

THE FUNDAMENTAL THEOREM OF ALGEBRA AND COMPLEXITY THEORY¹

BY STEVE SMALE

PART I

1. The main goal of this account is to show that a classical algorithm, Newton's method, with a standard modification, is a tractable method for finding a zero of a complex polynomial. Here, by "tractable" I mean that the cost of finding a zero doesn't grow exponentially with the degree, in a certain statistical sense. This result, our main theorem, gives some theoretical explanation of why certain "fast" methods of equation solving are indeed fast. Also this work has the effect of helping bring the discrete mathematics of complexity theory of computer science closer to classical calculus and geometry.

A second goal is to give the background of the various areas of mathematics, pure and applied, which motivate and give the environment for our problem. These areas are parts of (a) Algebra, the "Fundamental theorem of algebra," (b) Numerical analysis, (c) Economic equilibrium theory and (d) Complexity theory of computer science.

An interesting feature of this tractability theorem is the apparent need for use of the mathematics connected to the Bieberbach conjecture, elimination theory of algebraic geometry, and the use of integral geometry.

Before stating the main result, we note that the practice of numerical analysis for solving nonlinear equations, or systems of such, is intimately connected to variants of Newton's method; these are iterative methods and are called fast methods and generally speaking, they are fast in practice. The theory of these methods has a couple of components; one, proof of convergence and two, asymptotically, the speed of convergence. But, not usually included is the total cost of convergence.

On the other hand, there is an extensive theory of search methods of solution finding. This means that a region where a solution is known to exist is broken up into small subsets and these are tested in turn by evaluation; the process is repeated. Here it is simpler to count the number of required steps and one has a good knowledge of the global speed of convergence. But, generally speaking, these are slower methods which are not used by the practicing numerical analyst.

The contrast between the theory and practice of these methods, in my

Presented to the Symposium on the Mathematical Heritage of Henri Poincaré.

1980 *Mathematics Subject Classification*. Primary 00-02, 12D10, 68C25, 65H05, 58-02; Secondary 01A05, 30D10.

¹Partially supported by NSF Grant MCS-7717907 and the Miller institute at the University of California, Berkeley.

mind, has to do with the fact that search methods work inexorably and the analysis of cost goes by studying the worst case; but in contrast the Newton type methods fail in principle for certain degenerate cases. And near the degenerate cases, these methods are very slow. This motivates a statistical theory of cost, i.e., one which applies to most problems in the sense of a probabilistic measure on the set of problems (or data). There seems to be a trade off between speed and certainty, and a question is how to make that precise.

One clue can be taken from the problem of complexity in the discrete mathematics of theoretical computer science. The complexity of an algorithm is a measure of its cost of implementation. In these terms, problems (or algorithms) which depend on a "size" are said to be tractable provided the cost of solution does not increase exponentially as their size increases. The famous $P = NP$ problem of Cook and Karp lies in this framework.

In the case of a single polynomial the obvious "size" is the degree d . So these considerations pose the problem. Given $\mu > 0$, an allowable probability of failure, does the cost of computation via the modified Newton's method for polynomials in some set of probability measure $1-\mu$, grow at most as a polynomial in d ? Moreover, one can ask that as μ varies, this cost be bounded by a polynomial in $1/\mu$. I was able to provide an affirmative answer to these questions.

Let me be more precise. The problem is to solve $f(z^*) = 0$ where $f(z) = \sum_{i=0}^d a_i z^i$, $a_i \in \mathbf{C}$ and $a_d = 1$. The algorithm is the modified Newton's method given by: let $z_0 \in \mathbf{C}$ and define inductively $z_n = T_h(z_{n-1})$ where $T_h(z) = z - hf(z)/f'(z)$ for some h , $0 < h \leq 1$. If $h = 1$, this is exactly Newton's method.

We will say that z_0 is an *approximate* zero provided if taking $h = 1$, the sequence z_n is well defined for all n , z_n converges to z^* as $n \rightarrow \infty$, with $f(z^*) = 0$ and $|f(z_n)/f(z_{n-1})| < \frac{1}{2}$ for all $n = 1, 2, \dots$.

Practically and theoretically this is a reasonable definition. One could say that in this case, z_0 is in a strong Newton Sink.

Let \mathcal{P}_d be the space of polynomials f , $f(z) = \sum_{i=0}^d a_i z^i$, $a_d = 1$. Thus \mathcal{P}_d can be identified with \mathbf{C}^d , with coordinates $(a_0, \dots, a_{d-1}) = a \in \mathbf{C}^d$.

Define

$$P_1 = \{f \in \mathcal{P}_d \mid |a_i| < 1, i = 0, \dots, d-1\}$$

and use normalized Lebesgue measure on P_1 , for a probability measure.

MAIN THEOREM. *There is a universal polynomial $S(d, 1/\mu)$, and a function $h = h(d, \mu)$ such that for degree d and μ , $0 < \mu < 1$, the following is true with probability $1-\mu$. Let $x_0 = 0$. Then $x_n = T_h(x_{n-1})$ is well defined for all $n > 0$ and x_s is an approximate zero for f where $s = S(d, 1/\mu)$.*

More specifically we can say, if $s \geq [100(d+2)]^9/\mu^7$, then with probability $1-\mu$, x_s is well defined by the algorithm for suitable h and x_s is an approximate zero of f .

Note especially that h and s do not depend on the coefficients.

The use of probability is made more precise in the following very brief idea

of the proof. There is a certain subset $W_* \subset \mathcal{P}_d$ such that $f \in W_*$, $z_0 = 0$ is a “worst case” for the algorithm “in the limit” $h \rightarrow 0$. We don’t expect the algorithm to work in this case, no matter how small h is taken. But if $f \notin W_*$ the algorithm will converge for sufficiently small h . It will be shown that W_* is a real algebraic variety in \mathcal{P}_d using elimination theory. Then a certain family of open neighborhoods Y_σ of W_* , $0 < \sigma < 1$, are described, decreasing to W_* as $\sigma \rightarrow 0$.

Using a theorem of Weyl on the volume of tubes, and formulae of integral geometry (Santaló) we are able to estimate the volume of Y_σ .

The idea then is that if $\text{Vol } Y_\sigma / \text{Vol } P_1 < \mu$ and $f \notin Y_\sigma$, then the algorithm with a suitable choice of h will arrive at an approximate zero after s steps.

The algorithm is “tracked” in the target space of $f: \mathbb{C} \rightarrow \mathbb{C}$. Thus we want to estimate $|f(z_n)/f(z_{n-1})|$ in terms of values $f(\theta)$ of critical points θ where $f'(\theta) = 0$.

What is needed can be seen more precisely in terms of the Taylor series expansion

$$\frac{f(z')}{f(z)} = 1 - h + \sum_{d=2}^d (-h)^k \frac{f^{(k)}(z)f}{n!} \frac{f(z)^{k-1}}{f'(z)^k}$$

where $z' = z - hf(z)/f'(z)$.

This motivates

THEOREM (1A OF §2, PART II). *If f is a polynomial with $f'(z) \neq 0$, then there is a critical point θ (i.e. $f'(\theta) = 0$) such that*

$$\frac{|f^{(k)}(z)|}{k!} \frac{|f(\theta) - f(z)|^{k-1}}{|f'(z)|^k} \leq 4^{k-1}, \quad k = 2, 3, \dots$$

Thus if $z = 0$, $f(0) = 0$, $f(z) = \sum a_k z^k$, then

$$|f(\theta)| |a_k / a_1^k|^{1/(k-1)} \leq 4.$$

The proof uses mathematics related to the Bieberbach conjecture. The particular result used is due to Loewner.

In general, a number of related questions remain unsolved and new ones are suggested. Part III is devoted to these. The proof of the main result is in Part II, while the rest of Part I is devoted to background material.

Names which are italicized are listed in the bibliography at the end.

There is a final comment on the spirit of the paper. I feel one problem of mathematics today is the division into the two separate disciplines, pure and applied mathematics. Oftentimes it is taken for granted that mathematical work should fall into one category or the other. This paper was not written to do so.

I would like to acknowledge useful conversations, with a number of mathematicians including L. Blum, S. S. Chern, G. Debreu, D. Fowler, W. Kahan, R. Osserman, R. Palais, G. Schober and H. Wu.

Special thanks are due Moe Hirsch and Mike Shub.

2. There is a sense in which an important result in Mathematics is never

finished. In particular one might ask “Has the fundamental theorem of algebra been proved satisfactorily?”

What do historians of mathematics say about this? They most frequently assert that the first proof of the Fundamental Theorem of Algebra was in Gauss’ thesis. Here are some examples.

D. Struik (p. 115):

“GAUSS: THE FUNDAMENTAL THEOREM OF ALGEBRA. The first satisfactory proof of this theorem was presented by Carl Freidrich Gauss (1777–1855) in his Helmstädt doctoral dissertation of 1799”

D. E. Smith (pp. 473–474):

“Fundamental Theorem” After these early steps the statement was repeated in one form or another by various later writers, including Newton (c. 1685) and Maclaurin (posthumous publication, 1748). D’Alembert attempted a proof of the theorem in 1746, and on this account the proposition is often called d’Alembert’s theorem. Other attempts were made to prove the statement, notably by Euler (1749) and Lagrange but the first rigorous demonstration is due to Gauss (1799), with a simple treatment in (1849).”

H. Eves (p. 372):

“In his doctoral dissertation, at the University of Helmstädt and written at the age of twenty, Gauss gave the first wholly satisfactory proof of the *fundamental theorem of algebra*”

And finally Gauss himself, fifty years later, as related by D. E. Smith 1929, in *Source book in mathematics*, McGraw-Hill, New York, pp. 292–293:

“The significance of his first proof in the development of mathematics is made clear by his own words in the introduction to the fourth proof: ‘the first proof] . . . had a double purpose, first to show that all the proofs previously attempted of this most important theorem of the theory of algebraic equations are unsatisfactory and illusory, and secondly to give a newly constructed rigorous proof.’”

On the other hand, compare this with the following passages of Gauss’ thesis translated into English from the Latin in *Struik*:

“Now it is known from higher geometry that every algebraic curve (or the single parts of an algebraic curve when it happens to consist of several parts) either runs into itself or runs out to infinity in both directions, and therefore if a branch of an algebraic curve enters into a limited space, it necessarily has to leave it again.”⁷ . . . and ⁷[footnote by Gauss], “It seems to be sufficiently well demonstrated that an algebraic curve can neither be suddenly interrupted (as e.g. occurs with the transcendental curve with equation $y = 1/\log x$), not lose itself after an infinite number of terms (like the logarithmic spiral), and nobody, to my knowledge, has ever doubted it. But if anybody desires it, then on another occasion I intend to give a demonstration which will leave no doubt”

These passages from Gauss are interesting for several reasons and we will return to them. But for the moment, I wish to point out what an immense gap Gauss’ proof contained. It is a subtle point even today that a real algebraic plane curve cannot enter a disk without leaving. In fact even though Gauss redid this proof 50 years later, the gap remained. It was not until 1920 that

Gauss' proof was completed. In the reference *Gauss*, A. Ostrowski has a paper which does this and gives an excellent discussion of the problem as well (I am grateful to Horst Simon for providing me with an English translation of Ostrowski's paper).

One can understand the historical situation better perhaps from the point of view of *Imre Lakatos*. Lakatos in the tradition of Hegel, on one hand, and Popper, on the other, sees mathematics as a development which proceeds as a series of "proofs and refutations".

As an example of his critique, in connection with the origin of the concept of real function, *Lakatos* writes (p. 151):

"Some infallibilist historians of mathematics use here the ahistorical technique of condensing a long development full of struggle and criticism into one single action of infallible insight and attribute to Dirichlet the maturity of later analysts."

It seems as if the idea of this quote could also be applied to the written history of Gauss' thesis.

Another line of questioning of the proof of the fundamental theorem of algebra was undertaken by the Constructivists. *Brouwer* and *Weyl* in 1924 both published articles which gave constructive proofs for yielding a zero of a complex polynomial. But as is emphasized by the computer scientist, what good is a constructive solution if it takes 10^{10} years with the fastest computers (say even fastest in principle). Thus a Constructivist approach to be satisfactory today should be paired with a theorem on the speed or cost of computation. In fact in some of the literature on root finding methods, there is a successful effort to measure the number of steps. See for example *Dejon* and *Henrici*, *Henrici*, and *Collins*. Also, the book by *Ostrowski* gives a very useful account of fast algorithms for solving systems of equations in general and for finding a zero of a polynomial in particular.

3. Here we give some of the background of our project related to economics and numerical analysis. Why economics?! There are a couple of related reasons. I was brought to the complexity questions through my work in economics. Also economic theory, besides being concerned with the existence of equilibria, has seriously considered the computational question of finding equilibria as well.

About 100 years ago the great early economic theorist L. Walras saw economic equilibrium for several markets as a solution to a system of (nonlinear) equations, supply equals demand. In this model, supply and demand functions are generated by the individual agents of the economy, consumers and producers.

The modern rigorous development of this theory is due especially to Arrow and Debreu and can be found in *Debreu*. Arrow and Debreu transform the equation of supply equals demand into a fixed point problem, and then apply a fixed point theorem (that of Kakutani, or in simpler versions, Brouwer's). In this way a coherent unifying structure is given to classical economic theory.

Scarf has developed a technique for computing these economic equilibria, and fixed points generally. Although Scarf is himself a mathematical

economist, his method falls into the area of operations research (techniques of pivoting, etc.).

