THE FUNDAMENTAL THEOREM OF ALGEBRA

THE PROOFS OF GAUSS

WITH NOTES

By

RICHARD MOORE LOTSPEICH

Bachelor of Arts
Union College
Barbourville, Kentucky
1948

Submitted to the faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
MASTER OF SCIENCE
May, 1959

THE FUNDAMENTAL THEOREM OF ALGEBRA

THE PROOFS OF GAUSS

WITH NOTES

Report Approved:

Report Advisor

Dean of the Graduate School

PREFACE

The Fundamental Theorem of Algebra, well known to every student of elementary algebra, has been proved many times in a variety of ways. The purpose of this report is not to offer any new proof but merely to give the reader the opportunity of examining proofs of this famous theorem written by one of the greatest mathematicians of the nineteenth century, Carl Friedrich Gauss. The proofs are significant in that they represent the first rigorous proofs of the theorem and because they represent the work of a mathematician who made contributions to almost every leading field of pure mathematics, astronomy, electricity and magnetism.

Grateful acknowledgments are due Professor John Hoffman, my report advisor, and Professor Alton Juhlin, of the library staff, for their valuable help during the writing of this report. Thanks are also due Dr. James A. Zant and the National Science Foundation for making this year's study possible.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

A fundamental problem in the theory of equations is the solution of the general equation of the n th degree in one unknown. Only very special cases of this problem were solved by the ancient and medieval mathematicians. Instances of the solution of special cases are furnished by: the geometrical representation of roots of equations of small degree by the early Greeks, the finding of one positive rational root of quadratic equations by Diophantus[7], and the recognition by the Hindus and the Arabs of the fact that at least some numerical quadratic equations have two roots.

Starting from such special cases as these, mathematicians have gradually comprehended the fact that not only is every equation of the form $ax^2 + bx + c = 0$ solvable in the field of complex numbers, but far more is true: Every algebraic equation of any degree n with real or complex coefficients,

$$f(x) = x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1 x + a_o = 0,$$

has solutions in the field of complex numbers. This elegant theorem is commonly known as the Fundamental Theorem of Algebra.

For equations of the 3rd and 4th degrees this was established in the sixteenth century by Tartaglia, Carden, and others, who solved such equations by formulas essentially similar to that for the quadratic equation, although much more complicated. For almost two hundred years the general equations of the 5th and higher degrees were intensively studied, but all efforts to solve them by similar methods failed. It was a great achievement when Carl Friedrich Gauss in his doctoral dissertation (1799) succeeded in giving the first complete proof that solutions exist. The question of generalizing the classical formulas, which express the solutions of equations of degree less than 5 in terms of the rational operations plus root extraction, remained unanswered until early in the nineteenth century when Abel conceived the idea of proving the impossibility of the solution of the general algebraic equation of degree n by means of radicals.

Though Gauss apparently introduced the term, fundamental theorem of algebra, it is not certain to whom the credit belongs for first stating this theorem. Professor C. Raymond Adams[1] credits Peter Rothe (1608) with recognizing that an algebraic equation of the nth degree may have n roots, and Albert Girard (1629) with asserting that "every algebraic equation has as many solutions as the exponent of the highest term indicates"; adding the qualification "unless the equation is incomplete" (that is, does not contain all powers of x from n down to zero). Girard did point out that if an equation has fewer roots than its degree indicates, it is useful to introduce as many "impossible" (i.e., complex) solutions as will make the total number of roots and impossible solutions equal the degree of the equation.

Before Gauss, several faulty proofs of the theorem were devised, notably by d'Alembert (1740), whose proof was so widely accepted that the theorem came to be known, at least in France, as d'Alemberts' theorem; by Euler (1749); by Lagrange (1772); and by LaPlace (1795).

