OPTIMAL SOLUTION OF NONLINEAR EQUATIONS

UUCS-85-103

K. Sikorski

Department of Computer Science
University of Utah
October, 1985

Abstract.

We survey recent worst case complexity results for the solution of nonlinear equations. Notes on worst and average case analysis of iterative algorithms and a bibliography of the subject are also included.

1. INTRODUCTION.

We study the approximate solution of the equation $f(x) = 0$ and also the problem of computing the topological degree of f. It is assumed that f belongs to the class of smooth functions F defined on a real interval or on the unit simplex or on the unit cube in m dimensions. The information on f consists in general of n values of arbitrary linear functionals which are computed sequentially (adaptively). The topological degree or an approximation to α, a zero of f, is constructed by an algorithm which uses these evaluations.

Several classes of functions F and the two error criteria are studied to determine whether or not it is possible to find an approximation to α which satisfies the error criterion to within a specified tolerance ϵ. The topological degree is computed exactly for the class of Lipschitz functions.

Complexity, i.e., the minimal cost of solving these problems, or lower and upper bounds on the complexity are found, and information and algorithms are developed which solve the problem with cost close to the complexity.

In Sections 2, 3 and 4 we include some history and list recent complexity results for the worst case setting. In Section 5 we exhibit future directions of research, and in Sections 6 and 7 include notes on worst and average case analysis of iterative algorithms.

2. Overview and Brief History.

In this section we discuss the worst case setting for the solution of nonlinear equations. We consider here algorithms which use general sequential information and whose error and cost are defined by the worst case performance. Sections 2–5 deal with non-iterative algorithms.

There are numerous papers in this area where optimal or nearly optimal algorithms are found. A partial list includes Booth (1967), Chernousko (1968), Eichhorn (1968), Gross and Johnson (1959), Hyafil (1977), Kiefer (1953, 1957), Micchelli and Miranker (1975), Sikorski (1982,1984b), Sikorski and Woźniakowski (1983) and Sukharev (1976), who considered the scalar case, and Boult and Sikorski (1984, 1985a, 1985b, 1985c), Nemirovsky and Yudin (1983), Majstrovskij (1972), Sikorski (1984a) and Todd (1978) who consider the multivariate case. We include here papers dealing with computing extrema since this problem is closely related to zero finding.

We briefly recall some of the results. The first results may be found in the Master's thesis of J. Kiefer, Kiefer (1953). He considers the search for the maximum in the class of scalar unimodal functions. The information is the values of f at n points. He proves that Fibonacci search is the optimal information and the optimal algorithm.

In the 1960's people worked on classes of functions which were convex or unimodal and obtained optimal algorithms for approximating zeros of functions and also derivatives; see Gross and Johnson (1959), Booth (1967), Chernousko (1968) and Eichhorn (1968).

Gross and Johnson consider the solution of scalar nonlinear equations for the class of convex continuos functions changing sign at the endpoints of an interval. The information is sequential function evaluations. They study optimal evaluation points and optimal algorithms.

Booth studies location of zeros of derivatives, Chernousko the search for a zero in the class of scalar functions with bounded difference quotients and Eichhorn the search for the maximum or a zero in the class of unimodal or montone nonincreasing functions.

In the 1970's the focus shifted to classes of functions with bounded derivatives. Examples here include the work of Micchelli and Miranker (1975) on "envelope methods", Sukharev (1976) on Lipschitz functions and Majstrovskij (1972) on optimality of Newton's method.

An important contribution is the work of Nemirovsky and Yudin (1983). They study the multivariate minimization problem and are interested in finding the minimal number $N(\epsilon)$ of function or derivative evaluations in order to determine ϵ-approximation to the extremum in the residual sense. The authors find sharp estimates for $N(\epsilon)$ for classes of convex and strongly convex functions as well as for nonconvex smooth functions defined n convex and/or compact set G of dimension n. For example $N(\epsilon) \simeq n \ln \frac{1}{\epsilon}$ for the convex class and $N(\epsilon) \simeq \left(\frac{1}{\epsilon}\right)^{n/k}$ for the nonconvex class of k-times continuously differentiable functions. The authors only occasionally deal with the problem of combining these $N(\epsilon)$ evaluations to compute an approximate solution.

There is much current research on optimal algorithms for solving nonlinear equations. A precise formulation of the worst case setting is given in Section 3 and selected recent results are stated in Section 4.