Working in equilibrium theory and following the work of Scarf, I perceived the problem of the existence of an equilibrium in terms of solving a system of equations (closer to Walras). In particular, one tries to find a price system $p = (p_1, \dots, p_l) \in \mathbf{R}_+^l$ which makes the excess demand Z (supply minus demand) zero. Now the values of Z are commodities, so that $Z: \mathbf{R}_+^l \rightarrow \mathbf{R}^l$ is a *morphism* whose source space consists of price systems and target space consists of commodity bundles. Fixed point theory deals with *endomorphisms* or maps of a space into itself. Thus I found it more natural to solve the equation $Z(p) = 0$ by directly constructing a solution in contrast to the method of artificially transforming it into a fixed point problem, and then either using a fixed point theorem or Scarf's method. In fact, eventually, the existence theorems of economic equilibrium theory could be proved directly and constructively this way (see *Smale* (to appear)).

I abstracted these ideas to the general problem of solving a system of nonlinear equations. The method was analogous to Scarf's method in some ways, but it emphasized the equation approach (versus fixed points) on one hand and differentiability on the other hand. I called it the "Global Newton Method", because locally it was essentially Newton's method and it worked globally (see *Smale* (1976)). It was developed further, in several respects including polynomial systems in *Hirsch-Smale*.

Having proved convergence theorems, it was natural to consider the question of speed of convergence and thus the problems of the present paper.

In some sense then there is a little of the Global Newton methodology in Newton, but what is more interesting is its connection to the thesis of Gauss.

To see that connection better, consider the basic idea for the case of a complex polynomial $f: \mathbf{C} \rightarrow \mathbf{C}$ (although, the same construction works for polynomial maps $f: \mathbf{C}^m \rightarrow \mathbf{C}^m$). To emphasize the "morphism" aspect we can write $f: \mathbf{C}_{\text{source}} \rightarrow \mathbf{C}_{\text{target}}$, and we want to solve $f(z) = 0_{\text{target}}$. Suppose $f(z_0)$ does not lie on one of the finite number of rays in $\mathbf{C}_{\text{target}}$ which contain the critical values $f(\theta)$ (i.e., $f'(\theta) = 0$). Of course such $z_0 \in \mathbf{C}_{\text{source}}$ exists. Now take the segment from $f(z_0)$ to 0 in $\mathbf{C}_{\text{target}}$ and lift it back by f^{-1} to a curve starting at z_0 in $\mathbf{C}_{\text{source}}$. This lifting cannot go to ∞ or to a critical point and therefore the lifted curve goes to z with $f(z) = 0$. This proves the fundamental theorem of algebra; it also is the basic idea of the "Global Newton Method" and the basic idea of this paper. For sufficiently differentiable maps $f: \mathbf{R}^m \rightarrow \mathbf{R}^m$ which are not polynomial maps, one uses Sard's theorem to choose the z_0 , and the argument becomes a bit more subtle.

Now we can compare this construction with the passage in Gauss' thesis quoted earlier in §2. There is a reasonable sense in which Gauss attempts to find a zero of a polynomial by following an algebraic curve.

Pont (p. 32) is right in relating this footnote to the precursors of algebraic topology. More exactly, *Pont* refers to a sentence later in the same footnote where Gauss speaks of the principles of the geometry of position "which are no less valid than the principles of the geometry of magnitudes".

Besides the mentioned references, there are points of contact between this

Global Newton method I have described previously, and the work of Branin, Davidenko, *Hirsch*, *Kellog-Li-Yorke* and many others. See *Hirsch-Smale* for more information.

4. I would like to say a few words on the discrete mathematics of complexity theory of computer science. Even though the mathematics is somewhat removed from mine, it was an important element in my motivation here. Complexity theory made me aware of the problem of computational cost of algorithms that I had been working with (i.e., the “Global Newton” method).

It was the computer scientists who focused on the key question of polynomial versus exponential bounds on the cost. And of course that is important in formulating my theorem here. Also computer scientists have seen the importance of dealing with problems statistically and not just studying the (often very slow) worst case.

More particularly, let me refer to *Hartmanis* for a useful perspective and history with references. One can see *Garey and Johnson* for an account of the problem $P = NP$ associated to Cook and Karp. Traub and Wozniakowski have explicitly discussed complexity in the analytic framework; see *Traub*. In some of the work in numerical analysis and algebraic root finding, one finds results on the cost. See §2 of this paper.

5. We give some further details on the geometric background to the algorithm used in this paper. This geometry is based on an idea which is known, but not usually explicated. Namely, *Newton’s method for solving $f(z) = 0$ is an Euler approximation to the ordinary differential equation*

$$dz/dt = -\frac{1}{2} \text{grad}|f|^2. \quad (*)$$

The prior idea can be restated to be true quite generally; see *Smale* (1976) and *Hirsch-Smale*. Here we restrict ourselves to the case of a single complex polynomial f .

In (*) $\text{grad}|f|^2$ denotes the gradient vector field of the real valued function $|f|^2$ on $\mathbf{R}^2 = \mathbf{C}$, $\text{grad}|f|^2 = (\partial|f|^2/\partial x, \partial|f|^2/\partial y)$. Thus $dz/dt = -\frac{1}{2} \text{grad}|f|^2$ is a real ordinary differential equation (1st order, autonomous) on \mathbf{R}^2 . Let f' denote the (complex) derivative of f and \bar{f} the complex conjugate. Then it is easily shown that

$$\text{LEMMA. } \frac{1}{2} \text{grad}|f|^2 = f\bar{f}'.$$

As a consequence $-\frac{1}{2} \text{grad}|f|^2 = \rho(z)(-f(z)/f'(z))$ where $\rho: \mathbf{R}^2 \rightarrow \mathbf{R}$ is the positive scalar function, $\rho(z) = f'(z)\bar{f}'(z)$.

Therefore the vector fields on \mathbf{R}^2 defined by $-f(z)/f'(z)$ and $-\frac{1}{2} \text{grad}|f|^2$ differ only by a rescaling, which has the effect of desingularizing $-f(z)/f'(z)$ at the critical points z where $f'(z) = 0$.

An Euler approximation (see *Hurewicz*) of one of these two vector fields may thus be regarded as an Euler approximation of the other, by changing the step size by ρ .

In particular, if z_0 is fixed then

$$z_{n+1} = z_n - h_n f(z_n) / f'(z_n), \quad n = 1, 2, \dots, \quad (**)$$

with $h_n > 0$ is such an Euler approximation, and if $h_n = 1$, this is Newton's method. Generally speaking (**) describes an algorithm, and that is the algorithm studied in this paper where $h_n = h$ is independent of n ; we will also assume $0 < h \leq 1$.

From these considerations some geometrical insight into the algorithm can be obtained by looking at the solution structure of (*) in $\mathbf{R}^2 = \mathbf{C}$. We will do this by stating some facts which can be proved without great difficulty, and by giving an example. First the example.

Let $f(z) = z^2 - a^2$, $a > 0$. Then the "phase portrait" of (*) is given by Figure 1.

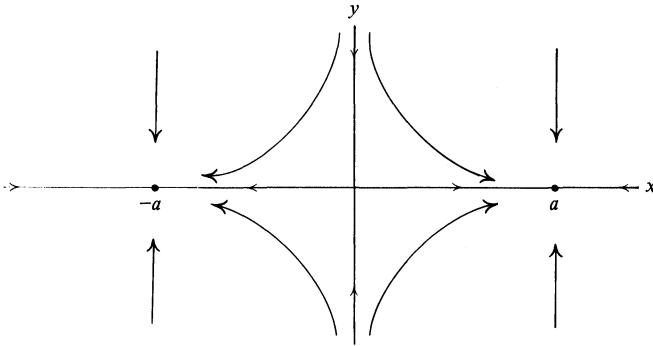


FIGURE 1

Thus the point at ∞ is a single source, and there are two sinks, a and $-a$, the zeros of f . There is one saddle point, namely the origin which is a critical point of f .

The solution curves, which are purely imaginary don't approach the zeros of f , while all others do. Thus one may expect that if the real part of z_0 is not zero and $h > 0$ is small enough, then (*) converges to a or $-a$. That is true and is the basis for the general success of (*).

For general complex polynomials f , ∞ is the single source, and the sinks are precisely the zeros of f . The zeros of the vector field, i.e., the equilibria of the equation (*) other than the zeros of f , are the zeros of f' .

These zeros of f' or the critical points (when distinct from the zeros of f) are saddle points when their multiplicity is one.

One other fact of geometrical importance for us here is that the solution curves of (*) are the same as the inverse image of rays in \mathbf{C} under the map $f: \mathbf{C} \rightarrow \mathbf{C}$. This explains why, in the subsequent analysis, we are especially concerned with what happens in the "target" space. See also §3 of this paper.

I emphasize that the above considerations are not used in our proof of the main theorem, but may provide helpful background geometry.

PART II

1. Here is a simple consequence of Theorem 1. Suppose that the polynomial

$$f(z) = z + a_2z^2 + \dots + a_dz^d$$

has no critical values $f(\theta)$ in the unit disk in \mathbf{C} (i.e., $|f(\theta)| \geq 1$ if $f'(\theta) = 0$). Then $|a_2| \leq 2$ and it is perhaps true that $|a_2|$ is always less than $\frac{1}{2}$. It is somewhat subtle even that $|a_2|$ has any bound (independent of d).

THEOREM 1. *Let $f(z) = \sum_{i=0}^d a_i z^i$, a complex polynomial with $a_0 = 0$, $a_1 \neq 0$. Then there is a critical point $\theta \in \mathbf{C}$ (i.e., $f'(\theta) = 0$) such that for every $k = 1, 2, 3, \dots$*

$$|a_k/a_1|^{1/(k-1)}|f(\theta)|/|a_1| < 4.$$

REMARK 1. We will sometimes write instead

$$|a_k/a_1|^{1/(k-1)}|f(\theta)|/|a_1| < K.$$

Thus K could be 4. On the other hand 4 may not be the sharpest value. I suspect one could take $K = 1$, but, we will give examples which show that $K \geq 1$. See also problem number 1 of Part III.

REMARK 2. The proof yields sharper results for each k . More precisely

$$|a_k/a_1|^{1/(k-1)}|f(\theta)|/|a_1| \leq \beta_k$$

where $\beta_2 = 2$, $\beta_3 = \sqrt{5}$, $\beta_4 = (14)^{1/3}$, $\beta_5 = (42)^{1/4} \sim 2.55$, $\beta_6 \sim 2.61$, $\beta_7 \sim 2.65$, $\beta_{20} \sim 3.29$ and of course $\beta_k \leq 4$.

REMARK 3. The theorem is false in a quite extreme way for entire functions. There is no bound at all for $k = 2$ or any k . This is even true for entire functions which have no critical points in \mathbf{C} . For example let $f(z) = (1/\alpha)e^{\alpha z} - 1/\alpha$, $\alpha > 1$.

The proof of Theorem 1 goes by a theorem of Loewner. See *Hayman Jenkins*, or *Schober*. Loewner proved this theorem at the same time that he showed that $|a_3| \leq 3$ in the Bieberbach conjecture (1923).

Define D_r to be the disk $\{z \in \mathbf{C} \mid |z| < r\}$ throughout.

THEOREM (LOEWNER). *Let $g: D_1 \rightarrow \mathbf{C}$ be a "schlicht" function, i.e. $g(z)$ can be expressed as a convergent power series, $g(z) = \sum_{i=0}^{\infty} b_i z^i$, $|z| < 1$, with $b_0 = 0$, $b_1 = 1$ and g is 1-1. Suppose that $f: \Omega \rightarrow D_1$ is an inverse to g with $0 \in \Omega$. Let $f(z) = \sum_{i=0}^{\infty} a_i z^i$ near 0 (therefore $a_0 = 0$, $a_1 = 1$). Then*

$$|a_k| \leq B_k, \quad k = 1, 2, \dots$$

where

$$B_k = 2^k \frac{1 \cdot 3 \dots (2k-1)}{1 \cdot 2 \dots (k+1)}.$$

We need to extend the theorem to the case where D_1 is replaced by D_R and $b_1 \neq 0$ is arbitrary.

EXTENDED LOEWNER'S THEOREM. *Let $g: D_R \rightarrow \mathbf{C}$, ($R > 0$) be 1-1, $g(z) = \sum_{i=0}^{\infty} b_i z^i$, $b_0 = 0$, $b_1 \neq 0$ and $f: \Omega \rightarrow D_R$ be an inverse to g with $0 \in \Omega$. Let*

$f(z) = \sum_{i=0}^{\infty} a_i z^i$ near 0. Then

$$|a_k/a_1|^{1/(k-1)} R/|a_1| \leq B_k^{1/(k-1)}.$$

PROOF. First let f, g be as in the extended theorem with $R = 1$. Note that $a_1 \neq 0$ and in fact $a_1 = f'(0) = 1/b_1 = 1/g'(0)$. Let $g_0(\omega) = (1/b_1)g(\omega)$ and $f_0(z) = f(z/a_1)$. So

$$f_0(g_0(\omega)) = f\left(\frac{1}{a_1}\left(\frac{1}{b_1}g(\omega)\right)\right) = \omega \quad \text{for } |\omega| < 1.$$

Since

$$f_0(z) = z + \frac{a_2}{a_1^2} z^2 + \frac{a_3}{a_1^3} z^3 + \dots,$$

and since Loewner's theorem applies to g_0 , we have

$$|a_k/a_1^k| \leq B_k.$$

Thus we conclude

LEMMA 1. If $g: D_1 \rightarrow \mathbf{C}$ is 1-1, $g(0) = 0$, $g'(0) \neq 0$ and $f: \Omega \rightarrow D_1$, $0 \in \Omega$, $f(0) = 0$, $f(z) = a_1 z + a_2 z^2 + \dots$ is its inverse, then

$$|a_k/a_1|^{1/(k-1)} 1/a_1 \leq B_k^{1/(k-1)}.$$

Next let $g: D_R \rightarrow \mathbf{C}$, and f be as in the extended theorem and let $g_1(w) = g(Rw)$. Then g_1 and f_1 satisfy Lemma 1 where $f_1(z) = (1/R)f(z)$, since $f_1(g_1(w)) = (1/R)f(g(Rw)) = w$. Also $f_1(z) = (a_1/R)z + (a_2/R)z^2 + \dots$. Apply Lemma 1 to obtain

$$|a_k/a_1|^{1/(k-1)} R/|a_1| \leq B_k^{1/(k-1)}.$$

This proves the theorem.

