive meshing schemes used by the computational fluid mechanics community to avoid unwanted harmonics. The algorithm is fully described in [7] and [12]. A brief summary is presented here.

Gradient projection algorithms are variations of the steepest descent method. The goal is to solve the constrained optimization problem

$$\min_{\mathbf{x}\in\boldsymbol{\Omega}} f(\mathbf{x}) \tag{2}$$

where $\boldsymbol{x}$ is the vector of unknown parameters to be determined. The feasible set of solutions Ω is defined by simple bounds

$$\boldsymbol{\Omega} = \left\{\mathbf{x} \middle| l_i \le x_i \le u_i \right\}. \tag{3}$$

Here x_i is the ith component of the vector $\boldsymbol{x}$. Central to the gradient projection algorithm is the ease with which a vector $\boldsymbol{x}$ can be projected onto $\boldsymbol{\Omega}$. The projection of $\mathbf{P}$ onto $\boldsymbol{\Omega}$ is defined by

$$\mathrm{P}_i(x) = \begin{cases} l_i & \text{if } x_i \le l_i \\ u_i & \text{if } x_i \ge u_i \\ x_i & \text{if } l_i < x_i < u_i \end{cases} \tag{4}$$

where l_i is the ith component of the vector representing lower bounds on the solution and u_i is the ith component of the vector representing upper bounds on the solution. The gradient projection iteration from a current approximation $\mathbf{x}_c$ to a better approximation $\mathbf{x}_+$ is defined by

$$\mathbf{x}_+ = \mathbf{P}(\mathbf{x}_c - a\nabla f(\mathbf{x}_c)) \tag{5}$$

where α is a step length parameter determined so that $f(\mathbf{x}_+)$ is smaller than $f(\mathbf{x}_c)$.

If the gradients are approximated by finite differences, ∇f is replaced by some difference approximation $\nabla_h f$ where $\nabla_h f$ could be forward, backward, or centered differences. In the case of centered differences, which are used in this study, the ith component of $\nabla_h f(\mathbf{x})$ is

$$(\nabla_{\text{h}} f(\mathbf{x}))_i = \frac{f(\mathbf{x}+\text{h}\mathbf{e}_i) - f(\mathbf{x}-\text{h}\mathbf{e}_i)}{2\text{h}} \qquad (6)$$

where $\mathbf{e}_i$ is the ith canonical basis vector. For such an approximation, accuracy of the order of h^2 is expected. The convergence theory of Gilmore and Kelley [13] applies to both centered and one sided differences. Experience in using this algorithm in semiconductor applications indicates that centered differences work far better and thus are used in this work.

The iterative scheme involves calculating the gradient using Equation 6. This is then used in Equation 3 and 4 to determine a *candidate solution* $\mathbf{x}_+$. This *candidate solution* $\mathbf{x}_+$ is then checked for sufficient descent by

$$f(\mathbf{x}_c) - f(\mathbf{x}_+) \geq \frac{\sigma \|\mathbf{x}_c - \mathbf{x}_+\|^2}{\alpha} \qquad (7)$$

If Equation 7 is satisfied, the *candidate solution* $\mathbf{x}_+$ is accepted and the iteration continues by repeating the above steps up to a limiting value *kmax* times. If Equation 7 is not satisfied, α is reduced as

$$\alpha = \beta\alpha \qquad (8)$$

,with $0<\beta<1$, a *new candidate solution* $\mathbf{x}_+$ is calculated from eqs. 3 and 4 and checked using Equation 7. This is repeated up to *jmax* times until an acceptable solution $\mathbf{x}_+$ is found. If an acceptable solution is not found in this process, is usually means that the step size h (or scale) is too large and needs to be reduced. This is done in the implicit filtering algorithm.

The iteration is terminated whenever

$$\|x_+ - x_{\mathbf{c}}\| \leq \tau\, h \qquad (9)$$

is below a desired tolerance.

The adaptation of the gradient projection algorithm to implicit filtering is rather simple. Successive applications of the gradient search algorithm are performed using ever decreasing scales. This is done a total of *m* times. This has the effect of filtering out higher harmonics of the cost function in the early stages of the iteration while admitting them in the final stages of the iteration to improve final convergence. Once the convergence criteria (7) is satisfied for all scales, the algorithm is restarted using the best-to-date solution and rechecked to assure convergence for all scales.

COMPUTATIONAL RESULTS

The implementation of implicit filtering is contained in a code named IFFCO. IFFCO can be obtained by anonymous ftp to math.ncsu.edu in the directory pub\kelley\iffco. Users are advised to get both the FORTRAN code iffco.f and the users' guide ug.ps. The inputs to the code are:

1. the objective function f
2. the initial guess of the solution
3. the bounds $\{\mathbf{u}_i\}$ and $\{\boldsymbol{l}_i\}$
4. an estimate f_{scale} for the minimum absolute value of f
5. τ, the constant used in the termination criteria
6. maximum number of step size reductions *m*
7. the maximum number of step length reductions *jmax*
8. the maximum number of iterates for each use of the gradient projection algorithm *kmax*

Of these, f_{scale}, τ, *m*, *jmax* and *kmax* are parameters that the user can tune to improve the performance of IFFCO.