3. FORMULATION OF THE PROBLEM.

Let $\alpha(f)$ be a zero of the nonlinear function f and let $x(f)$ be an approximation to $\alpha(f)$. What we mean by approximation is determined by which error criterion we choose. Two commonly used criteria are the root criterion and the residual criterion.

Let $\epsilon > 0$ and a norm $\|\cdot\|$ be given. Then the *root criterion* is defined by

$$\|x(f) - \alpha(f)\| \leq \epsilon, \tag{3.1}$$

and the *residual criterion* is defined by

$$\|f(x(f))\| \leq \epsilon. \tag{3.2}$$

For each of these error criteria we study whether we can compute an approximation to within a tolerance ϵ, for every function f in a class F. This is called the worst case setting.

The information commonly used to solve this problem are evaluations of f and its derivatives. We will, more generally, assume the information consists of linear functionals. Namely, the information on f, $N(f)$, consists of n values of arbitrary linear functionals which are computed sequentially (adaptively):

$$N(f) = [L_1(f), ..., L_n(f)], \tag{3.3}$$

where the choice of L_i depends on $L_1(f), ..., L_{i-1}(f)$. If L_1 are given a priori the N is called *parallel* or *nonadaptive*, otherwise it is called *sequential* or *adaptive*. It is important to stress that this information is *partial* i.e., there are, in general, many functions sharing the same information. Knowing $N(f)$ we compute $x(f) = \varphi(N(f))$ by an *algorithm* φ which is a mapping

$$\varphi : N(F) \to D, \tag{3.4}$$

where D is the domain of the functions f. The error of an algorithm φ using N is given by

$$e(\varphi, N) = \sup_{f \epsilon F} \|\alpha(f) - \varphi(N(f))\| \tag{3.5}$$

in the case of the root criterion, and

$$e(\varphi, N) = \sup_{f \epsilon F} \|f(\varphi(N(f)))\| \tag{3.6}$$

in the case of residual criterion.

We next define *problem complexity.*

We assume that functional evaluations cost c and that arithmetic operations and comparisons cost unity.

To compute $\varphi(N(f))$ we must

$$\text{Compute } y = N(f), \tag{3.7}$$

and

(3.8) Compute $\varphi(y)$.

Thus the cost of an algorithm φ, $\text{cost}(\varphi, N)$, is given by

$$\text{cost}(\varphi, N) = \sup_{f \in F}(\text{cost}\, N(f) + \text{cost}\, \varphi(N(f))). \tag{3.9}$$

The *problem complexity* $\text{comp}(\epsilon)$ is defined as the minimal cost of an algorithm which uses arbitrary information N and computes an approximation $x(f)$ to within tolerance ϵ, i.e.,

$$\text{comp}(\epsilon) = \inf\{\text{cost}(\varphi, N) : \forall \varphi, N : e(\varphi, N) \leq \epsilon\} \tag{3.10}$$

By optimal information and an optimal algorithm we mean N and φ for which $\text{cost}(\varphi, N)$ $= \text{comp}(\epsilon)$ and $\epsilon(\varphi, N) \leq \epsilon$.

4. Recent Results.

In this section we list some recent complexity results. We conclude that for smooth functions the zero-finding problem with the root criterion is much more difficult than with the residual criterion. For example, for infinitely differentiable real functions with bounded arbitrary semi-norm it is impossible to solve the problem with the root criterion. It is, however, possible to solve it for the residual criterion. Thus additional restrictions on the functions are needed or the use of some nonlinear information is required to obtain positive results for the root criterion for these classes. We show that bisection is an almost optimal complexity algorithm for the class of functions changing sign at the end points of an interval.

We also present recent results on the complexity of computing the topological degree of a mapping, which can be used to determine the existence of a zero of f in a given domain.

Finally, we exhibit some results for the miultivariate case, both for the root and residual criteria.

4.1 ROOT CRITERION.

First, we consider the class F of smooth real functions having a real zero:

$$F_1 = \{f : [a,b] \to R,\ \exists \alpha : f(\alpha) = 0, f \in C^\infty[a,b], \|f\| \leq 1\}, \tag{4.1}$$

where $\|\cdot\|$ is an arbitrary semi-norm.