LEMMA 2. $B_k^{1/(k-1)} \leq 4$.

PROOF.

$$B_k = 2^{k-1} \cdot \frac{2}{k+1} \cdot \frac{1 \cdot 3 \cdot 5 \cdots 2k-1}{1 \cdot 2 \cdot 3 \cdots k} \leq 2^{k-1} \left(\frac{2}{k+1}\right) 2^{k-1} < 2^{k-1} \cdot 2^{k-1}.$$

It follows.

LEMMA 3. Let $f(z)$ be a polynomial of exact degree d , $f(0) = 0$ and such that

$$R = \min_{\substack{\theta \\ f'(\theta)=0}} |f(\theta)| > 0.$$

Then there exists an analytic function $g: D_R \rightarrow \mathbf{C}$ with $g(0) = 0$, and $f(g(w)) = w$ for all $w \in D_R$ and g is 1-1.

PROOF. First note that

$$f: \mathbf{C} - \bigcup_{\substack{\theta \\ f'(\theta)=0}} f^{-1}(f(\theta)) \rightarrow \mathbf{C} - \bigcup_{f'(\theta)=0} f(\theta)$$

is proper (the inverse image of a compact set is compact) and a local

homeomorphism. It is a covering map. Therefore we may lift it back by g over D_R , uniquely if $g(0) = 0$. This proves Lemma 3.

Combining Lemmas 2 and 3 with the Extended Loewner Theorem yields Theorem 1.

We give some lower bounds for the constant $K = K_{d,k}$ of Remark 1 of Theorem 1. Thus for each k, d as in Theorem 1, let $C_{d,k}$ be the lower bound exhibited by $f(z) = (z - 1)^d - (-1)^d$ if $k < d$ and $f(z) = z^d - dz$ if $k = d$. Then

PROPOSITION. $C_{d,d} = ((d - 1)/d)(1/d)^{d-1}$.

$$C_{d,k} = \left(\frac{(d - 1)!}{k! (d - k)!} \right)^{1/(k-1)} \frac{1}{d} \quad \text{if } k < d.$$

COROLLARY. K OF (REMARK 1). *Theorem 1 must be at least 1 (it is independent of k, d).*

The proof of the Proposition is a direct calculation which we omit.

Note that $C_{d,2} = \frac{1}{2}(d - 1)/d$ for all $d > 2$ and that $C_{d,k} \leq (d - 1)/d$ for all k, d . Note also that

$$\begin{aligned} \max_d C_{d,2} &= \frac{1}{2}, & \max_d C_{d,3} &= \left(\frac{1}{6}\right)^{1/2}, \\ \max_d C_{d,k} &= \left(\frac{k - 1}{k}\right) \left(\frac{1}{k}\right)^{1/(k-1)}, & k &> 3. \end{aligned}$$

This gives the Corollary.

We end this section by remarking that there is a large related literature on the critical points of polynomials; see *Marden*.

2. Here we make some estimate on the effect of one step of our algorithm. As usual we are taking a target space perspective, i.e., estimating values. In the next section the result is used to estimate the effect of many steps.

Let K be as in Remark 1 after Theorem 1, e.g., $K = 4$.

THEOREM 2. *Let f be a polynomial and $z \in \mathbb{C}$ such that $f(z)$ and $f'(z)$ are not zero. Let*

$$h_0 = \min_{f'(\theta)=0} \frac{1}{K} \left| \frac{f(\theta) - f(z)}{f(z)} \right|$$

and $0 < h < h_0$. If $z' = z - hf(z)/f'(z)$, then

$$f(z')/f(z) = 1 - h + \alpha h^2 / (h_0 - h)$$

for some $\alpha \in \mathbb{C}$, $|\alpha| \leq 1$.

The proof uses the following version of Theorem 1.

THEOREM 1A. *Let f be a polynomial and $z \in \mathbb{C}$ such that $f'(z) \neq 0$. There is*

a critical point θ of f such that for all $k = 2, 3, \dots$,

$$\left| \frac{f^{(k)}(z)}{k!f'(z)} \right|^{1/(k-1)} \leq \frac{K}{|f(z) - f(\theta)|}.$$

Furthermore, if $f(z) \neq 0$, and h_0 is as in Theorem 2, then for $k = 2, 3, \dots$

$$\left| \frac{f^{(k)}(z)}{k!} \frac{f(z)^{k-1}}{f'(z)^k} \right|^{1/(k-1)} \leq \frac{1}{h_0}.$$

Note first that the last statement of Theorem 1A is a consequence of the first part. We now check the first part. For this let f, z be as in Theorem 1A and expand f about z so that $f(w) = \sum_{k=0}^d (f^{(k)}(z)/k!)(w-z)^k$. Define the polynomial g by $g(u) = \sum_{k=1}^d (f^{(k)}(z)/k!)u^k$. So $g(w-z) = f(w) - f(z)$ and $g'(w-z) = f'(w)$. Apply Theorem 1 to g to obtain

$$\left| \frac{f^k(z)}{k!f'(z)^k} \right|^{1/(k-1)} |g(\sigma)| \leq K \quad \text{for some } \sigma \text{ with } g'(\sigma) = 0.$$

Thus

$$\left| \frac{f^k(z)}{k!f'(z)^k} \right|^{1/(k-1)} |f(\theta) - f(z)| \leq K \quad \text{for } \theta = z + \sigma.$$

This proves 1A.

We now prove Theorem 2. Again expand f by Taylor's series about z , using $z' = z - hf(z)/f'(z)$, to obtain $f(z')/f(z) = 1 - h + h\gamma$ where

$$\gamma = \sum_{k=2}^d (-h)^{k-1} \frac{f^{(k)}(z)f(z)^{k-1}}{k!f'(z)^k}.$$

Now by Theorem 1A we obtain

$$|\gamma| \leq \sum_{k=2}^d \left| \frac{h}{h_0} \right|^{k-1}.$$

Thus

$$|\gamma| \leq \sum_{i=1}^{\infty} \left| \frac{h}{h_0} \right|^i = \sum_{i=0}^{\infty} \left| \frac{h}{h_0} \right|^i - 1 \leq \frac{1}{1 - h/h_0} - 1 = \frac{h}{h_0 - h}.$$

Let $\alpha = ((h_0 - h)/h)\gamma$; then $|\alpha| \leq |\gamma| |(h_0 - h)/h| \leq 1$.

This proves Theorem 2.

We apply Theorem 2 to obtain information about Newton's method.

COROLLARY A. Let $C \geq 1$ and $\rho_f = \rho = \min_{\theta: f'(\theta)=0} |f(\theta)|$ for a polynomial f . If $|f(z)| < \rho/CK + K + 1$ then $|f(z')/f(z)| < 1/C$ where $z' = z - f(z)/f'(z)$ (Newton's method).

PROOF OF COROLLARY A. We first remark that for any critical point θ , $|f(\theta) - f(z)| + |f(z)| \geq |f(\theta)|$ and $|f(\theta) - f(z)| \geq |f(\theta)| - |f(z)| \geq \rho - |f(z)|$. By hypothesis, $|f(z)|(CK + K + 1) < \rho$, so $|f(z)|K(1 + C) < \rho -$

$|f(z)|$. Thus by the prior remark, for any critical point θ , $|f(z)|K(1 + C) < |f(\theta) - f(z)|$.

Thus we have shown that for any critical point θ ,

$$|f(\theta) - f(z)|/K|f(z)| > 1 + C.$$

So for h_0 of Theorem 2, we have $h_0 > 1 + C$ and $1/(h_0 - 1) < 1/C$.

Apply Theorem 2 with $h = 1$ to obtain

$$|f(z')/f(z)| \leq 1/(h_0 - 1) < 1/C.$$

This proves Corollary A.

COROLLARY B. *Let K and ρ be as before with $|f(z)| < \rho/(2K + 1)$. Then Newton's method starting at z converges to z^* , a solution of $f(z^*) = 0$. (Recall we could take $K = 4$.)*

PROOF. Choose $C > 1$ such that $|f(z)| < \rho/(CK + K + 1)$. Use Corollary A. If $z_1 = z - f(z)/f'(z)$, then $|f(z_1)/f(z)| < 1/C$. Inductively let $z_{n+1} = z_n - f(z_n)/f'(z_n)$. Thus $|f(z_{n+1})| < (1/C)|f(z_n)| < \rho/(2K + 1)$, so $f'(z_n)$ is never zero and $|f(z_{n+1})/f(z)| < (1/C)^{n+1}$. This proves Corollary B.

Note that Corollaries A and B give a new criterion for Newton's method to work. This is used strongly in our main result.

3. Here we make estimates in the target space or on the values $f(z_0), f(z_1), f(z_2), \dots$ as the algorithm constructs z_1, z_2, \dots . Under suitable initial conditions, it must be shown that $f(z_n)$ tends to zero sufficiently faster than it veers off toward some critical value $f(\theta)$. At the critical points, the algorithm is not defined, and near them goes very slowly. Thus this analysis is central. We use heavily Theorem 2 of the preceding section.

Throughout K will be the constant in the remark after Theorem 1. Certainly $K \geq 1$ and K could be taken as 4. As in §2, ρ_f is the number

$$\rho_f = \min_{\substack{\theta \\ f'(\theta)=0}} |f(\theta)|.$$

We now define for each polynomial f and $z_0 \in \mathbb{C}$ such that $f(z_0) \neq 0$ and $\rho_f > 0$,

$$\mathfrak{H} = \mathfrak{H}_{f,z_0} = \min_{\substack{\theta \\ f'(\theta)=0 \\ |f(\theta)| < 2|f(z_0)|}} \left\{ 1, \left| \arg \frac{f(\theta)}{f(z_0)} \right| \right\}.$$

Moreover, under the same conditions, define

$$\zeta = \zeta_{f,z_0} = \max \left\{ 2, \frac{(3K + 1)|f(z_0)|}{\rho_f} \right\}.$$

THEOREM 3. *There exist two universal functions $H: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{R}^+$, $S: \mathbb{R}^+ \times \mathbb{R}^+ \rightarrow \mathbb{Z}^+$, exhibited later, with these properties: If f, z_0 satisfy $\rho_f > 0$, $\mathfrak{H} = \mathfrak{H}_{f,z_0} > 0$ and $h = H(\mathfrak{H}, \zeta)$, $s = S(\mathfrak{H}, \zeta)$ then $z_k = z_{k-1} - hf(z_{k-1})/f'(z_{k-1})$ is well defined inductively for all k starting from $k = 1$ and*

$|f(z_s)| < \rho_f/(3K + 1)$. Moreover H and S have the form

$$H(\mathfrak{C}, \zeta) = \frac{1}{K} \frac{\sin^2(\mathfrak{C}/2)}{(3 \sin(\mathfrak{C}/2) + \log \zeta)}$$

$$S(\mathfrak{C}, \zeta) = K \left(3 + \frac{\log \zeta}{\sin(\mathfrak{C}/2)} \right)^2.$$

REMARK. According to Corollary A of Theorem 2, for $k > s$, one can change h to 1 (Newton's method) so that $z_k \rightarrow z^*$ as $k \rightarrow \infty$ with $f(z^*) = 0$. After s steps, we have found an approximate solution.

Define $h_* = (1/K) \sin(\mathfrak{C}/2)$. The following is the main lemma.

LEMMA 1. Suppose $s \in \mathbb{Z}^+$, $0 < h < h_*/2$, depending on ζ , \mathfrak{C} of Theorem 3 satisfy (1) and (2).

$$(1 - h + h^2/(h_* - h))^s < 1/\zeta, \quad (1)$$

$$s \left(\frac{h^2}{h_* - h} \right) \left(\frac{1}{1 - h} \right) < \sin \frac{\mathfrak{C}}{2}. \quad (2)$$

Then the conclusions of Theorem 3 (except the last sentence) are satisfied with these choices of h and s .

Note that once Lemma 1 has been proven, the rest of the proof of Theorem 3 amounts to solving equations (1) and (2) for h , s and exhibiting their dependence on \mathfrak{C} and ζ .

For the proof of Lemma 1, we use

LEMMA 2. (a) For any $\tau, \beta \in \mathbb{C}$, $\beta \neq 0$, $|\sin \arg((\beta + \tau)/\beta)| \leq |\tau/\beta|$,
 (b) $|\sin \arg(f(\theta)/f(z))| \leq |(f(z) - f(\theta))/f(z)|$ if $f(z) \neq 0$,
 (c) for $0 < h < h_*/2$, $\alpha \in \mathbb{C}$, $|\alpha| \leq 1$,

$$\left| \sin \arg \left(1 - h + \frac{\alpha h^2}{h_* - h} \right) \right| \leq \left(\frac{h^2}{h_* - h} \right) \left(\frac{1}{1 - h} \right).$$

PROOF OF LEMMA 2. (a) can be checked via trigonometry. (b) follows from (a) with $\tau = f(\theta) - f(z)$, $\beta = f(z)$. For (c), take in (a), $\beta = 1 - h$, $\tau = \alpha h^2/(h_* - h)$.