The bounds on the variables are scaled, using $\{\mathbf{u}_i\}$ and $\{\boldsymbol{l}_i\}$, to be between 0 and 1. The function f is divided by f_{scale} in order to make the estimated maximum value of f approximately 1. These two scalings are intended to assist the gradient projection algorithm avoid terminations on the initial iteration or after too many step length reductions. The algorithm appears to be relatively insensitive to f_{scale} and using the function value from the initial guess is reasonable.

τ reflects the estimate of the amplitude of the high frequency noise in the problem. If the amplitude of τ is small and the noise amplitude is large, the convergence criteria (eq. 7) might never be satisfied. Small amplitude noise would require a small value of τ. If τ is too large, the convergence criteria will be satisfied too easily causing the scale to be reduced too soon, admitting high frequency components into the cost function evaluation. This will cause IFFCO to become trapped in a local minimum. Thus, this is the most important parameter and must be determined by experimentation with each problem undertaken. For this problem, $\tau=0.1$ worked well.

Since the problem is now bounded by 0 and 1, the scale (step size) h per application of the gradient projection algorithm is defined as $h=2^{-i}$. Other variations are possible, but this one worked well for this problem. The number of step reductions *m* should be large enough to provide good resolution of the solution. For this problem *m* = 10 was used.

The maximum number of iterates *kmax* and *jmax* are tuned to allow the iterative schemes to proceed long enough for successful termination without being so large as to use excessive computational effort. For this problem 5 was used for both.

Figure 6 shows the results using IFFCO. The parameter plotted is the magnitude of the first valve bounce experience by the mechanism after closing as a function of engine speed. The initial guess was unusable in that it completely failed to predict the instability in the system at 9,200 rpm. IFFCO was able to obtain a very good solution, predicting the instability region with good precision. However, the adaptive random search technique [14] is also capable of finding good solutions, given enough time. Thus, one must look at computational efficiency to assess the superior technique.

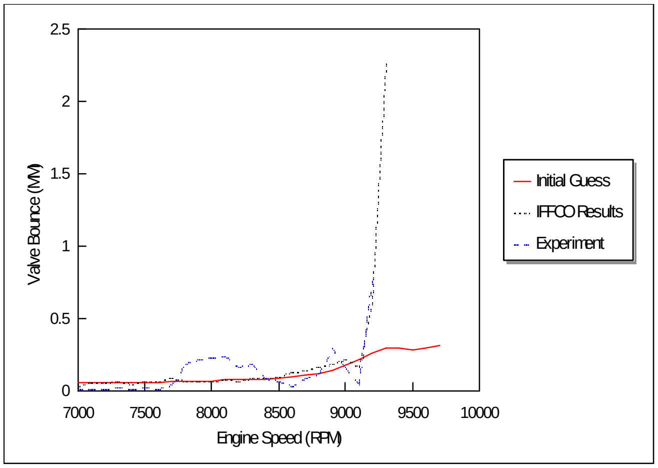

Figure 6: Experimental Results, Initial Guess and IFFCO Results

Figure 7 shows the results of a comparison between IFFCO and the adaptive random search (ARS) technique. Upper bounds and lower bounds are the same for both techniques and z represents the distance between the upper and lower bounds used for the initial guess; i.e.

$$x_i = l_i + (u_i - l_i)z \tag{10}$$

For both methods, iteration was stopped when the cost function reached the same goal. In all cases, IFFCO obtained solutions in 1/3 to 1/4 of the number of iterations as ARS.

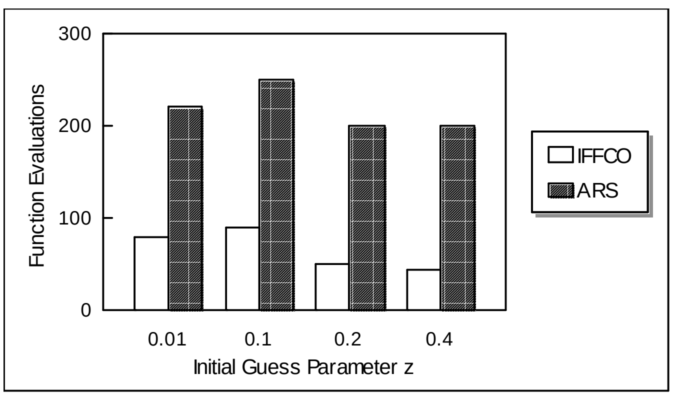

Figure 7: Cost Function Evaluations as a Function of Initial Guess

Figure 8 shows a comparison of cost function evaluations obtained after 60 iterations. In all cases, the solution obtained by IFFCO is half the value obtained by ARS.