Of course, if $\epsilon \geq (b-a)/2$, then the midpoint $x(f) = \frac{b+a}{2}$ solves the problem for every f in F. Thus the only interesting case is for $\epsilon < (b-a)/2$. Then Sikorski and Woźniakowski(1983) show

Theorem 4.1. $\mathrm{comp}(\epsilon) = +\infty,\ \forall \epsilon < \frac{b-a}{2}$. ∎

This means that for ϵ less than half of the length of the interval there *exists no* information of the general form (3.3) and *no* algorithm to solve this problem.

A very different result holds when we assume that the functions change sign at the endpoints of an interval. Namely consider the class

$$F_2 = \{f : [a,b] \to R, f(a) \leq 0, f(b) \geq 0, f \in C^j\ [a,b]\}, \tag{4.2}$$

where j is an integer, $j = 0, 1, 2, ...,$ or $j = +\infty$. It is known that if the information consists of function evaluations, then the optimal algorithm and information are given by bisection. It is proved in Sikorski (1982) that bisection remains optimal even if the general information of (3.3) is permitted. We have

Theorem 4.2. $\mathrm{comp}(\epsilon) = (c + a_1)\left\lfloor log \frac{b-a}{\epsilon} \right\rfloor + a_2$, *where* $a_1 \in [0,3]$ *and* $a_2 \in [0,1]$. ∎

The upper bound is realized by bisection information and the bisection algorithm. Thus the bisection algorithm using bisection information enjoys almost minimal complexity. Note that $\mathrm{comp}(\epsilon)$ does not depend on function smoothness. The same result holds for the class of polynomials of unbounded degree assuming the information is continuous, see Sikorski(1984b).

Do similar results hold in the multidimensional case? The analagous assumption to sign change is an assumption on non-zero topological degree(deg). Namely let

$$
\begin{aligned}
&F_3 = \{f \epsilon C^{\infty}(\Delta), \Delta - \text{unit triangle in } R^2, \\
&f : \Delta \to R^2, f \text{ has exactly one zero, which is} \\
&\text{simple, } deg(f, \Delta, \theta) \neq 0, \theta = (0,0) \text{and } \Delta \text{ is} \\
&\text{completely labeled under } f\}.
\end{aligned} \tag{4.3}
$$

It turns out that even in the two dimensional case it is impossible to solve this problem. Thus one additional dimension makes the problem noncomputable. Boult and Sikorski(1984) proved

Theorem 4.3. *If* $\epsilon < diam(\Delta)/2$ *then*

$$\text{comp}(\epsilon) = +\infty.$$

∎

Thus there *exists no* information and *no* algorithm to solve this problem to within $\epsilon < diam(\Delta)/2$. In particular, the complexity of all methods which use as information function and/or derivative evaluations *is infinite*. Examples of such methods are continuation and simplicial continuation methods; see Allgower and Georg(1980).

This negative result suggests that additional restrictions on the class of functions or the use of nonlinear information are needed to obtain positive results in the multidimensional case.

4.2 Complexity of Computing Topological Degree.

A problem related to zero finding is the computation of topological degree. The crucial property of the degree , expressed by Kronecker's theorem, is the following: if the topological degree of f is not zero on some domain D then the function f has a zero

in D. Thus we can approximate a zero of f by the following algorithm. We start with some domain with nonzero degree. Subdivide this domain, compute the degree, if it is well defined, for the smaller domains, choose a subdomain of nonzero degree, etc. In this way we construct a sequence of domains with decreasing diameters and nonzero degrees. This idea was investigated by many people, see Eiger, Sikorski and Stenger(1984), Harvey and Stenger (1976), Kearfott(1977,1979), Prüfer and Siegberg(1980), Stenger(1975) and Stynes(1979a,1979b,1981). Such an algorithm would not always work, since the degree is not defined if a funtion has a zero on the boundary of a subdomain. We believe, however, that it would work with high probability. This claim, of course, requires introducing a probability measure into the space of functions ; it will be a topic of future research.