LEMMA FROM TRIGONOMETRY.

$$|\sin \sum \alpha_i| \leq \sum |\sin \alpha_i|.$$

Let z_0 be given as in Theorem 3. We proceed by induction to prove Lemma 1. Let $k \leq s$ and suppose inductively for $i < k$ that $f'(z_i) \neq 0$, and hence $z_{i+1} = z_i - hf(z_i)/f'(z_i)$ is well defined, and moreover

$$h_i = \frac{1}{K} \min_{f'(\theta)=0} \left| \frac{f(z_i) - f(\theta)}{f(z_i)} \right| > h_*.$$

Note that if ever $f(z_k)$ (or $f(z_i)$, $i < k$) is zero, we are finished trivially. To

finish the induction, we will show that $f'(z_k) \neq 0$ and $h_k > h_*$.

LEMMA 3.

$$\left| \frac{f(z_k)}{f(z_0)} \right| < \left(1 - h + \frac{h^2}{h_* - h} \right)^k < 1, \quad (\text{a})$$

$$\left| \arg \frac{f(z_k)}{f(z_0)} \right| < \frac{\mathfrak{C}}{2}. \quad (\text{b})$$

PROOF OF LEMMA 3. First write $f(z_k)/f(z_0) = f(z_k)/f(z_{k-1}) \cdots f(z_1)/f(z_0)$. So we may apply Theorem 2 to obtain

$$\frac{f(z_i)}{f(z_{i-1})} = 1 - h + \frac{\alpha_i h^2}{h_{i-1} - h}, \quad i = 1, \dots, k, |\alpha_i| \leq 1, 0 < h < h_{i-1}.$$

By the induction hypothesis $h_{i-1} > h_*$, $i = 1, \dots, k$, and thus since $h < h_*/2$, we have

$$\left| 1 - h + \frac{\alpha_i h^2}{h_{i-1} - h} \right| < \left| 1 - h + \frac{h^2}{h_* - h} \right|.$$

Therefore

$$\left| \frac{f(z_k)}{f(z_0)} \right| \leq \left| 1 - h + \frac{h^2}{h_* - h} \right|^k.$$

This proves (a) of Lemma 3.

For part (b) we have

$$\arg \frac{f(z_k)}{f(z_0)} = \sum_{i=1}^k \arg \left(\frac{f(z_i)}{f(z_{i-1})} \right)$$

and thus by the lemma from trigonometry,

$$\left| \sin \arg \frac{f(z_k)}{f(z_0)} \right| \leq \sum_{i=1}^k \left| \sin \arg \frac{f(z_i)}{f(z_{i-1})} \right|.$$

But by Lemma 2(c) we may estimate

$$\left| \sin \arg \frac{f(z_i)}{f(z_{i-1})} \right| \leq \left(\frac{h^2}{h_{i-1} - h} \right) \left(\frac{1}{1 - h} \right) \leq \left(\frac{h^2}{h_* - h} \right) \left(\frac{1}{1 - h} \right),$$

thus

$$\left| \sin \arg \frac{f(z_k)}{f(z_0)} \right| \leq k \left(\frac{h^2}{h_* - h} \right) \left(\frac{1}{1 - h} \right).$$

By (2), since $k \leq s$,

$$\left| \sin \arg \frac{f(z_k)}{f(z_0)} \right| < \sin \frac{\mathfrak{C}}{2} < \frac{\mathfrak{C}}{2},$$

the last since $|\arg(f(z_k)/f(z_0))| \leq \mathfrak{C} \leq 1$. This proves Lemma 3.

Now $|\arg(f(\theta)/f(z_0))| \geq \mathfrak{C}$ for any critical point $\theta \in \mathbf{C}$ such that $|f(\theta)| < 2|f(z_0)|$. Since

$$\arg \frac{f(\theta)}{f(z_0)} = \arg \frac{f(\theta)}{f(z_k)} + \arg \frac{f(z_k)}{f(z_0)},$$

we conclude from Lemma 3 that $|\arg\{f(\theta)/f(z_k)\}| > \mathfrak{C}/2$ for any critical point θ with $|f(\theta)| \leq 2|f(z_0)|$. In particular, $|f(z_k)| \leq 2|f(z_0)|$ (Lemma 3(a)) so that z_k is not a critical point, $f'(z_k) \neq 0$.

Next, $|\sin \arg(f(\theta)/f(z_k))| > \sin(\mathfrak{C}/2)$, so using Lemma 2(b)

$$\left| \frac{f(z_k) - f(\theta)}{f(z_k)} \right| \geq \left| \sin \arg \frac{f(\theta)}{f(z_k)} \right| > \sin \frac{\mathfrak{C}}{2}$$

for any critical point θ such that $|f(\theta)| < 2|f(z_0)|$. On the other hand, if $|f(\theta)| > 2|f(z_0)|$, since $|f(z_k)| < |f(z_0)|$,

$$\left| \frac{f(z_k) - f(\theta)}{f(z_k)} \right| > 1 > \sin \frac{\mathfrak{C}}{2}.$$

Therefore

$$h_k = \min_{\substack{\theta \\ f'(\theta)=0}} \frac{1}{K} \left| \frac{f(z_k) - f(\theta)}{f(z_k)} \right| > \frac{1}{K} \sin \frac{\mathfrak{C}}{2} = h_*.$$

This finishes the induction. To finish the proof of Lemma 1, one has to show that $|f(z_s)| < \rho_f/(3K+1)$. Use Lemma 3(a) to get $|f(z_s)| < (1-h+h^2/(h_*-h))^s |f(z_0)|$. So $|f(z_s)| < (1/\xi)|f(z_0)|$ by (1) which in turn is less than $\rho_f/(3K+1)$ by the definition of ξ . This proves Lemma 1.

Now we solve (1) and (2) for h and s to finish the proof of Theorem 3. Recall

$$\left(1 - h + \frac{h^2}{h_* - h}\right)^s < \frac{1}{\xi}. \quad (1)$$

$$s \left(\frac{h^2}{h_* - h} \right) \left(\frac{1}{1-h} \right) < \sin \frac{\mathfrak{C}}{2} = Kh_*. \quad (2)$$

(1) is equivalent to

$$\frac{\log \xi}{\log(1 - h + h^2/(h_* - h))^{-1}} < s \quad (1')$$

and (2) to

$$s < Kh_* \left[\frac{(1-h)(h_*-h)}{h^2} \right]. \quad (2')$$

LEMMA 4 (EASY). $\log(1/(1-y)) > y$ if $0 < y < 1$.

In (1') let $y = h - h^2/(h_* - h) = h((h_* - 2h)/(h_* - h))$ and use Lemma

4 to see that (1'') implies (1').

$$\frac{\log \zeta}{y} < s. \quad (1'')$$

Next use $h = h_*/C$ where $C > 2$ to replace (1'') and (2') respectively by

$$(\log \zeta) \frac{C}{h_*} \left(\frac{C-1}{C-2} \right) < s, \quad (1''')$$

$$s < K(C - h_*)(C - 1). \quad (2'')$$

Besides (1''') and (2''), we have the condition that s be an integer.

TECHNICAL LEMMA. *If $C \geq 3 + \log \zeta / Kh_*$, then*

$$(\log \zeta) \frac{C}{h_*} \left(\frac{C-1}{C-2} \right) + 1 < K(C - h_*)(C - 1).$$

PROOF OF TECHNICAL LEMMA. Note first that if $C > 3$ then

$$\frac{C-2}{K} \frac{1}{C} \left(\frac{1}{C-1} + Kh_* \right) \leq 1$$

(using $h_* < \frac{1}{2}$, since $\mathfrak{C} \leq 1$).

Therefore from the hypothesis of the Technical Lemma,

$$\frac{\log \zeta}{Kh_*} + \frac{(C-2)}{K} \left(\frac{1}{C} \right) \left(\frac{1}{C-1} + Kh_* \right) \leq C - 2.$$

Next divide and multiply through this expression to obtain the Technical Lemma.

To finish we choose $C = 3 + \log \zeta / Kh_*$ and $h = h_*/C$. Then choose s to be the greatest integer in $(\log \zeta)(C/h_*)((C-1)/(C-2)) + 1$.

In fact we may calculate h and s more explicitly as follows.

$$h = Kh_*^2 / (\log \zeta + 3h_*K), \quad h_* = (1/K) \sin^2(\mathfrak{C}/2)$$

giving $H(\mathfrak{C}, \zeta)$ as in Theorem 3.

Going back to (2'') we could take s as simply KC^2 since increased s (fixed h) will still satisfy our conditions in Theorem 3. Then

$$s = \left(\frac{(1/K) \sin(\mathfrak{C}/2)}{h} \right)^2 = \frac{K(3 \sin(\mathfrak{C}/2) + \log \zeta)^2}{\sin^2(\mathfrak{C}/2)}.$$

This finishes the proof of Theorem 3.

4. This section gives the main probability (or volume) estimates which are used in our statistics. Our approach is more geometric than measure-theoretic although we do make use of Fubini's theorem. There is an interesting historical point about a theorem of *Herman Weyl* on the volume of tubes that we use. The statistician (and economist) *Harold Hotelling* proved a result on the volume of a tubular neighborhood of curves (see Lemma 2 below). This is motivated by problems of regression equations in theoretical statistics. Subsequently *Hotelling's* theorem was extended by *Weyl* (see Lemma 5 below) and

these two papers lie next to each other in the American Journal of 1939. *Weyl* writes: "In a lecture before the Mathematics Club at Princeton last year Professor Hotelling stated the following geometric problem as one of primary important for certain statistical investigations: . . ." and *Weyl* goes on to state and prove his theorem on the volume of tubes.

In the intervening years *Weyl's* result has become well known to geometers; see *Griffith's* extensive discussion. Now the result of *Weyl* is found useful in our statistical analysis.

There is another amusing aside in this connection. In the use of *Weyl's* theorem, I had need to estimate the integral of the Gaussian curvature K of a complex algebraic curve γ in \mathbf{C}^2 in terms of the degree δ . I took this problem to *Osserman* and after a little thought, using his joint paper with *Hoffman* he came up with the neat estimate:

$$\frac{1}{2\pi} \int K \leq -\delta(\delta - 1).$$

This is nice because the volume of γ is infinite, and it might also have singularities where the curvature blows up. Although I believe this estimate is not explicit in the literature, it now follows easily from standard results. *Griffiths* subsequently showed me another proof.

Eventually, I bypassed the need for the estimate by noticing a simple sign argument which did the trick.

We proceed to the theorems on volume estimates. Let \mathcal{P}_d be the space of polynomials of degree d with leading coefficient 1. Thus \mathcal{P}_d can be identified with complex Cartesian space $\mathbf{C}^d = \{(a_0, \dots, a_{d-1}) = a | a_i \in \mathbf{C}\}$ or we may write $f \in \mathbf{C}^d$ where $f(z) = \sum_{i=0}^d a_i z^i$, $a_d = 1$. Let $P(R)$ denote the polycylinder $\{a \in \mathcal{P}_d | |a_i| < R, i = 0, \dots, d-1\}$. Note that the volume of $P(R)$ or $\text{vol } P(R)$ is $(\pi R^2)^d$. Here we use the standard volume on $\mathbf{C}^d = \mathbf{R}^{2d}$ for \mathcal{P}_d .

Let $W_0 = \{f \in \mathcal{P}_d | f(\theta) = 0 \text{ for some } \theta \text{ with } f'(\theta) = 0\}$. It is easily seen that W_0 is the set of f with multiple roots. Finally define $U_\rho(W_0) = \cup_{f \in W_0} U_\rho(f_0)$ where $U_\rho(f_0) = \{f \in \mathcal{P}_d | |f(\theta) - f_0(\theta)| < \rho \text{ and } f'(z) = f_0'(z) \text{ for all } z\}$. Then $f \in U_\rho(f_0)$ if and only if $|f_0(0) - f(0)| < \rho$ and the coefficients a_i in f and f_0 , $i > 0$, coincide.

THEOREM 4A.

$$\frac{\text{Vol}[U_\rho(W_0) \cap P(R)]}{\text{Vol } P(R)} \leq \frac{d\rho^2}{R^2}.$$

This theorem is not used in the sequel; however it serves as a simple prototype of Theorems 4B and 4C which are important for our development.

The proof of Theorem 4A goes via

PROPOSITION 1. *The subset $W_0 \subset \mathcal{P}_d \cong \mathbf{C}^d$ has a representation as a complex algebraic hypersurface given by a single polynomial equation of degree $2d - 1$ which is of degree d in a_0 . More precisely, there is a polynomial $F: \mathbf{C}^d \rightarrow \mathbf{C}$,*

$$F(a_0, \dots, a_{d-1}) = \sum_{i=0}^d F_i(a_1, \dots, a_{d-1}) a_0^i,$$

where the F_i are polynomials, F has total degree $2d - 1$ and

$$W_0 = \{a \in \mathbf{C}^d | F(a) = 0\}.$$

PROOF OF PROPOSITION 1. Note that

$$W_0 = \{f \in \mathcal{P}_d | f(\theta) = 0, f'(\theta) = 0 \text{ for some } \theta \in \mathbf{C}\}.$$

Then the proof goes by the description of the resultant (see *Van der Waerden*, Vol. I (p. 84) or *Lang*). Thus F is the resultant of f and f' , $R(f, f')$. The proof of the proposition then follows from an examination of the form of $R(f, f')$ given in either reference.

For the proof of Theorem 4A, define $\chi: \mathbf{C}^d \rightarrow \mathbf{R}$ to be the characteristic function of $U_\rho(W_0)$, i.e., χ is 1 on $U_\rho(W_0) \subset \mathbf{C}^d$ and zero off of it.