Though van der Waerden[8] mentions five proofs, most historians agree that Gauss gave four proofs. The first was discovered in 1799 and constituted his dissertation, Demonstratio nova theorematis omnem functionem algebraicam rationalem integram unius variabilia in factores reales primi vel secundi gradus revolvi posse (A New Proof that Every Rational Integral Function of One Variable Can Be Resolved into Real Factors of the First or Second Degree). In this proof, Gauss gives one of the first coherent accounts of complex numbers, proving that all roots of any algebraic equation are "numbers" of the form $a + bi$, where a,b are real numbers, and i is the square root of -1. This new "number" was called complex. The significance of this first proof lies in the fact that it showed all previous proofs of this important theorem to be faulty and it gave a newly constructed rigorous proof. What Gauss actually proved is that any polynomial can be reduced to real factors of the first or second degree. The second and third proofs appeared in 1816; and the fourth, a simplification of the first, was published in 1850.

Gauss made the restriction that the coefficients in the equation be real; this, however, is not a serious defect since it can be shown that the case in which the coefficients are complex can be reduced to that in which they are real. (See note on fourth proof, page 29).

It is now generally agreed that this theorem does not belong to algebra. With the theorems of complex variable theory, the Fundamental Theorem can be established immediately as will be demonstrated in the closing chapter.

CHAPTER II

SECOND PROOF OF THE FUNDAMENTAL THEOREM OF ALGEBRA

This proof of the theorem, as translated by Professor C. Raymond Adams(1), has an introductory section which contains the proofs of certain theorems on the primality of rational integral functions and on symmetric functions. Only a brief resume of this section of the proof will be given as these theorems are now well known. From this point on the proof will be presented in full.

In the introductory section it is proved that if Y and Y' are any two integral functions of x, a necessary and sufficient condition that they have no common factor other than a constant is that there exist two other integral functions (the term integral function is used here in the sense of rational integral function) of x, Z and Z', satisfying the identity

$$ZY + Z'Y' = 1$$

It is pointed out that if a, b, c, ... is any set of m constants and if we define

$$v = (x-a)(x-b)(x-c)\ldots \equiv x^m - \lambda' x^{m-1} + \lambda'' x^{m-2} - \ldots,$$

each λ, or any function of the λ's is a symmetric function of a, b, c, ...

The next part of the introduction is devoted to proving that any integral symmetric function of a, b, c, ... is an integral function of the λ's; the uniqueness of this function of the λ's is established.

Next, the product

$$\pi = (a-b)(a-c)(a-d)\ldots x(b-a)(b-c)(b-d)\ldots x(c-a)(c-b)(c-d)\ldots x \ldots$$

is introduced. This is a certain integral function of λ', λ'',...; the same function of $\mathcal{L}'$, $\mathcal{L}''$, ... is denoted by p and is defined as the discriminant of the function

$$y = x^m - \mathcal{L}'x^{m-1} + \mathcal{L}''x^{m-2} - \ldots$$

This is regarded as any integral function of x of the mth degree with the leading coefficient 1, without regard to the question of factorability, and the $\mathcal{L}$'s are to be thought of as variables. On the other hand the function

$$Y = x^m - L'x^{m-1} + L''x^{m-2} - \ldots$$

is regarded as a particular, though arbitrary, function of the same type, with no restrictions on the coefficients, which are to be thought of as arbitrary constants. The value of p for $\mathcal{L}' = L'$, $\mathcal{L}'' = L''$, ... is denoted by P. It is with the factorability of Y that this proof is concerned. On the assumption that Y can be broken up into linear factors,

$$Y = (x-A)(x-B)(x-C)\ \ldots\ ,$$

the following theorems are proved:

Theorem I: If P, the discriminant of Y, is zero, Y and $Y' = \frac{dY}{dx}$ have a common factor.

Theorem II: If P, the discriminant of Y, is not zero, Y and Y' have no common factor.

This concludes the introductory section.

Theorem I and II will now be established. With the second, and simpler, we begin.