Here we exhibit lower and upper complexity bounds for the computation of degree in the class of Lipschitz functions. We stress that we compute the degree *exactly*. Let

$$(4.4)\quad F_4 = \{f:\ C \rightarrow\ R^n, \|f(x) - f(y)\|_\infty \leq\ K\|x - y\|_\infty, \|f(x)\| \geq\ d\ > 0,\ \forall x \epsilon\ \partial\ C\}$$

where $C\ =\ [0,1]^n$ is the unit cube in R^n. Since the degree is uniquely defined by function values on the boundary of a domain we assume as information the *sequential evaluation of function values* on the boundary $\partial\ C$ of C. For arbitrary $n, n\ \geq\ 2$, we have, Boult and Sikorski(1985b),

Theorem 4.4. *The complexity of computing degree, comp(deg), is bounded by:*

$$c_{low} \leq\ comp(deg) \leq\ c_{up},$$

$$c_{low}\ \simeq 2n(K/(8d))^{n-1}(c+n),$$

$$c_{up} \simeq 2n(K/(2d))^{n-1}(c\ +\ \frac{n^2}{2}\ (n-1)!),$$

whenever $K/(8d) \geq 1$.

Note that if $K/(2d) < 1$ then the functions in F_4 do not have zeros and therefore the degree is zero for every f. The case $1 \leq K/(2d) < 4$ is open. For the two dimensional case we have the stronger result, see Boult and Sikorski(1985a),

Theorem 4.5. *If* $n = 2$ *and* $K/(4d) \geq 1$ *then*

$$\text{comp}(deg) = 4\left\lfloor \frac{k}{4d} \right\rfloor (c + a) - 1$$

where a $\epsilon[2, 24]$. ∎

In both cases parallel function evaluations are almost optimal information. This information is used by an algorithm due to Kearfott(1979).

These results imply that the topological degree can be computed for small n and/or not too large $K/(2d)$. taking for example $n \leq 5$ and $K/(2d) \leq 10$ the degree can be computed with cost $\leq 10^5(c + 300)$. For large n and/or large $k/(8d)$ the problem is intractable. For example taking $n = 10$ and $K/(8d) = 10^3$ then the lower bound on complexity is $c_{low} \simeq 2.10^{28}(c + 10)$.

4.3 Residual criterion.

For the next two classes we show that the complexity of approximating a zero in the residual sense is finite. This contrasts with Theorem 4.1 in which we showed that with the root criterion the complexity is infinite. Namely consider the class F_5, which contains F_1 from (4.1) with the semi-norm $\|f\| = \|f^{(r)}\|_\infty$, and is defined by

(4.5) $F_5 = \{f : [a, b] \to R, f^{(r-1)}$absolutely continuous, $\|f^{(r)}\|_\infty \leq 1$ and $\exists \alpha : f(\alpha) = 0\}$.

It is proved in Sikorski and Woźniakowski (1983):

Theorem 4.5. $comp(\epsilon) = \theta(\epsilon^{-1/r})$. ∎

It turns out that optimal information is parallel and that the optimal algorithm is easy to implement for small r.

Next we report on the multivariate case. Namely, take Lipschitz class:

$$(4.6)\quad F_6 = \{f : C \to R^n : \|f(x) - f(y)\|_\infty \le K\|x - y\|_\infty, \text{and} \exists \alpha \in C : f(\alpha) = (0, ..., 0)\},$$

where $C = [0, 1]^n$.

It is proved in Sikorski (1984a):

THEOREM 4.6. $comp(\epsilon) = \theta((\frac{K}{2\epsilon})^n)$. ∎

We also showed that optimal information is parallel and that an optimal algorithm is an easy to implement search procedure. In fact this information and algorithm were already implemented in IMSL library by Aird and Rice (1977) in their ZSRCH routine.

5. FUTURE DIRECTIONS.

We list here a number of directions for future research.

(i) Optimal information and algorithms in probabilistic and average case models should be studied.

(ii) For a number of classes that we have investigated, the optimal information is parallel. We wish to obtain general conditions under which parallel information is optimal.

(iii) For some problems linear information is too weak. Therefore, restricted non-linear information should be investigated.

(iv) Classes of functions which are piecewise smooth should be considered since these classes arise in practice.

(v) Optimal information and algorithms should be implemented in software and hardware.

6. ITERATIVE MODEL OF COMPUTATION: WORST CASE.

In the 1960's and 1970's most of the research on optimal solution of nonlinear equations was devoted to the *iterative* model. In this model one constructs a sequence of points convergent to a zero of a function and wants to obtain information and an algorithm which guarantee the fastest possible convergence for every function in a given class. Thus, this can be viewed as an asymptotic worst case model with stationary iterative information;

stationary iterative information here means that a fixed set of linear functionals is repeatedly used in generating a sequence of points. Thus, for example, Newton information is evaluations of a function and its derivative on a sequence of points. Research monographs are due to Traub(1964), Ortega and Rheinboldt(1970), and Traub and Woźniakowski(1980, part B). Although our paper is not devoted to the iterative model we present a brief survey of this field.