Now observe that for almost every (in the sense of Lebesgue measure) (a_1, \dots, a_{d-1}) , W_0 intersected with the one (complex) dimensional coordinate plane through (a_1, \dots, a_{d-1}) consists of at most d points. This follows from the proposition, with the exceptional (a_1, \dots, a_{d-1}) being those points where all the F_i vanish. From this, using the definitions, we may conclude that for almost every (a_1, \dots, a_{d-1}) ,

$$\left| \int_{|a_0| < R} \chi(a_0, \dots, a_{d-1}) da_0 \right| \leq d\pi\rho^2.$$

Thus by Fubini's theorem

$$\begin{aligned} \frac{\text{Vol } U_\rho(W_0) \cap P(R)}{\text{Vol } P(R)} &= \frac{1}{(\pi R^2)^d} \int_{P(R)} \chi(a) da \\ &= \frac{1}{(\pi R^2)^d} \int_{|a_1| < R} \left[\int_{|a_0| < R} \chi(a_0, \dots, a_{d-1}) da_0 \right] da_1 \cdots da_{d-1} \\ &\quad \vdots \\ &\quad |a_{d-1}| < R \\ &\leq \frac{1}{(\pi R^2)^d} \int_{|a_1| < R, \dots, |a_{d-1}| < R} (d\pi\rho^2) da_1 \cdots da_{d-1} \\ &\leq \frac{1}{(\pi R^2)^d} (d\pi\rho^2)(\pi R^2)^{d-1} \leq \frac{d\rho^2}{R^2}. \end{aligned}$$

This proves Theorem 4A.

We will now use $\text{Re } z$ and $\text{Im } z$ systematically for the real and imaginary parts of a complex number z .

Consider the subspace W_* of \mathcal{P}_d , $W_* = \{f \in \mathcal{P}_d | \text{Im}(\overline{f(0)}f(\theta)) = 0 \text{ for some critical point } \theta \text{ of } f\}$. So $f \in W_*$ provided $\text{Re } f'$, $\text{Im } f'$ and $\text{Im}(\overline{f(0)}f)$ have a common zero. Let $U_\rho(W_*) = \bigcup_{f_0 \in W_*} U_\rho(f_0)$, where $U_\rho(f_0)$ is defined earlier in this section.

THEOREM 4B.

$$\frac{\text{Vol}[U_\rho(W_*) \cap P(R)]}{\text{Vol } P(R)} \leq \frac{3\rho(d-1)^2}{R}.$$

The proof of this theorem uses the following Algebraic Lemma.

ALGEBRAIC LEMMA. *The set $W_* \subset \mathbb{P}_d$ is a closed subset of a real algebraic variety in $\mathbf{R}^{2d} = \mathbf{C}^d$. In fact this variety is the set of zeros of a real polynomial G in the $2d$ variables $(\text{Re } a_i, \text{Im } a_i)_{i=0}^{d-1}$ and G has degree $(d-1)^2$ in $(\text{Re } a_0, \text{Im } a_0)$.*

To prove the Algebraic Lemma, we use elimination theory which can be found in *Van der Waerden*, Vol. II. See especially p. 15.

THEOREM FROM ELIMINATION THEORY. *n generic homogeneous polynomials in n -variables have a resultant F which is an integral polynomial in their (indeterminate) coefficients. The vanishing of this resultant for particular f_1, \dots, f_n with coefficients in a field is a necessary and sufficient condition for the existence of a solution (perhaps in an extension field) of the system of equations $f_1 = 0, \dots, f_n = 0$, distinct from the zero solution. The resultant is homogeneous in the coefficients of f_k of degree $D_k = \prod_{i \neq k} d_i$ (the product) where degree $f_i = d_i$.*

We need to apply this result to the case of n inhomogeneous polynomials in $(n-1)$ variables. This is done as usual by adding a new variable to homogenize the polynomials. New solutions may be introduced, however. We obtain in this way

COROLLARY. *Let $\mathcal{F}_i, i = 1, \dots, n$, be given spaces of inhomogeneous polynomials f_i of degree d_i in $(n-1)$ variables over a fixed field K . Thus \mathcal{F}_i can be identified with the space of coefficients of f_i . Then there is a polynomial (the resultant) $F: \prod_{i=1}^n \mathcal{F}_i \rightarrow K$ in the coefficients of these f_i with the property that if particular f_1, \dots, f_n have a common zero, then $F(f_1, \dots, f_n) = 0$. Furthermore F is homogeneous in the coefficients of f of degree $D_k = \prod_{i=1; i \neq k}^n d_i$.*

Now fix the field to be \mathbf{R} , the real number field. Let \mathcal{F}_1 and \mathcal{F}_2 each be the space of real (inhomogeneous) polynomials of 2 variables of degree $d-1$. Let \mathcal{F}_3 be the space of real polynomials of 2 variables of degree d .

The above corollary applies to yield a polynomial $F: \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3 \rightarrow \mathbf{R}$ such that $F(g_1, g_2, g_3) = 0$ whenever g_1, g_2, g_3 have a common zero $(x, y) \in \mathbf{R}^2$.

Next define a map $\Phi: \mathcal{P}_d \rightarrow \mathcal{F}_1 \times \mathcal{F}_2 \times \mathcal{F}_3$, $\Phi = (\Phi_1, \Phi_2, \Phi_3)$ by $\Phi(f) = (\text{Re } f', \text{Im } f', \text{Im}(\overline{f(0)}f))$. Here f' , the derivative of f , may be naturally interpreted as the sum of 2 polynomials, $f' = \text{Re } f' + i \text{Im } f'$, where $\text{Re } f'$ is a real polynomial of 2 real variables ($\mathbf{C} = \mathbf{R}^2$) of degree $d-1$. Similarly $\text{Im}(\overline{f(0)}f)$ has a natural interpretation as a real polynomial of 2 real variables of degree d . Explicitly $\text{Im}(\overline{f(0)}f)$ is the imaginary part of the polynomial

$$\sum_{k=0}^d (\overline{\alpha_0 + i\beta_0})(\alpha_k + i\beta_k)(x + iy)^k$$

where $a_k = \alpha_k + i\beta_k$, $k = 0, \dots, d$ and $i = \sqrt{-1}$ (recall $a_d = 1$).

Thus Φ is well defined and moreover $\Phi_1 f$, $\Phi_2 f$ are independent of α_0, β_0 and Φ_3 is linear in α_0, β_0 .

From the construction of $F: \mathfrak{F}_1 \times \mathfrak{F}_2 \times \mathfrak{F}_3 \rightarrow \mathbf{R}$, if $f \in W_*$, then $F \circ \Phi(f) = 0$. Let $G = F \circ \Phi$. Then G is a polynomial in the $\alpha_0, \beta_0, \dots, \alpha_{d-1}, \beta_{d-1}$ which has degree $(d - 1)^2$ in (α_0, β_0) . This finishes the proof of the Algebraic Lemma.

Now let γ be a plane real algebraic curve of degree δ . Thus γ can be thought of as a locus of zeros in \mathbf{R}^2 of a real polynomial $F(x, y)$ of degree δ ($\delta = (d - 1)^2$ in our application). The *singular points* of γ are those points in \mathbf{R}^2 which are common zeros of $\partial F/\partial x$ and $\partial F/\partial y$. Thus by Bezout's theorem there are at most $(\delta - 1)^2$ singular points.

Let $N_\rho(\gamma)$ be the set of all points in \mathbf{R}^2 within ρ of some point of γ . Let $D_R \subset \mathbf{R}^2$ be the open disk of radius $R > 0$ about zero.

PROPOSITION 2. *If γ is a real plane algebraic curve of degree δ , then*

$$\text{area}[N_\rho(\gamma) \cap D_R] \leq 3\pi R\delta\rho.$$

For the proof introduce the set $T_\rho(\gamma) \subset \mathbf{R}^2$ of all points which lie on line segments of length $< \rho$ normal to a nonsingular point of γ .

LEMMA 1.

$$N_\rho(\gamma) \cap D_R \subset V_\rho \cup T_\rho(\gamma \cap D_{R+\rho})$$

where

$$V_\rho = \cup D_\rho(z), \quad z \text{ a singular point of } \gamma.$$

Here $D_\rho(z)$ is the disk of radius ρ about z .

For the proof of Lemma 1, let $x \in N_\rho(\gamma)$, $x \in D_R$. Then there is $y \in \gamma$ such that $d(x, y) < \rho$ and we may suppose $d(x, y)$ is minimal. Therefore $y \in D_{R+\rho}$. If y is a singular point, then $x \in D_\rho(y) \subset V_\rho$. If y is not singular, then since $d(x, y)$ is minimal, x lies on a normal segment to γ at y of length less than ρ , i.e., $x \in T_\rho(\gamma \cap D_{R+\rho})$. This proves Lemma 1.

The following is a special case of a theorem of *Hotelling*.

LEMMA 2. *Let γ_0 be a smooth plane curve with nonvanishing tangent (but not necessarily connected) of total length l . Then*

$$\text{Area } T_\rho(\gamma_0) \leq 2\rho l.$$

LEMMA 3. *Let γ be a real algebraic plane curve of degree δ . Then the length $l(\gamma \cap D_R)$ is given by*

$$l(\gamma \cap D_R) \leq \pi\delta R.$$

For the proof of Lemma 3, we follow *Santalo* pp. 17–31 and especially pp. 30–31. Let dG be the standard measure on the space of lines in the plane. The

length L of a curve C satisfies (Santalo p. 31)

$$2L = \int n \, dG$$

where n is the number of points of intersection of a line and the curve. The integral is over all lines. Now in our case, almost every line (in the sense of our measure on the space of lines in \mathbf{R}^2) meets C in at most δ points. This is a very special case of Bezout's theorem. Moreover, the integral is just over the lines which meet D_R . So

$$2L \leq \delta \int_{\text{lines meeting } D_R} dG \leq \delta 2\pi R.$$

The last evaluation, that

$$\int_{\text{lines meeting } D_R} dG = 2\pi R,$$

is shown on Santalo p. 30. This proves Lemma 3.

We now give the proof of Proposition 2. Use Lemma 1 to obtain

$$\text{Area}(N_\rho(\gamma) \cap D_R) \leq \text{Area } V_\rho + \text{Area } T_\rho(\gamma \cap D_{R+\rho}).$$

Clearly, $\text{Area } V_\rho \leq \pi\rho^2(\delta - 1)^2$ since there are at most $(\delta - 1)^2$ singularities (again we use Bezout's theorem). By Lemma 2, $\text{Area } T_\rho(\gamma \cap D_{R+\rho}) \leq 2\rho l(\gamma \cap D_{R+\rho})$ and by Lemma 3, this is less than $2\rho\pi\delta(R + \delta)$. Putting these estimates together yields

$$\begin{aligned} \text{Area}(N_\rho(\gamma) \cap D_R) &\leq \pi\rho^2(\delta - 1)^2 + 2\rho(\delta\pi(R + \rho)) \\ &\leq \pi[2\rho\delta R + \rho^2[2\delta + (\delta - 1)^2]] \\ &\leq \pi[2\rho\delta R + \rho^2(\delta + 1)]. \end{aligned}$$

Now observe that if $\rho \geq R/3\delta$, then $3\pi\delta R\rho \geq \pi R^2$ and since

$$\text{area}[N_\rho(\gamma) \cap D_R] \leq \text{area } D_R \leq \pi R^2,$$

the proposition is true in this case.

Thus for the proof we may assume $\rho < R/3\delta$. Since $(R\rho/3\delta)(\delta^2 + 1) \leq \delta R\rho$ (easily checked), $\rho^2(\delta^2 + 1) \leq \delta R\rho$. Together with our previous estimate this yields the proposition.

Now the proof of Theorem 4B proceeds as the proof of Theorem 1A. Let $\chi: \mathbf{C}^d \cong \mathbf{R}^{2d} \rightarrow \mathbf{R}$ be the characteristic function of $U_\rho(W_*)$. Now represent W_* as in the algebraic lemma. The almost everywhere considerations used in the proof of Theorem 4A apply here as well. Thus by Proposition 2, for almost every (a_1, \dots, a_{d-1}) ,

$$\left| \int_{|a_0| < R} \chi(a_0, a_1, \dots, a_{d-1}) \, da_0 \right| \leq 3(d-1)^2 \pi \rho R.$$

Then by Fubini's theorem,

$$\begin{aligned} \frac{\text{Vol } U_\rho(W_*) \cap P(R)}{\text{Vol } P(R)} &= \frac{1}{(\pi R^2)^d} \int_{\substack{|a_1| < R \\ \vdots \\ |a_{d-1}| < R}} \int \chi(a_0, \dots, a_{d-1}) da_0 da_1 \dots da_{d-1} \\ &< \frac{1}{(\pi R^2)^d} 3(d-1)^2 \pi \rho R \int_{|a_1| < R, \dots, |a_{d-1}| < R} da_1 \dots da_{d-1} \\ &< \frac{3(d-1)^2}{(\pi R^2)^d} \pi \rho R (\pi R^2)^{d-1} = 3(d-1)^2 \frac{\rho}{R}. \end{aligned}$$

This proves Theorem 4B.

Define

$$W_1 = \{f \in \mathcal{P}_d \mid f(\theta) = f(0) \text{ for some critical point } \theta\}.$$

Let

$$L_\rho(f_0) = \left\{ f \in \mathcal{P}_d \mid |f'(0) - f'_0(0)| < \rho \text{ and } \left| \frac{f''(0)}{2} - \frac{f''_0(0)}{2} \right| < \rho \right\}$$

and

$$L_\rho(W_1) = \cup L_\rho(f_0), \quad f_0 \in W_1.$$

THEOREM 4C.

$$\frac{\text{Vol}[L_\rho(W_1) \cap P(R)]}{\text{Vol } P(R)} < 4 \left(\frac{\rho}{R} \right)^2 (d+2).$$

As in the proof of Theorem 4A we use the resultant to describe W_1 as a variety in \mathbb{C}^d . The basic fact is that $f \in W_1$ if and only if f' and $f - f(0)$ have a common zero. As in *van der Waerden*, Vol. I, and *Lang*, we consider the resultant $R(f', f - f(0))$ as a polynomial F in the coefficients $(a_0, a_1, \dots, a_{d-1})$. Note here that of course $a_d = 1$, and also that F is independent of a_0 . Furthermore it can be seen from the form of the resultant that F has degree $(d+2)$ in (a_1, a_2) .