We will denote by ρ the function

$$\frac{\pi(x-b)(x-c)(x-d)\ldots}{(a-b)^2(a-c)^2(a-d)^2\ldots} + \frac{\pi(x-a)(x-c)(x-d)\ldots}{(b-a)^2(b-c)^2(b-d)^2\ldots} + \frac{\pi(x-a)(x-b)(x-d)\ldots}{(c-a)^2(c-b)^2(c-d)^2\ldots} + \ldots,$$

which, since π is divisible by the individual demominators, is an integral function of the unknowns $x, a, b, c, \ldots$ Furthermore, we set $dv/dx = v'$, where $v = (x-a)(x-b)(x-c)\ldots$, obtaining $v' = (x-b)(x-c)(x-d)\ldots + (x-a)(x-c)(x-d)\ldots + (x-a)(x-b)(x-d)\ldots + \ldots$ For $x = a$ we clearly have $\rho v' = \pi$, from which we conclude that the function $\pi - \rho v'$ is exactly divisible by $x-a$ (an integral function will be said to be exactly divisible by a second integral function of the same variables if the quotient of the first by the second is a third integral function of these variables). $\pi - \rho v'$ is also exactly divisible by $x-b$, $x-c$, ... and consequently also by the product v. If we set

$$\frac{\pi - \rho v'}{v} = \sigma,$$

σ is an integral function of the unknowns $x, a, b, c, \ldots$ and symmetric in the unknowns $a, b, c, \ldots$ Accordingly, there can be found two integral functions r and s of the unknowns $x, \lambda', \lambda'', \ldots$ which, when we make the substitutions $\lambda' = \lambda''$, $\lambda'' = \lambda''$, ..., become ρ and σ, respectively. If analogously we denote the function

$$mx^{m-1} - (m-1)\lambda' x^{m-2} + (m-2)\lambda'' x^{m-3} - \ldots,$$

that is, the derivative dy/dx, by y', so that y' also goes over by those substitutions in v', then clearly by those same substitutions $p - sy - ry'$ goes over into $\pi - \sigma v - v'$, that is, into zero, and must therefore vanish identically. Hence we have the identity

$$p = sy + ry'$$

If we assume that by the substitution $\mathcal{L}' = L'$, $\mathcal{L}'' = L''$,... r and s become respectively R and S, we have also the identity

$$P = SY + RY' ;$$

and since S and R are integral functions of x, and P is a definite quantity or number, it follows at once that Y and Y' can have no common factor if P is not zero.

The proof of Theorem I will be constructed by showing that if Y and Y' have no common factor, P can certainly not be zero. To this end we determine two integral functions of the unknown x, say f(x) and φ(x), such that the identity

$$f(x)\cdot Y + \varphi(x)\cdot Y' = 1$$

holds; this we can also write as

$$f(x)\cdot v + \varphi(x)\cdot v' = 1 + f(x)\cdot(v-Y) + \varphi(x)\,\frac{d(v-Y)}{dx}$$

or, since we have

$$v' = (x-b)(x-c)(x-d)\,\ldots + (x-a)\,\frac{d\,[(x-b)(x-c)(x-d)\ldots]}{dx} ,$$

in the form

$$\varphi(x)\;(x-b)(x-c)(x-d)\ldots + \varphi(x)\cdot(x-a)\frac{d\,[(x-b)(x-c)(x-d)\ldots]}{dx}$$

$$+\; f(x)\cdot(x-a)(x-b)(x-c) = 1 + f(x)\cdot(v-Y) + (x)\;\frac{d(v-Y)}{dx}$$

For brevity we will denote the expression

$$f(x)\cdot(y-Y) + \varphi(x)\;\frac{d(v-Y)}{dx}$$

which is an integral function of the unknowns x, $\mathcal{L}'$, $\mathcal{L}''$, ... , by

$$F(x, \mathcal{L}', \mathcal{L}'', \ldots) ;$$

hence we have identically

$$1 + f(x)\cdot(v-Y) + \varphi(x)\cdot\frac{d(v-Y)}{dx} = 1 + F(x, \lambda', \lambda'', \ldots),$$

and therefore the identities

$$(1) \quad \varphi(a)\cdot(a-b)(a-c)(a-d)\ldots = 1 + F(a, \lambda', \lambda'', \ldots),$$
$$\varphi(b)\cdot(b-a)(b-c)(b-d)\ldots = 1 + F(b, \lambda', \lambda'', \ldots),$$
$$\ldots\ldots\ldots\ldots$$