The study of iterative complexity was intitiated by the work of Traub(1961,1964). In his 1964 monograph Traub considers iterations for approximating a simple or multiple zero of a scalar nonlinear function f. The information is the values of f and it derivatives. He introduces the classification of iterations according to the information used as one-point, one-point with memory, multipoint and multipoint with memory. He proves the maximal order theorem for one-point iterations and introduces the idea of interpolatory iteration. He conjectures that memory always adds less than one to order for a one-point iteration. He introduces multipoint iterations and shows that it differs significantly from one-point iteration. He considers several complexity measures.

Woźniakowski(1972-1976) generalizes the problem of maximal order to the multivariate and infinite dimensional cases, establishes the maximal order of interpolatory algorithms for the scalar case and shows that memory does not in general increase order for the multivariate case. He introduces the concept of order of information which provides a general tool for establishing the maximal order of an iteration. He shows that maximal order depends only on information used by an algorithm and not on the structure of the algorithm, see also Traub and Woźniakowski(1976a,1976b,1979,1980a,1980b,1980c). Significant papers on optimal iteration include Brent (1973, 1976a, 1976b), Brent, Winograd and Wolfe (1973), Kacewicz (1976a, 1976b, 1979), Kung (1976) Kung and Traub (1974a,1974b,1976), Meersman (1976a, 1976b), Saari and Simon (1978), Trojan (1980a, 1980b) and Wasilkowski (1980, 1981a, 1981b, 1983).

There is also a very interesting stream of work for the approximate solution of scalar or

multivariate polynomial equations. It is assumed that complete information is available and given by the degree and coefficients of a polynomial. There is a number of interesting papers in that area. Some of them deal with average or probabilistic settings, some of them deal with different models of computation. A partial list includes: Hirsch and Smale (1979), Kim (1985), McMullen (1985), Murota (1982), Renegar (1984,1985a,1985b,1985c), Schönhage (1982), Shub and Smale (1985), Smale (1981,1985) and Wongkew (1985).

7. Iterative Model of Computation: Average Case.

There have been a number of important recent papers on average case algorithmic analysis of iterative methods for computing polynomial zeros.

Pioneering work is due to Shub and Smale(1985) who analyze the average behavior of Newton-type method for approximating zeros of complex polynomials. Also for polynomial zeros Renegar(1984, 1985a,1985b) investigates the simplicial-continuation algorithm due to Kuhn,(Kuhn et al. (1984)), and the multivariate Newton method.

Shub and Smale (1985) show that, on the average, six starts of a modified Newton's method are sufficient to obtain a point z, with $\mid f(z) \mid < \epsilon$, at cost proportional to $d(d + log1/\epsilon)$ for a complex normalized polynomial f of degree d. Renegar(1984) assumes the normalized Lebesque measure on the coefficients of complex polynomials and proves that for sets of polynomials of large measure the Kuhn's simplicial algorithm finds a point z with $|f(z)| < \epsilon$ and also $\mid z - \alpha \mid < \epsilon$, where $f(\alpha) = 0$, within $0(log1/\epsilon)$ steps. In Renegar(1985a, 1985b,1985c) he generalizes these results to the multivariate case.

We believe that a significant area for future research will be to obtain optimal imformation and algorithms in average case and probabilistic settings.

REFERENCES

Aird, T. J. and J. R. Rice, *Systematic Search in High Dimensional Sets*, SIAM J. Num. Anal. **14** (1977), 296–312.

Allgower, E. and K. Georg 1980, January, *Simplicial and Continuation Methods for Approximating Fixed Points and Solutions to Systems of Equations*, SIAM Review **22, No. 1**,, 28–85.

Booth, R. S. (1976) ,, *Location of Zeros of Derivatives*, SIAM J. Appl. Math., **15**,, 1495–1501.

Boult, T. and K. Sikorski (1984), *Can We Approximate Zeros of Functions with Non-zero Topological Degree?*, Report, Dept. Comp. Science, Columbia University.