On \mathbb{C}^2 we will use the norm $\| \cdot \|_B$, defined by $\|z\|_B = \|(z_1, z_2)\|_B = \max(|z_1|, |z_2|)$.

PROPOSITION 3. Let γ be a complex algebraic plane curve of degree $\delta > 3$ and define

$$N_\rho(\gamma) = \{z \in \mathbb{C}^2 \mid \|z - y\|_B < \rho \text{ for some } y \in \gamma\}.$$

Then

$$\text{Vol}[N_\rho(\gamma) \cap P(R)] < 4\delta(\pi\rho R)^2.$$

Here $P(R)$ is the polycylinder $D_R \times D_R$, D_R the disk of radius R in \mathbf{C} .

The proof of Proposition 3 is analogous to the proof of Proposition 2 with Lemmas 4, 5, 6 playing the role of Lemmas 1, 2, 3 respectively. Also Proposition 3 plays the same role in the proof of Theorem 4C that Proposition 2 did in the proof of Theorem 4B.

Define $T_\rho(\gamma)$ just as in Lemma 1, in terms of distance along line segments normal to the surface γ in \mathbf{R}^4 . The same applies to $T_\rho(\gamma \cap P(R + \rho))$. Define

$$V_\rho = \bigcup_{z \in \Sigma} P_z(\rho)$$

where Σ is the set of singular points of γ and $P_z(\rho)$ is the polycylinder

$$P_z(\rho) = \{y \in \mathbf{C}^2 \mid \|y - z\|_B < \rho\}.$$

LEMMA 4.

$$N_\rho(\gamma) \cap P(R) \subset V_\rho \cup T_\rho(\gamma \cap P(R + \rho)).$$

REMARK. It seems likely that this is true omitting V_ρ . However, this doesn't affect our estimates much and we leave it in. P. Griffiths, in conversation, has confirmed the possibility of omitting V_ρ .

PROOF OF LEMMA 4. Let $x \in P(R) \cap N_\rho(\gamma)$. Suppose $y \in \gamma$ minimizes $\|x - y\|_B$. Thus $\|x - y\|_B < \rho$ and if $x \notin V$, then y is nonsingular. Therefore the segment xy is normal to γ . Finally $\|x\|_B < R$, $\|x - y\|_B < \rho$ so $\|y\|_B < R + \rho$ and $y \in P(R + \rho)$. This proves Lemma 4.

See *Weyl* (1939) (and also *Griffiths*) for the following

LEMMA 5 (WEYL).

$$\text{Vol } T_\rho(\gamma) \leq \pi \left[\frac{\rho^2}{4} \text{ area } \gamma + \frac{\rho^4}{24} \int_\gamma K \right].$$

Here γ is supposed to be a smooth surface in \mathbf{R}^2 with equality in the case of no overlapping. Actually *Weyl* supposes γ is compact, no boundary (or singularities). But his analysis applies to our situation as long as we interpret T_ρ via normal segments. In the integral K is the Gaussian curvature of γ .

We first note that since γ is a complex algebraic curve, it is a minimal surface; therefore everywhere $K \leq 0$ and we may discard this term. See also the introduction to this section.

LEMMA 6. If γ is a complex plane algebraic curve of degree δ , then

$$\text{area}[\gamma \cap P(R)] \leq 6\pi\delta R^2.$$

PROOF OF LEMMA 6. Here we follow *Santalo* once more. We integrate functions over the space of 2 dimensional planes in $\mathbf{R}^4 = \mathbf{C}^2$.

This goes as follows. We write L_2 for a typical 2-plane in \mathbf{R}^4 and dL_2 for the measure over the space of such planes. Then

$$\frac{O_4 O_3}{O_2 O_1} \text{ area}(\gamma \cap P(R)) \leq \int_{L_2} \#(\gamma \cap L_2) dL_2$$

where $\#(\gamma \cap L_2)$ is the number of points of intersection of γ and L_2 (see

Santalo, 14.70, p. 245). Here O_k is the surface area of the unit k -sphere and in particular $O_0 = 2$, $O_1 = 2\pi$, $O_2 = 4\pi$, $O_2/O_4 = 3/2\pi$ (*Santalo*, p. 9).

Now $\#(\gamma \cap L_2) \leq \delta$ by Bezout's theorem except for a set of L_2 of measure zero. Then

$$\int_{L_2} \#(\gamma \cap L_2) dL_2 \leq \delta \int_{L_2 \cap P(R) \neq \emptyset} dL_2.$$

Now for any convex set K (*Santalo*, p. 233), $\int_{L_2 \cap K} dL_2$ is evaluated in terms of a certain *quermassintegrals* $W_r(K)$. Thus

$$\int_{L_2 \cap K} dL_2 = 2 \frac{O_2 O_1}{O_1 O_0} W_2(K).$$

I was not able to use *Santalo's* formula (p. 225) for $K = P(R)$; so crudely use the fact that $P(R) \subset D_{\sqrt{2}R}$, the ball, and *Santalo's* evaluation on p. 224

$$W_2(D_{\sqrt{2}R}) = \frac{O_3 R^2}{2}.$$

Thus

$$\int_{L_2 \cap K} dL_2 \leq \frac{O_3 O_2}{O_0} R^2.$$

Putting these estimates together we obtain

$$\text{area}(\gamma \cap P(R)) \leq \frac{O_2 O_1}{O_4 O_3} \delta \frac{O_3}{O_0} O_2 R^2 \leq 6\pi\delta R^2.$$

This proves Lemma 6.

We now are able to prove Proposition 3 in much the manner of Proposition 2. First use Lemma 4.

$$\text{Vol}(N_\rho(\gamma) \cap P(R)) \leq \text{Vol } V_\rho + \text{Vol } T_\rho(\gamma \cap P(R + \rho)),$$

where

$$\text{Vol } V_\rho \leq (\delta - 1)^2 \text{Vol } P(\rho) = (\delta - 1)^2 (\pi\rho^2)^2.$$

By Lemma 5 (and comments),

$$\text{Vol } T_\rho(\gamma \cap P(R + \rho)) \leq \pi \frac{\rho^2}{4} \text{area}(\gamma \cap P(R + \rho));$$

then using Lemma 6,

$$\text{Vol } T_\rho(\gamma \cap P(R + \rho)) \leq \frac{\pi\rho^2}{4} \cdot 6\pi\delta(R + \rho)^2,$$

so

$$\text{Vol}[N_\rho(\gamma) \cap P(R)] \leq \pi^2 \rho^4 (\delta - 1)^2 + \frac{3}{2} \pi^2 \rho^2 \delta (R + \rho)^2. \quad (*)$$

Now we prove the estimate of Proposition 3 in two cases. First suppose that

$\rho > (2/3\delta)^{1/2}R$. Then

$$4\delta\pi^2R^2\rho^2 \geq (\pi R^2)^2 \geq \text{Vol } P(R) \geq \text{Vol}[N_\rho(\gamma) \cap P(R)].$$

On the other hand, suppose $\rho \leq (2/3\delta)^{1/2}R$. Use our previous estimate (*). So we have to show that if $\rho \leq (2/3\delta)^{1/2}R$, then

$$\pi^2\rho^4(\delta - 1)^2 + \frac{3}{2}\pi^2\rho^2\delta(R + \rho)^2 \leq 4\delta\pi^2\rho^2R^2$$

or

$$\rho^2(\delta - 1)^2 + \frac{3}{2}\delta(R + \rho)^2 \leq 4\delta R^2$$

or yet

$$\frac{2}{3\delta^2}(\delta - 1)^2 + \frac{3}{2}\left(1 + \left(\frac{2}{3\delta}\right)^{1/2}\right)^2 \leq 4.$$

But this last is easily checked using the fact that $\delta \geq 4$. This proves Proposition 3.

It remains to prove Theorem 4C, which will follow closely the proof of Theorem 4B. Thus let $\chi: \mathbf{C}^d \rightarrow \mathbf{R}$ be the characteristic function of $L_\rho(W_1)$.

Now we have shown that W_1 is a complex hypersurface in \mathbf{C}^d with the property that for almost every $(a_0, a_3, a_4, \dots, a_{d-1})$, it has degree $(d + 2)$ in a_1, a_2 . Furthermore

$$L_\rho(W_1) = \cup \{(a_0, a'_1, a'_2, a_3, \dots, a_{d-1}) | \\ \times |a'_1 - a_1| < \rho, |a'_2 - a_2| < \rho\}, \quad (a_0, a_1, \dots, a_{d-1}) \in W_1.$$

By Proposition 3, for almost every $(a_0, a_3, \dots, a_{d-1})$

$$\int_{\substack{|a_1| < R \\ |a_2| < R}} \chi(a_0, \dots, a_{d-1}) da_1 da_2 \leq 4(d + 2)(\pi\rho R)^2.$$

Then

$$\frac{\text{Vol } L_\rho(W_1) \cap P(R)}{\text{Vol } P(R)}$$

$$\leq \frac{1}{(\pi R^2)^d} \int_{\substack{|a_0| < R \\ |a_3| < R \\ \vdots \\ |a_{d-1}| < R}} \left[\int \chi(a_0, \dots, a_{d-1}) da_1 da_2 \right] da_0 \dots da_{d-1} \\ \leq \frac{1}{(\pi R^2)^d} \left[\int_{\substack{|a_0| < R \\ |a_3| < R \\ |a_{d-1}| < R}} da_0 da_3 \dots da_{d-1} \right] 4(d + 2)(\pi\rho R)^2$$

or

$$\begin{aligned} \frac{\text{Vol}(L_\rho(W_1) \cap P(R))}{\text{Vol } P(R)} &\leq \frac{(\pi R^2)^{d-2}}{(\pi R^2)^d} 4(d+2)\pi^2 R^2 \rho^2 \\ &\leq 4(d+2) \frac{\rho^2}{R^2}. \end{aligned}$$

This proves Theorem 4c.

5. Here we will prove our main result, as stated in the introduction and in Theorem 6. Toward this end we combine Theorem 3 and Theorems 4B, 4C. This is done by first using Theorems 4B and 4C to prove Theorem 5. Then the main result will follow easily from Theorems 3 and 5.

We take $U_{\sigma R}(W_*)$ as defined in §4 and define

$$\begin{aligned} Q_\sigma &= \{f \in \mathcal{P}_d \mid |f(\theta) - f(0)| < \sigma \text{ for some critical point } \theta\}, \\ Y_\sigma &= Q_\sigma \cup U_{\sigma R}(W_*). \end{aligned}$$

THEOREM 5. (1) If $R > \frac{1}{3}$,

$$\frac{\text{Vol } Y_\sigma \cap P(R)}{\text{Vol } P(R)} \leq 150(d+2)^{4/3} \sigma^{2/3},$$

(2) Let $\sigma < 1$ and $f \notin Y_\sigma$. Then

- (a) $\rho_f > \sigma R$,
- (b) $4R \sin \mathfrak{H}_f \geq \sigma^2$.

Here (see also §3)

$$\rho_f = \min_{f'(\theta)=0} |f(\theta)|, \quad \mathfrak{H} = \mathfrak{H}_f = \min_{f'(\theta)=0} \left| \arg \frac{f(\theta)}{f(0)} \right|.$$

For the proof of (1) we use Lemma 1.

LEMMA 1. If $\sigma < \frac{1}{8}$, $Q_\sigma \cap P(R) \subset L_\alpha(W_1)$,

$$\alpha = \max\{2\sigma^{2/3}, \sigma^{1/3}, 6R\sigma^{1/3}\}$$

where $L_\alpha(W_1)$ is defined in Theorem 4C.

PROOF. Let $f \in Q_\sigma \cap P(R)$, $f(z) = \sum_{i=0}^d a_i z^i$, $f(\theta) - f(0) = b$, $|b| < \sigma$, $f'(\theta) = 0$. We will construct $f_0 \in W_1$ such that $f \in L_\alpha(f_0)$. There are two cases to consider; first suppose that $|\theta| \geq \sigma^{1/3}$. Define in this case $f_0(z) = f(z) - c_1 z - c_2 z^2$ where $c_1 = 2b/\theta$, $c_2 = -b/\theta^2$. One can compute directly that $|c_1| < \alpha$, $|c_2| < \alpha$, so that it remains to show $f_0 \in W_1$. To do that it is sufficient to prove that $f'_0(\theta) = 0$ and $f_0(\theta) = f_0(0)$. But both of these are checked by direct calculations.

To finish the proof of Lemma 1, it remains to construct the desired f_0 in case $|\theta| < \sigma^{1/3}$. For that let $f_0(z) = f(z) - a_1 z$ and choose $\theta' = 0$. Then clearly $f'_0(\theta') = 0$ and $f_0(\theta') = f_0(0)$, so $f_0 \in W_1$. To show $f \in L_\alpha(f_0)$, it is sufficient to prove that $|a_1| < 6R\sigma^{1/3}$.