If then, we assume that the product of all the functions

$$1 + F(a, \ell', \ell'', \ldots),\ 1 + F(b, \ell', \ell'', \ldots),\ \ldots,$$

which is an integral function of the unknowns $a, b, c, \ldots, \ell'$, ℓ'', ... and indeed a symmetric function of $a, b, c, \ldots$, is denoted by

$$\psi(\lambda', \lambda'', \ldots, \ell', \ell'', \ldots),$$

there follows from the multiplication of all the equations of (1) the new identity

$$(2) \quad \pi\varphi a\cdot\varphi b\cdot\varphi c \ldots = \psi(\lambda', \lambda'', \ldots, \lambda', \lambda'', \ldots).$$

It is furthermore clear that since the product $\varphi a\cdot\varphi b\cdot\varphi c \ldots$ involves the unknowns $a, b, c, \ldots$ symmetrically, an integral function of the unknowns $\ell' = \lambda', \ell'' = \lambda''$, ... goes over into $\varphi a\cdot\varphi b\cdot \varphi c \ldots$ If t is this function we have identically

$$(c) \quad pt = \psi(\ell', \ell'', \ldots, \ell', \ell'', \ldots),$$

for by the substitution $\ell' = \lambda', \ell'' = \lambda''$, ... this equation becomes the identity (2).

From the definition of the function F follows immediately the identity

$$F(x, L', L'', \ldots) = 0$$

Hence we have successively the following identities.

$$1 + F(a, L', L'', \ldots) = 1,\ 1 + F(b, L', L'', \ldots) = 1, \ldots,$$
$$\psi(\lambda', \lambda'', \ldots, L', L'', \ldots) = 1,$$

and

(4) $$\psi(\mathfrak{L}', \mathfrak{L}'', \ldots, L', L'', \ldots,) = 1$$

From equations (3) and (4) jointly, if we set $\mathfrak{L}' = L', \mathfrak{L}'' = L'', \ldots$, follows the relation

(5) $$PT = 1$$

where T denotes the value of the function t that corresponds to those substitutions. Since this value must be finite, P can certainly not be zero.

From the foregoing it is apparent that every integral function Y of an unknown x whose discriminant is zero can be broken up into factors of which none has a vanishing discriminant. In fact, if we find the greatest common divisor of the function Y and $\frac{dY}{dx}$, Y is thereby broken into two factors. If one of these factors again has the discriminant zero, it may in the same way be broken into two factors, and so we shall proceed until Y is finally reduced to factors no one of which has the discriminant zero.

Moreover, one sees that of those factors into which Y has been broken, at least one has the property that among the factors of its degree index the factor 2 is present no more frequently than it occurs among the factors of m, the degree index of Y; accordingly, if we set $m = k \cdot 2^{\mu}$, where k is odd, there will be among the factors of Y at least one whose degree is $k' 2^{v}$, k' being odd and $v = \mu$ or $v \geq \mu$. The validity of this assertion follows immediately from the fact that m is the sum of the numbers which indicate the degree of the individual factors of Y.

Before proceeding further, we will explain an expression whose introduction is of the greatest use in all investigations of symmetric functions and which will be exceedingly convenient also for our purposes. We assume that M is a function of some of the unknowns a,b,c, ... Let μ be the number of those which enter into the expression M, without reference to other unknowns which perhaps are present in M. If these μ unknowns are permuted in all possible ways, not only among themselves but also with the $m-\mu$ remaining unknowns of the set a,b,c, ... , there arise from M other expressions similar to M, so that we have in all

$$m(m-1)(m-2) \ldots (m-\mu+1)$$

expressions, including M itself; the set of these we call simply the set of all M. From this, it is clear what is to be understood by the sum of all M, the product of all M, ... Thus, for example, π can be called the product of all a-b, v' the product of all x-a, v' the sum of all $\frac{v}{x-a}$, etc.