Boult, T. and K. Sikorski (1985a), *Complexity of Computing Topological Degree of Lipschitz Functions in Two Dimensions*, Report, Dept. Comp. Science, Columbia University and University of Utah.

Boult, T. and K. Sikorski (1985b), *Complexity of Computing Topological Degree of Lipschitz Functions in n Dimensions*, Report, Dept. Comp. Science, Columbia University and University of Utah.

Boult, T. and K. Sikorski (1985c), *A Fortran Subroutine for Computing Topological Degree of Lipschitz Functions*, in progress.

Brent, R. P. (1973), "Algorithms for Minimization Without Derivatives", Prentice Hall,, Englewood Cliffs, New Jersey.

Brent, R.P. (1976a), "A Class of Optimal-Order Zero-Finding Methods Using Derivative Evaluations, in Analytic Computational Complexity(J.F. Traub, ed.)", Academic Press,, New York, pp. 59–73.

Brent, R. P. (1976b), "Multiple Precison Zero-Finding Methods and the Complexity of Elementary Function Evaluation, In Analytic Computational Complexity(J. F. Traub, ed.)", Academic Press, New York, pp. 151–176.

Brent, R. P. Winograd, S. and P. Wolfe (1973), *Optimal Iterative Processes for Root Finding*, Numer. Math. **20**, 327–341.

Chernousko, F. L. (1968), *An Optimal Algorithm for Finding the Roots of an Approximately Computed Function*, USSR Comp. Math. and Math. Phys. **1**, 1–24.

Eichhorn, B.H. (1968), "On Sequential Search, Selected Statistical Papers", Math. Centrum, Amsterdam, pp. 81–85.

Eiger, A., Sikorski, K. and F. Stenger (1984), *A Bisection Method for Systems of Nonlinear Equations*,, ACM TOMS **10, No. 4**, 367–377.

Gross, O. and S. M. Johnson (1959), *Sequential Minimax Search for a Zero of a Convex Function*, MTAC(now Math. Comp.), 13, 44-51.

Harvey, C. and F. Stenger (1976), January, *A Two-Dimensional Analogue to the Method of Bisections for Solving Nonlinear Equations*, Quant. Appl. Math., 351–367.

Hirsch, M. and Smale, S. (1979), *On Algorithms for Solving* $f(x) = 0$, Comm. Pure. Appl. Math. **2**, 281–312.

Hyafil, L. (1977), *Optimal Search for the Zero of the* $(n-1)$*-st Derivative*, IRIA/LABORIA Rep. no. 247.

Kacewicz, B. (1976), "The Use of Integrals in the Solution of Nonlinear Equations in N-dimensions, in Analytic Computational Complexity (J.F. Traub, ed.,)", Academic Press, New York, pp. 127–141.

Kacewicz, B. (1976b), *An Integral Interpolation Iterative Method for the Solution of Scalar Equations*, Numer. Math. 26, 355–365.

Kacewicz, B. (1979), *Integrals with a Kernel in the Solution of Nonlinear Equations in N Dimensions*, J. Assoc. Comput. Mach. 26 233–249.

Kearfott, R. B. (1977), *Computing the Degree of Maps and a Generalized Method of Bisection, Doctoral Dissertation*, University of Utah.

Kearfott, R. B. (1979), *An Efficient Degree Computation Method for a Generalized Method of Bisection*, Numer, Math. 32, 109–127.

Kiefer, J. (1953), *Sequential Minimax Search for a Maximum*, Proc. Amer. Math. Soc. 4, 502–505.

Kiefer, J. (1957), *Optimum Sequential Search and Approximation Methods Under Minimum Regularity Assumptions*, J. Soc. Indust. Appl. math. 5,, 105–136.

Kim, M. H., "Ph.D. thesis", C.U.N.Y. Graduate School, 1985.

Kuhn, H. W., Wang, Z. and S. Xu (1984), *On the Cost of Computing Roots of Polynomials*, Math. Prog. **28**, 156–163.

Kung, H. T. (1976), "The Complexity of Obtaining Starting Points for Solving Operator Equations by Newton's Method, in Analytic Computational Complexity (J. F. Traub, ed.)", Academic Press, New York, pp. 33–57.

Kung, H. T. and J. F. Traub (1974a), *Optimal Order of One-Point and Multipoint Iteration, in Complexity of Computation(R.H. Karp, ed.)*, SIAM-AMS Proc. **7**, 149-160.