Since $f'(\theta) = 0$, $\sum_{k=0}^d k a_k \theta^{k-1} = 0$, so

$$|a_1| = \left| \sum_{k=2}^d k a_k \theta^{k-1} \right| \leq R \sum_{k=2}^d k |\theta|^{k-1}.$$

The last uses the fact that $|a_k| < R$ since $f \in P(R)$. Thus

$$|a_1| \leq R \sum_{k=2}^{\infty} k (\sigma^{1/3})^{k-1} \leq R \left\{ \sum_{k=0}^{\infty} k (\sigma^{1/3})^{k-1} - 1 \right\}.$$

But for $0 < \beta < 1$

$$\sum_{k=0}^{\infty} k \beta^{k-1} = \frac{d}{d\beta} \sum_{k=0}^{\infty} \beta^k = \frac{d}{d\beta} \left(\frac{1}{1-\beta} \right) = \frac{1}{(1-\beta)^2}.$$

We conclude for $\beta = \sigma^{1/3}$

$$|a_1| \leq R \left\{ \frac{1}{(1-\beta)^2} - 1 \right\} \leq R \frac{\beta(2-\beta)}{(1-\beta)^2}.$$

Now

$$\frac{2-\beta}{(1-\beta)^2} \leq 6 \quad \text{if } \beta \leq \frac{1}{2} \text{ or } \sigma \leq \frac{1}{8}.$$

Thus

$$|a_1| \leq R 6\beta = 6R\sigma^{1/3}.$$

This finishes the proof of Lemma 1.

The proof of (1) of Theorem 5 now goes as follows. We may assume $\sigma < \frac{1}{8}$. Use Lemma 1 with $\alpha = 6R\sigma^{1/3}$ taking into account the limitations $R > \frac{1}{3}$, $\sigma < \frac{1}{8}$. First

$$\frac{\text{Vol}(Y_\sigma \cap P(R))}{\text{Vol } P(R)} \leq \frac{\text{Vol}(U_{\sigma R}(W_*) \cap P(R))}{\text{Vol } P(R)} + \frac{\text{Vol}(Q_\sigma \cap P(R))}{\text{Vol } P(R)}.$$

By Theorem 4B,

$$\frac{\text{Vol}(U_{\sigma R}(W_*) \cap P(R))}{\text{Vol } P(R)} \leq 3\sigma(d-1)^2.$$

By Lemma 1 and Theorem 4C,

$$\frac{\text{Vol}(Q_\sigma \cap P(R))}{\text{Vol } P(R)} \leq \frac{\text{Vol}(L_{6R\sigma^{1/3}}(W_1) \cap P(R))}{\text{Vol } P(R)} \leq 144(d+2)\sigma^{2/3}.$$

We may now suppose that $150(d+2)^{4/3}\sigma^{2/3} < 1$. This implies $d^2\sigma < (1/150)^3 < 2^{3/2}$ so that $3\sigma(d-1)^2 \leq 6(d+2)^{4/3}\sigma^{2/3}$. On the other hand

$$144(d+2)\sigma^{2/3} \leq 144(d+2)^{4/3}\sigma^{2/3};$$

thus

$$3\sigma(d-1)^2 + 144(d+2)\sigma^{2/3} \leq 150(d+2)^{4/3}\sigma^{2/3}.$$

With our previous estimates, this proves (1) of Theorem 5.

LEMMA 2. $f \in U_\rho(W_0)$ if and only if $\rho_f < \rho$ (here $U_\rho(W_0)$ is defined in §4).

PROOF OF LEMMA 2. We have to show $f \in U_\rho(W_0)$ if and only if $|f(\theta)| < \rho$ for some critical point θ . By definition, $f \in U_\rho(W_0)$ implies there is some $f_0 \in W_0$ with $f \in U_\rho(f_0)$. Thus $|f(z) - f_0(z)| < \rho$ for all z . But there is some critical point θ with $f_0(\theta) = 0$ since $f_0 \in W_0$. Also f and f_0 have the same critical points. Thus $f \in U_\rho(W_0)$ implies $|f(\theta)| < \rho$ some critical point θ . Conversely if $|f(\theta)| < \rho$ for some critical point θ , define $f_0(z) = f(z) - f(\theta)$. Then $f_0 \in W_0$ and $f \in U_\rho(f_0)$. This proves Lemma 2.

Since $W_0 \subset W_*$, $U_\rho(W_0) \subset U_\rho(W_*)$ and $U_{\rho R}(W_0) \subset U_{\rho R}(W_*)$. This yields (2)(a) in Theorem 5.

Incidentally (not used here), we have as a Corollary (use Theorem 4A),

PROPOSITION.

$$\frac{\text{Vol}\{f \in \mathcal{P}_d | \rho_f < \rho\} \cap P(R)}{\text{Vol } P(R)} < \frac{d\rho^2}{R^2}.$$

For the proof of part (2)(b) of Theorem 5, suppose that $f \notin Q_\sigma$. We will show that under the estimate $4R \sin \mathfrak{I}C_f < \sigma^2, f \in U_{\sigma R}(W_*)$.

Consider (see Figure 2) (a similar picture with β reversed if $|f(\theta)| < |f(0)|$)

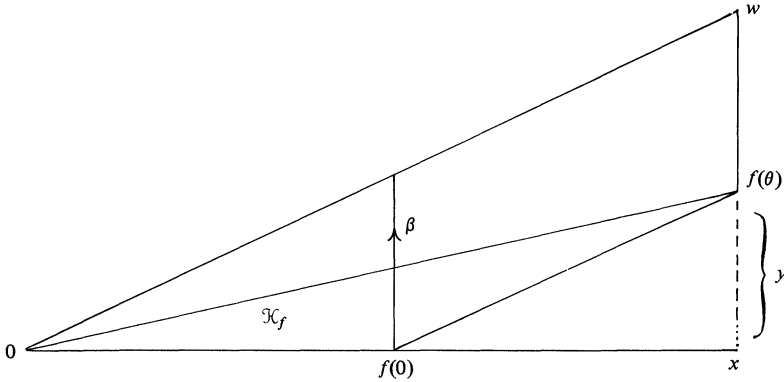


FIGURE 2

The segment $\overline{0w}$ is drawn parallel to $\overline{f(0)f(\theta)}$. We define β to be the vector starting at $f(0)$, perpendicular to $\overline{0f(0)}$ which meets the segment $\overline{0w}$, and $\beta = w - f(\theta)$ defines the magnitude of w .

Since $f(0) + \beta$ and $f(\theta) + \beta$ lie on the same ray,

$$\text{Im}((f(0) + \beta) / (f(\theta) + \beta)) = 0.$$

By a similar triangle argument $|\beta|/|f(0)| = y/(x - |f(0)|)$ where (x, y) are as in the figure. If $f(0) > 0$, then (x, y) are the Cartesian coordinates of $f(\theta)$.

We obtain,

$$\frac{|\beta|}{R} \leq \frac{\beta}{|f(0)|} = \frac{y}{|x - f(0)|} \leq \frac{y}{(\sigma^2 - y^2)^{1/2}}.$$

The last inequality is true since $|f(\theta) - f(0)| \geq \sigma$ ($f \notin Q_\sigma$) and thus

$$|x - f(0)|^2 + y^2 \geq \sigma^2.$$

So $|x - f(0)| \geq (\sigma^2 - y^2)^{1/2}$.

Furthermore

LEMMA 3.

$$\frac{y}{(\sigma^2 - y^2)^{1/2}} < \sigma.$$

PROOF OF LEMMA 3. By multiplying and squaring the inequality becomes

$$y^2 \leq \sigma^2(\sigma^2 - y^2) \quad \text{or} \quad y^2(1 + \sigma^2) \leq \sigma^4$$

or yet $y \leq \sigma^2/(1 + \sigma^2)^{1/2}$. Since $\sigma < 1$, $\frac{1}{2}\sigma^2 \leq \sigma^2/(1 + \sigma^2)^{1/2}$ and it is sufficient to prove that

$$2y \leq \sigma^2.$$

Now $y = x \sin \mathcal{I}_f$, $x \leq |f(\theta)| \leq 2R$, so by our hypothesis $2y = 2x \sin \mathcal{I}_f \leq 2|f(\theta)| \sin \mathcal{I}_f \leq 4R \sin \mathcal{I}_f < \sigma^2$. This proves Lemma 3.

Now by Lemma 3 and the preceding inequality we have

$$\frac{|\beta|}{R} < \sigma \quad \text{or} \quad |\beta| < \sigma R.$$

Let $f_0(z) = f(z) + \beta$. Then $f_0 \in W_1, f \in \overline{U_{|\beta|}(W_1)} \subset U_{\sigma R}(W_1)$.

This finishes the proof of Theorem 5.

THEOREM 6. Suppose $0 < \mu < 1$, d are given and $R > \frac{1}{3}$. Let σ be chosen so that $\text{Vol}(Y_\sigma \cap P(R))/\text{Vol } P(R) < \mu$ and let $f \in P(R)$ with $f \notin Y_\sigma$. Then for a suitable choice of h depending on (μ, d) , and $z_0 = 0$, $z_k = z_{k-1} - hf(z_{k-1})/f'(z_{k-1})$ is well defined for all k and z_s is an approximate zero of f where

$$s = 4 \left\{ 3 + \frac{\log(15/\sigma)8R}{\sigma^2} \right\}^2, \quad \sigma = \left(\frac{\mu}{150} \right)^{3/2} \frac{1}{(d+2)^2}.$$

Thus an approximate zero is obtained after s steps with probability of at least $1 - \mu$.

Define $\sigma(\mu)$ by

$$\sigma(\mu) = \left(\frac{\mu}{150} \right)^{3/2} \frac{1}{(d+2)^2}.$$

Therefore by Theorem 5(1),

$$\frac{\text{Vol}(Y_{\sigma(\mu)} \cap P(R))}{\text{Vol } P(R)} < \mu$$

for all μ , $0 < \mu < 1$.

This construction complements the second sentence of Theorem 6.

Now for the proof, suppose $f \notin Y_\sigma$, $f \in P(R)$. By Theorem 5(2), $\mathcal{I}_f, \rho_f > 0$, so Theorem 3 applies. In fact we obtain for a suitable h , the

conclusion of Theorem 6 relative to the s defined by

$$s = K \left\{ 3 + \frac{\log \zeta}{\sin(\mathfrak{C}/2)} \right\}^2.$$

Thus to finish the proof, we have to translate this function of ζ and \mathfrak{C} to our function of μ and d . For this we use the definition of ζ , the full strength of Theorem 5(2) and the explicit $\sigma = \sigma(\mu)$ stated previously.

Briefly then, since $|f(0)| < R$, $K = 4$, and $\rho_f > \sigma R$, we can take $\zeta = 15/\sigma$. Since, $4R \sin \mathfrak{C} \geq \sigma^2$, take $\sin(\mathfrak{C}/2) = \sigma^2/R$. This yields Theorem 6.

Note that crude estimates from Theorem 6 yield that if $R = 1$,

$$s = \frac{[100(d + 2)]^9}{\mu^7}$$

is sufficient. Thus the theorem stated in the introduction follows.

PART III

We devote this part to a discussion of open problems related especially to Part II.

PROBLEM 1A. Reduce the constant K of Theorem 1 (Remark 1) from 4. It is possible that $K = 1$. In the same body of ideas there are a number of other problems about polynomials, their critical points and values; we will go into this in more detail and suggest some way of looking at these problems.

In Theorem 1, the coefficients a_k are symmetric functions of the critical points $\theta_1, \dots, \theta_{d-1}$ of $f(z) = \sum_{k=0}^d a_k z^k$ provided $f(0) = 0$ (as we have assumed) if we take $a_d = 1$, as we will now. To see this, just solve these equations for a_k ,

$$\sum_1^d k a_k \theta_i^{k-1} = 0, \quad i = 1, \dots, d - 1$$

or better

$$\sum_1^d a_k z^k = \frac{1}{d} \int \prod_{i=1}^{d-1} (z - \theta_i) - \text{const.}$$

Recall the elementary symmetric functions which are defined by

$$S_k(z_1, \dots, z_{d-1}) = \sum_{i_1 < i_2 < \dots < i_k} z_{i_1} z_{i_2} \dots z_{i_k}.$$

It follows that

$$\left| \frac{a_k}{a_1} \right|^{1/(k-1)} \frac{1}{|a_1|} = \left| \frac{1}{k} S_{k-1} \left(\frac{1}{\theta_1}, \dots, \frac{1}{\theta_{d-1}} \right) \right|^{1/(k-1)} \frac{1}{|f'(0)|}$$

where $f'(0) = a_1 = d\theta_1 \dots \theta_{d-1}$.

Thus for example we get from Theorem 1 for $k = 2$ (recall $K = 2$ if $k = 2$).

PROPOSITION. If $f(0) = 0$, degree $f = d$, $a_d = 1$, then for some critical point θ ,

$$\left| \frac{1}{2} \sum_{i=1}^{d-1} \frac{1}{\theta_i} \left| \frac{f(\theta)}{f'(\theta)} \right| \right| \leq 2.$$

This brings our subject close to the theory of critical points of polynomials (see *Marden* for a survey). It also suggests an alternate approach to Theorem 1. One parametrizes all polynomials $f(z) = \sum a_i z^i$ of degree d with $f(0) = 0$, $a_d = 1$ by their critical points $\theta_1, \dots, \theta_{d-1}$ (of course we list a multiple critical point repeatedly).

Thus each nontrivial $(d-1)$ tuple $(\theta_1, \dots, \theta_{d-1})$ determines such a polynomial (and conversely up to an ordering of the θ_i).

Let $\theta = (\theta_1, \dots, \theta_{d-1}) \in \mathbb{C}^{d-1} - 0$, and suppose now $d > 2$. The previous proposition motivates the definition of these functions,

$$\psi_m(\theta) = \left(\frac{1}{2} \sum_{i=1}^{d-1} \frac{1}{\theta_i} \right) \frac{f(\theta_m)}{f'(\theta)}, \quad m = 1, \dots, d-1.$$

It can be checked that each ψ_m is homogeneous of degree 0 in the θ_i (it is a rational function) and can be thought of as $\psi_m: P_{d-2}(\mathbb{C}) - \Sigma \rightarrow \mathbb{C}$. Here $P_{d-2}(\mathbb{C})$ is complex projective space and $\theta \in \Sigma$ if some $\theta_i = 0$.