If M is a symmetric function of some of the μ unknowns which it contains, the permutations of these among themselves will not alter the function M; hence in the set of all M, every term is multiple and in fact, is present $1\cdot 2 \ldots v$ times if v stands for the number of unknowns in which M is symmetric. But if M is symmetric not only in v unknowns but also in v' others, and in v'' still different unknowns, etc., then M is unchanged if any two of the first v unknowns are permuted among themselves, or any two of the following v' among themselves, or any two of the next v'' among themselves, etc., so that identical terms always correspond to

$$1\cdot 2 \ldots v\cdot 1\cdot 2 \ldots v'\cdot 1\cdot 2 \ldots v'' \ldots$$

permutations. If then, from these identical terms we retain only one of each, we have in all

$$\frac{m(m-1)(m-2)\ \ldots\ (m-\mu+1)}{1\cdot 2\ \ldots\ v\cdot 1\cdot 2\ \ldots\ v'\cdot 1\cdot 2\ \ldots\ v''\ \ldots}$$

terms, the set of which we call the set of all M without repetitions to distinguish it from the set of all M with repetitions. Unless otherwise stated, we shall always admit the repetitions.

Furthermore the sum of all M, or the product of all M, or more generally, any symmetric function whatever of all M is always a symmetric function of the unknowns a,b,c,... , whether repetitions are admitted or excluded.

We will now consider the product of all $u-(a+b)x+ab$ without repetitions, where u and x indicate unknowns, and denote the same by f . Then f will be the product of the following $1/2\,m(m-1)$ factors:

$$u-(a+b)x+ab,\ u-(a+c)x+ac,\ u-(a+d)x+ad,\ \ldots\ ;$$
$$u-(b+c)x+bc,\ u-(b+d)x+bd,\ \ldots\ ;$$
$$u-(c+d)x+cd,\ \ldots\ ;\ \ldots$$

Since this function involves the unknowns a,b,c,... symmetrically, it determines an integral function of the unknowns $u, x, \ell', \ell'', \ldots$, which shall be denoted by z, with the property that it goes over into f if the unknowns $\ell', \ell'', \ldots$ are replaced by $\lambda', \lambda'', \ldots$ Finally we will denote by Z the function of the unknowns u and x alone to which z reduces if we assign to the unknowns ℓ', ℓ'', ... the particular values L', L'', ...

These three functions f , z, and Z can be regarded as integral functions of degree $1/2m(m-1)$ of the unknown u with undetermined coefficients; these coefficients are

for f , functions of the unknowns x,a,b,c,...

for z, functions of the unknowns $x, \Omega', \Omega'', \ldots$

for Z, functions of the single unknown x.

The individual coefficients of z will go over into the coefficients of f by the substitutions $\Omega' = \lambda', \Omega'' = \lambda'', \ldots$ and likewise into the coefficients of Z by the substitutions $\Omega' = L', \Omega'' = L'', \ldots$ The statements made here for the coefficients hold also for the discriminants of the functions f, z, and Z. These we will examine more closely for the purpose of obtaining a proof of the following theorem.

Theorem: Whenever P is not zero, the discriminant of the function Z cannot vanish identically.

The proof of this theorem will be omitted.

The discriminant of the function f is the product of all differences between pairs of quantities $(a+b)x - ab$, the total number of which is

$$1/2\, m(m-1)\left[1/2\, m(m-1) - 1\right] = 1/4\,(m+1)\, m(m-1)(m-2).$$

This number also expresses the degree in x of the discriminant of the function f. The discriminant of the function z will be of the same degree, while the discriminant of the function Z can be of lower degree if some of the coefficients of the highest power of x vanish. Our problem is to prove that in the discriminant of the function Z certainly not all the coefficients can be zero.