Kung, H. T. and J. F. Traub (1976), *Optimal Order and Efficiency for Iterations with Two Evaluations*, SIAM J. Numer. Anal. 13,, 84–89.

Majstrovskij, G. D. (1972), *On the Optimality of Newton's Method*, Soviet, Math. Dokl. 13, 832–840.

McMullen, C. (1985), *Families of Rational Map s and Iterative Root-Finding Algorithms*, "Ph.D. thesis", Harvard University.

Meersman, R. (1976a), "On Maximal Order of Families of Iterations for Nonlinear Equations, Doctoral Thesis", Vrije Univ., Brussels.

Meersman, R. (1976b), "Optimal Use of Information in Certain Iterative Processes , in Analytic Computational Complexity(J.F. Traub. ed.)", Academic Press, New York, pp. 109–125.

Micchelli. C. A. and W. L. Miranker (1975), *High Order Search Methods for Finding Roots*, J. Assoc. Comput. Mach. 22, 52–60.

Murota, K., *Global Convergence of a Modified Newton Iteration for Algebraic Equations*, SIAM Journal of Numerical Analysis **19** (1982), 793–799.

Nemirovsky, A. S. and D. B. Yudin (1983), "Problem Complexity and Method Efficiency in Optimization", John Wiley and Sons, translated from Slozhnost Zadach I Effektivnost Metedov Optimizatsii, Nauka 1979.

Ortega, J. M. and W. C. Rheinboldt (1970), "Iterative Solution of Nonlinear Equations in Several Variables", Academic Press, New York.

Prüfer, M. and H. W. Siegberg (1980), *On Computational Aspects of Topological Degree in* R^n, Sonderforschungsbereich 72, Approximation und Optimierung, Universitat Bonn, preprint No. 252.

Renegar, J. (1984), *On the Efficiency of Newton's Method in Approximating all zeros of a System of Complex Polynomials*, Report, Math. Dept., Colorado State University.

Renegar, J. (1985a), *On the Complexity of a Piecewise Linear Algorithm for Approximating Roots of Complex Polynomials*, to appear.

Renegar, J. (1985b), *On the Cost of Approximating all Roots of a Complex Polynomial*, to appear.

Renegar, J. (1985c), *On the Efficiency of a Piecewise Linear Homotopy Algorithm in Approximating All Zeros of a System of Complex Polynomials*, Report, Math. Dept., Colorado State University.

Saari, D. G. and C. P. Simon (1978), *Effective Price Mechanisms*, Econometrica 46, 1097–1125.

Schönhage, A., *The Fundamental Theorem of Algebra of Alegbra in Terms of Computational Complexity*, Report , University of Tubingen,(1982).

Shub, M. and S. Smale (1985), *Computational Complexity: On the Geometry of Polynomials and a Theory of Cost*, Submitted for publication.

Sikorski, K. (1979), *A Three Dimensional Analogue to the Method of Bisection for Solving Nonlinear Equations*, Math. Comp. **33**, 722–738.

Sikorski, K. (1982), *Bisection is Optimal*, Numer. Math. 40, 111–117.

Sikorski, J. (1984a), *Optimal Solution of Nonlinear Equations Satisfying a Lipschitz Condition*, Numer. Math. 43, 225–240..

Sikorski, J. (1984b), *Study of Linear Information for Classes of Polynomial Equations*, Report, Dept. Comp. Science, Columbia University.

Sikorski, K. and G. M. Trojan (1984), *Asymptotic Optimality of the Bisection Method*, Numer. Math., to appear.

Sikorski, K. and H. Woźniakowski (1983), *For What Error Criteria Can We Solve Nonlinear Equations*, Report, Dept. Comp. Science, Columbia University.

Smale, S. (1981), *The Fundamental Theorem of Algebra and Complexity Theory*, Bull. Amer. Math. Soc. **4**, 1–36.

Smale, S. (1985), *On the Efficiency of Algorithms of Analysis*, Bull. Amer. Math. Soc. **13** (1985), 87–121.

Stenger, F. (1975), *Computing the Topological Degree of a Mapping in R^n*, Numer. Math. 25, 23–38.

Stynes, M. (1979a), *An Algorithm for Numerical Calculation of Topological Degree*, Appl. Anal. 9,, 63–77.