The preceding proposition translates to

$$\min_m |\psi_m(\theta)| \leq 2.$$

Consistent with our preceding problem is

PROBLEM 1B. Can 2 be replaced by $\frac{1}{2}((d-1)/d)$ in the previous proposition?

Note that the function $\theta \rightarrow \min_m |\psi_m(\theta)|$ is a continuous function on $P_{d-2}(\mathbb{C})$ and this gives a proof that $\min_m |\psi_m(\theta)|$ is bounded. A bound can be computed directly to be approximately 2^d , but this is too weak for the main theorem.

We are naturally led to the problem of finding the maximum over $\theta \in P_{d-2}(\mathbb{C})$ of $\min_m |\psi_m|$. Finding this maximum could lead to the solution of several of the problems in this section.

PROBLEM 1C. Does the point $\theta = (\theta_1, \dots, \theta_{d-1}) = (1, \dots, 1)$ maximize $\min_m |\psi_m|$?

An affirmative answer to Problem 1C would yield an affirmative answer to Problem 1B. Note that $\theta = (1, \dots, 1)$ corresponds to $f(z) = (z-1)^d - (-1)^d$.

An easier question is

PROBLEM 1D. Is $\theta = (1, \dots, 1)$ a *local* maximum for the function $\min_m |\psi_m|$?

Recall that Problems 1B, 1C, 1D all have to do with $k = 2$ in Theorem 1 and that for each k , there is a similar set of problems.

One should keep in mind for these questions, that very general maximum principles hold for complex analytic functions; see *Whitney*. Also I checked that these problems do have affirmative answers if $d = 3$ (and a number of other cases).

It is noteworthy that among the maxima (which might be unique) is one

which is *Pareto optimal* as defined in mathematical economics (see *Smale* (1974) or *Wan*). One says that $\theta^* \in P_{d-2}(\mathbb{C})$ is Pareto Optimal for $|\psi|$, $m = 1, \dots, d - 1$, provided there is no $\theta \in P_{d-2}(\mathbb{C})$ such that $|\psi_m(\theta)| > |\psi_m(\theta^*)|$ for all m with strict inequality for some m . Thus a Pareto optimality study of the $|\psi_m|$ might be in order.

Before we state the next problem, we prove a theorem which is an idealized version of the previous proposition, proved with crucial help from Mike Shub.

THEOREM. *If f is a polynomial with $f(0) = 0, f'(0) \neq 0$ then*

$$\min_{\substack{\theta \\ f'(\theta)=0}} \left| \frac{f(\theta)}{\theta} \right| \left| \frac{1}{f'(0)} \right| < 4.$$

The proof goes via the Koebe constant. Note first that the expression on the left has numerator, denominator each linear in a_d and homogeneous of degree d in the critical points $\theta_1, \dots, \theta_{d-1}$ of f . Here degree $f = d$.

We make the substitution $\theta_i \rightarrow \lambda\theta_i$, for all i , to reduce to the case $f'(0) = 1$. Next make a further substitution $f \rightarrow f(Rz)/R$ to reduce to the case $\min_{\theta} |f(\theta)| = 1$. Compare this to the proof of Theorem 1.

Now take the inverse of f to obtain a Schlicht function g with $g(0) = 0, g'(0) = 1, w \rightarrow g(w)$ 1-1 for $|w| < 1$. Now the Koebe theorem (see *Hille*) asserts that the image of the unit disk D_1 under g contains a disk D_κ of radius $\kappa > 0$ (κ is the Koebe constant). Bieberbach evaluated $\kappa = \frac{1}{4}$ at the time he proved $|a_2| \leq 2$ and proposed the Bieberbach conjecture. Thus $|\theta| \geq \frac{1}{4}$ in our situation so $1/|\theta| \leq 4$ and our theorem follows.

PROBLEM 1E. Can one replace 4 by 1 in the theorem (or yet $(d - 1)/d$)?

One cannot do better because of the example $f(z) = z^d - dz$.

The following is an incidental problem which came up as I was working on some of the previous problems.

PROBLEM 1F. What is the set of polynomials f of degree d (for each d) such that $f(0) = 0$, the leading coefficient a_d is 1, and $f(\theta) = \theta$ for each critical point θ ? That is to say, the critical points are fixed points.

If f' has simple zeros it must be $f(z) = z^d - (d/(1 - d))z$. Other examples are of the form $f(z) = (z - a)^d - (-a)^d$ or $(z^d - a)^l$ -constant. But there are other more subtle examples if $d \geq 4$. I believe that I computed them for $d \leq 5$ and showed that for each d there were only finitely many.

PROBLEM 2. Extend the main result to polynomial maps $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$ for each n . This is quite a nice problem. At first, I thought necessary estimates would involve some kind of Bieberback conjecture mathematics of several variables. But eventually I noticed the following counterexample to the Koebe theorem for $\mathbb{C}^2 \rightarrow \mathbb{C}^2$:

$$(z_1, z_2) \rightarrow (z_1, z_2 + \lambda z_1^2), \quad \lambda > 0.$$

This map is globally invertible. But the image of the unit polydisk gets thin as $\lambda \rightarrow \infty$. On the other hand since there are no critical points in \mathbb{C}^2 , the example is not bad from the point of view of computing by Newton type algorithms.

PROBLEM 3. Find a bound on cost for real polynomial maps $\mathbb{R}^n \rightarrow \mathbb{R}^n$. Here

the natural algorithm for $n > 1$ is the "Global Newton". See *Hirsch-Smale*. One must take into account negative Jacobian determinants.

PROBLEM 4. Reduce the number of steps (and/or find a shorter proof!) of the main theorem. One might vary h in §3. Also §§4 and 5 might be developed in a different way to reduce the degree in d . Also §4 suggests various geometric problems.

PROBLEM 5. Analyse Part II in terms of round off error. The algorithm is robust, but still there is a question here. One can see Wilkinson on this subject. Then there are related purely discrete or algebraic problems.

PROBLEM 6. Find an analogue for our main theorem for the simplex method of linear programming. This is a very well-known problem. For example in discussing Khachian's recent work, *Wolfe* writes . . . "Dantzig's 'simplex method' has been shown not to run in polynomial, but in exponential time, in the worst case. 'Worst case' behaviour is always the easiest to study; a theory of 'average' behaviour, which would explain the fact that, in practice, the simplex method acts like a highly efficient polynomial time algorithm, does not exist".

PROBLEM 7. Estimate the probability of (strict) Newton's method converging. From results and with the notations in Part II, this could be accomplished by estimating

$$\frac{\text{Volume}\{f \in P(R) \mid |f(0)| < \rho_f / (2K + 1)\}}{\text{Volume } P(R)}.$$

PROBLEM 8. Prove an analogue of our theorem for the Scarf-Eaves algorithm (see *Eaves and Scarf*). In fact there are a large number of problems in operations research and numerical analysis that suggest themselves.

PROBLEM 9. It is a fact essentially due to *Barna* that for a polynomial f with all roots real, Newton's method itself converges to a zero starting with almost every real number. The exceptional set of starting points is homeomorphic to the Cantor set.

In fact $T: P_R \rightarrow P_R$ defined by $T(x) = x - f(x)/f'(x)$ is an Axiom A dynamical system where P_R is real 1-dimensional projective space. One can use a theorem of Bowen and Ruelle to prove the measure-theoretic statement. The problem is to find an analogue of our main result for this situation.

We summarize by noting some points about our main result. One might ask what about letting the initial point of the algorithm z_0 vary as well as the polynomial. I believe everything goes through directly with no problem. What about letting the leading coefficient a_d vary, not just stay at 1? I haven't checked this and don't know what happens.

One can also ask to what extent is it a reasonable computational problem to let the degree of a polynomial grow large. In fact there can be a certain amount of ill-posedness by taking high powers. On the other hand there has been much successful numerical work on this, cf. *Dejon-Henrici*. For example, in this book of conference proceedings, there is reported work by Dejon and Nickel on a rather random polynomial of degree 100. All the roots were found. Hirsch and I worked on a number of simple examples with a PDP11 with good success. I chose this problem of working with a polynomial

equation because of the tradition associated to the fundamental theorem of algebra as well as its being a prototype of the fundamental and central problem of solving a (nonlinear) system of equations.

REFERENCES

- B. Barna, 1956, *Über die Divergenzpunkte des Newtonsches Verfahrens zur Bestimmung von Wurzeln Algebraischen Gleichungen*. II, *Publicationes Mathematicae*, Debrecen, vol. 4, pp. 384–397.
- L. E. J. Brouwer, 1924 (w. B. de Loor), *Intuitionistischer Beweis des Fundamentalsatzes der Algebra*, Coll. Works, vol. 1 (1975), North-Holland, Amsterdam.
- G. Collins, 1977, *Infallible calculation of polynomials to specified precision in mathematical software*. III, Academic Press, New York.
- G. Debreu, 1959, *Theory of value*, Yale Univ. Press, New Haven, Conn.
- B. Dejon and P. Henrici, 1969, *Constructive aspects of the fundamental theorem of algebra*, Wiley, New York.
- C. Eaves and H. Scarf, 1976, *The solution of systems of piecewise linear equations*, *Math. Operations Res.*, vol. 1, pp. 1–27.
- H. Eves, 1976, *An introduction to the history of mathematics* (4th ed.), Holt, Rinehart and Winston, New York.
- M. Garey and D. Johnson, 1979, *Computers and intractability*, Freeman, San Francisco.
- C. F. Gauss, 1973, *Werke*, Band X, Georg Olms Verlag, New York.
- P. Griffiths, 1978, *Complex differential and integral geometry and curvature integrals associated to singularities of complex analytic varieties*, *Duke Math. J.* **45**, pp. 427–512.
- J. Hartmanis, 1979, *Observations about the development of theoretical computer science*, 20th Annual Sympos. on Foundations of Computer Science, IEEE, Long Beach, Calif.
- W. Hayman, 1958, *Multivalent functions*, Cambridge Univ. Press, Cambridge, England.
- P. Henrici, 1977, *Applied and computational complex analysis*, Wiley, New York.
- E. Hille, 1962, *Analytic function theory*. II, Ginn, Boston.
- M. Hirsch, 1963, *A proof of the non-retractability of a cell onto its boundary*, *Proc. Amer. Math. Soc.* **14**, pp. 364–365.
- M. Hirsch and S. Smale, 1979, *On algorithms for solving $f(x) = 0$* , *Comm. Pure Appl. Math.* **32**, pp. 281–312.
- D. Hoffman and R. Osserman, (to appear) *The geometry of the generalized Gauss map*.
- H. Hotelling, 1939, *Tubes and spheres in n -space and a class of statistical problems*, *Amer. J. Math.* **61**, pp. 440–460.
- W. Hurewicz, 1958, *Lectures on ordinary differential equations*, MIT Press, Cambridge, Mass.
- J. Jenkins, 1965, *Univalent functions and conformal mapping*, Springer, New York.
- R. Kellog, T. Li and J. Yorke, 1976, *A constructive proof of the Brouwer fixed point theorem and computational results*, *SIAM J. Numer. Anal.* **13**, pp. 473–483.
- I. Lakatos, 1976, *Proofs and refutations*, Cambridge Univ. Press, Cambridge, England.
- S. Lang, 1965, *Algebra*, Addison-Wesley, Reading, Mass.
- M. Marden, 1966, *Geometry of polynomials*, *Math. Surveys*, no. 3, Amer. Math. Soc., Providence, R. I.
- A. Ostrowski, 1973, *Solutions of equations in Euclidean and Banach spaces*, Academic Press, New York.
- Jean-Claude Pont, 1974, *La topologie algébrique, des origines à Poincaré*, Presses Universitaires de France, Paris.
- L. Santalo, 1976, *Integral geometry and geometric probability*, Addison-Wesley, Reading, Mass.
- H. Scarf, 1973, *The computation of economic equilibria* (in collaboration with T. Hansen) Yale Univ. Press, New Haven, Conn.
- G. Schober, (to appear), *Coefficient estimates for inverses of Schlicht functions*, Proc. of NATO-LMS Conference on Aspects of Contemporary Complex Analysis, Academic Press, New York.

S. Smale, 1974, *Sufficient conditions for an optimum*, Dynamical Systems—Warwick 1974, Lecture Notes in Math., vol. 468, Springer-Verlag, Berlin and New York.

_____, 1976, *A convergent process of price adjustment and global Newton methods*, J. Math. Econom., 3, pp. 107–120.

_____, (to appear), *Global analysis and economics*, Handbook of Mathematical Economics (Arrow and Intrilligator, eds.), North-Holland, Amsterdam.

D. Smith, 1953, *History of mathematics*, vol. II, Dover, New York.

D. Struik, 1969, *A source book in mathematics, 1200–800*, Harvard Univ. Press, Cambridge, Mass.

J. Traub, ed., 1976, *Analytic computational complexity*, Academic Press, New York.

B. Van der Waerden, 1953, *Modern algebra*, Vol. I, Ungar, New York.

_____, 1950, *Modern algebra*, Vol. II, Ungar, New York.

Y.-H. Wan, 1975, *On local Pareto optima*, J. Math. Econom., 2, pp. 35–42.

H. Weyl, 1924, *Randbemerkungen zu Hauptproblem der Mathematik*, Math. Z. 20, pp. 131–150.

_____, 1939, *On the volume of tubes*, Amer. J. Math. 61, pp. 461–472.

H. Whitney, 1972, *Complex analytic varieties*, Addison-Wesley, Reading, Mass.

J. H. Wilkinson, 1963, *Rounding errors in algebraic processes*, Prentice-Hall, Englewood Cliffs, N. J.

P. Wolfe, 1980, *The ellipsoid algorithm* (letter to the editor), Science, 208, pp. 240–242.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA, BERKELEY, CALIFORNIA 94720