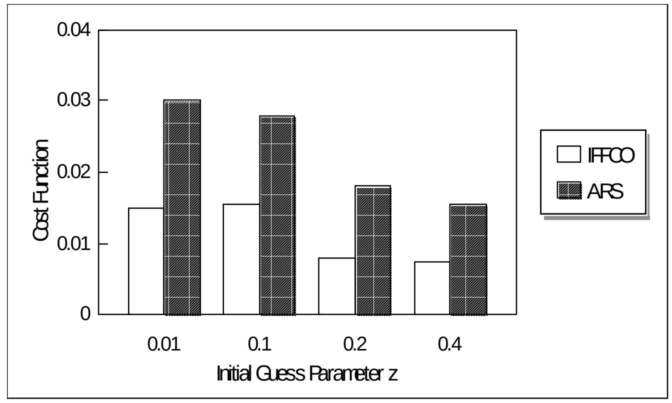

Figure 8: Cost Function Value as a Function of Initial Guess for 60 Iterations

CONCLUSIONS

Implicit filtering appears to be an efficient method for parameter identification of high-speed, nonlinear valve trains. Reductions of computation effort as high as 75% were realized over the adaptive random search technique. Some numerical experiments are required, however, to adjust the tuning parameters to achieve optimal results.

ACKNOWLEDGMENTS

The authors wish to gratefully acknowledge Mr. Jim Covey of the GM Motorsports Technology Group for his support of the valve train research and the National Science Foundation for their support of the optimization research under grant DMS-9321938.

REFERENCES

[1] P. Barkan, *Calculation fof High-Speed Valve Motion with a Flexible Overhead Linkage*, **SAE Transactions,** Vol. 61, 1953, pp. 687-700

[2] G.I. Johnson, *Studying Valve Dynamics with Electronic Computers*, **SAE Progress in Technology,** Vol. 5, 1963, pp. 10-28.

[3] A.P. Pisano and F. Freudenstein, *An Experimental and Analytical Investigation of the Dynamic Response of a High-Speed Cam Follower System. Part 2: A Combined, Lumped/Distributed Parameter Dynamic Model*, **ASME Journal of Mechanisms, Transmissions, and Automation in Design,** Vol. 105, Dec. 1983, pp. 699-704.

[4] D. Kim and J.W. David, *A Combined Model for High-Speed Valve Train Dynamics, Partly Linear and Partly Nonlinear*, **SAE Technical Paper Series 901726.**

[5] D. Kim, *Dynamics and Optimal Design of High Speed Valve Train Systems*, Ph. D thesis, Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC, March, 1990.

[6] J.W. David and D. Kim, *Optimal Design of High Speed Valve Train Systems*, **SAE Paper 942502,** 1994 Motor Sports Conference Proceedings, Vol. 2, Engines and Drivelines, pp. 103-109.

[7] P. Gilmore and C.T. Kelley, *An Implicit Filtering Algorithm for Optimization of Functions with Many Local Minima,* **SIAM Journal of Optimization,** Vol. 5, No. 2, May 1995.

[8] D. Stoneking, G. Bilbro, R. Trew, P. Gilmore and C.T. Kelley, *Yield Optimization Using a GaAs Process Simulator Coupled to a Physical Device Model,* **Proceedings IEEE/Cornell Conference on Advanced Concepts in High Speed Devices and Circuits,** IEEE, 1991, pp. 374-383

[9] D. Stoneking, G. Bilbro, R. Trew, P. Gilmore and C.T. Kelley, *Yield Optimization Using a GaAs Process Simulator Coupled to a Physical Device Model,* **IEEE Transactions on Microwave Theory and Techniques,** 40 (1992), pp. 1353-1363.

[10] D. Stoneking, G. Bilbro, R. Trew, P. Gilmore and C.T. Kelley, *Simulated Performance Optimization of GaAs MESFET Amplifiers,* **Proceedings IEEE/Cornell Conference on Advanced Concepts in High Speed Devices and Circuits,** IEEE, 1991, pp. 393-402.

[11] D.B. Bertsekas, *On the Goldstein-Levitin-Polyak Gradient Projection Method,* **IEEE Transactions on Automatic Control,** 21 (1976), pp. 174-184.

[12] P. Gilmore and C.T. Kelley, , *An Implicit Filtering Algorithm for Optimization of Functions with Many Local Minima,* Tech. Report CRSC-TR94-23, North Carolina State University, Center for Research in Scientific Computation, December, 1994.

[13] C.T. Kelley, *Iterative Methods for Linear and Nonlinear Equations,* SIAM Frontiers in Applied Mathematics, Philadelphia, 1995 (to appear).

[14] A. Liu, *System Identification and Optimal Design of High Speed Valve Train Systems,* Ph. D thesis, Mechanical and Aerospace Engineering, North Carolina State University, Raleigh, NC, April, 1994.