Stynes (1979b), *A Simplification of Stenger's Topological Degree Formula*, Numer. Math. 33, 147–156.

Stynes (1981), *On the Construction of Sufficient Refinements for Computation of Topological Degree*, Num. Math., 453–462.

Sukharev, A. G. (1976), *Optimal Search for a Zero of Function Satisfying a Lipschitz Condition*, USSR Comp. Math. and Math. Phys., 16 no. 1, 17–26.

Todd, M. J. (1978), *Optimal Dissection of Simplices*, SIAM J. Appl. Math., **34, no. 4**, 792–803.

Traub, J. F. (1961), *On Functional Iteration and the Calculation of Roots, Preprints of Papers*, 16th National ACM. Conf. Session 5A-1, 1-4, Los Angeles, California.

Traub, J. F. (1964), "Iterative Methods for the Solution of Equations", Prentice Hall, Englewood Cliffs, New Jersey.

Traub, J. F. and H. Woźniakowski (1976a), "Strict Lower and Upper Bounds on Iterative Computational Complexity, in Analytic Computational Complexity(J. F. Traub, ed.)", Academic Press, New York, pp. 15–34.

Traub. J F. and H. Woźniakowski (1976b), "Optimal Linear Information for the Solution of Nonlinear Operator Equations, in Algorithms and Complexity, New Directions and Recent Results (J. F. Traub, ed.)", Academic Press, New York, pp. 103–109.

Traub, J. F. and H. Woźniakowski (1979), *Convergence and Complexity of Newton's Iteration for Operator Equations*, J. Assoc. Comput. Mach. 26, 250–258.

Traub, J. F. and H. Woźniakowski (1980a), "A General Theory of Optimal Algorithms", Academic Press, New York.

Traub, J. F. and H. Woźniakowski (1980b), *Convergence and Complexity of Interpolatory-Newton Iterations in Banach Space*, Comput. Math. Appl. 6, 385–400.

Traub, J. F. and H. Woźniakowski (1980c), *Optimal Radius of Convergence of Interpolotary Iterations for Operator Equations*, Aequationes Mathematicae 21, no. 2-3, 159–172.

Traub, J. F. and H. Woźniakowski (1984), *Information and Computation*, Advances in Computers **23**, 35–92.

Traub, J. F., Wasilkowski, G. W. and H. Woźniakowski (1983), "Information, Uncertainty, Complexity", Addison, Wesley, Reading, Massachusetts.

Trojan, G. M. (1980a), *Optimal Iterative Methods for the Solution of Nonlinear Equations, Doctoral Thesis (in Polish)*, Institute of Informatics, University of Warsaw.

Trojan, G. M. (1980b), *Tight Bounds on the Complexity Index of One Point Iteration*, Comp. Meths. Appl. 6, 431–433.

Wasilkowski, G. W. (1980), *Can Any Stationary Iteration Using Linear Information be Globally Convergent?*, J. Assoc. Comput. Mach. 27, 263–269.

Wasilkowski, G. W. (1981a), *n-Evaulation Conjecture for Multipoint Itrations for the Solution of Scalar Nonlinear Equations*, J. Assoc. Comput. Mach. 28, 71–80.

Wasilkowski, G. W. (1981b), *The Strength of Nonstationary Iteration*, Aequationes Mathematicae 24, 243–260.

Wongkew, R., "Ph.D. thesis", University of California at Berkeley,(1985).

Woźniakowski. H. (1972), *On Nonlinear Iterative Processes in Numerical Methods, Doctoral Thesis (in Polish)*, University of Warsaw.

Woźniakowski, H. (1974), *Maximal Stationary Iterative Methods for the Solution of Operator Equations*, SIAM J. Numer. Anal. 11, 934–949.

Woźniakowski, H. (1975a), *Generalized Information and Maximal Order of Iteration for Operator Equations*, SIAM J. Numer. Anal. 12, 121–135.

Woźniakowski, H. (1975b), *Properties of Maximal Order Methods for the Solution of Nonlinear Equations, Z. Agnew*, Math. Mech. 55, 268–271.

Woźniakowski, H. (1976), "Maximal Order of Multipoint Iterations Using n Evaluations, in Analytic Computational Complexity (J. F. Traub, ed.)", Academic Press, New York, pp. 75–107.

Woźniakowski, H. (1985), *A Survey of Information-Based Complexity*, J. Complexity **1**.