If we examine more closely the differences whose product is the discriminant of the function f, we notice that a part of them (that is, those differences between two quantities $(a+b)x - ab$ which have a common element) provides the product of all $(a-b)(x-c)$; from the others (that is, those differences between two quantities $(a+b)x - ab$ which have no common element) arises the product of all

$$(a+b-c-d)x-ab+cd$$

without repetitions. The first product contains each factor $a-b$ clearly $m-2$ times, whereas each factor $x-c$ is contained $(m-1)\cdot(m-2)$ times; from this it is easily seen that the value of this product is

$$\pi^{m-2} v^{(m-1)(m-2)} .$$

If further we indicate by r that functions of the unknowns x, $\mathcal{L}'$, $\mathcal{L}''$, ... which by the substitutions $\mathcal{L}'=\lambda'$, $\mathcal{L}''=\lambda''$,... goes over into ρ, and by R that function of x alone into which r goes over by the substitutions $\mathcal{L}'=L'$, $\mathcal{L}''=L''$, ..., the discriminant of the function z will be equal to

$$p^{m-2} y^{(m-1)(m-2)} r,$$

while the discriminant of the function Z will be

$$P^{m-2} Y^{(m-1)(m-2)} R .$$

Since by hypothesis P is not zero, it now remains to be shown that R cannot vanish identically.

To this end we introduce another unknown w and will consider the product of all

$$(a+b-c-d)w+(a-c)(a-d)$$

without repetitions; since this involves the $a,b,c,\ldots$ symmetrically, it can be expressed as an integral function of the unknowns w, λ', λ'', ... We denote this function by $f(w,\lambda',\lambda'',\ldots)$. The number of the factors $(a+b-c-d)w+(a-c)(a-d)$ will be

$$1/2\, m(m-1)(m-2)(m-3),$$

from which easily follow in succession the equalities

$$f(0,\lambda',\lambda'',\ldots) = \pi^{(m-2)(m-3)}$$

$$f(0, \mathcal{L}', \mathcal{L}'', \ldots) = p^{(m-2)(m-3)}$$

and

$$f(0, L', L'', \ldots) = p^{(m-2)(m-3)}$$

The function $f(w,L',L'',\ldots)$ must in general be of degree

$$1/2\, m(m-1)(m-2)(m-3) ;$$

only in particular cases can it reduce to lower degree, if some coefficients of the highest power of w vanish; it is, however, impossible for it to be identically zero, since as the above equation shows, at least the last term of the function does not vanish. We will assume that the highest term of the function $f(w,L',L'',\ldots)$ to have a non-vanishing coefficient is Nw^v. If we make the substitution $w = x - a$, it is clear that $f(x - a, L', L'', \ldots)$ is an integral function of the unknowns x and a, or what is the same thing, an integral function of x whose coefficients depend upon the unknown a; its highest term is Nx^v and it therefore has a coefficient that is independent of a and different from zero. In the same way $f(x - b,L',L'',\ldots)$, $f(x - c,L',L'',\ldots)$, ... are integral functions of the unknown x which individually have Nx^v as highest term, while the coefficients of the remaining terms depend upon $a,b,c,\ldots$ Hence the product of the m factors $f(x - a, L',L'',\ldots)$, $f(x - b,L',L'',\ldots)$, $f(x - c, L',L'',\ldots)$,... will be an integral function of x whose highest term is $N^m x^{mv}$, whereas the coefficients of the subsequent terms depend upon $a,b,c,\ldots$

We now consider the product of the m factors

$$f(x-a,\mathcal{L}',\mathcal{L}'',\ldots), f(x-b,\mathcal{L}',\mathcal{L}'',\ldots), f(x-c,\mathcal{L}',\mathcal{L}'',\ldots),\ldots,$$

which as a function of the unknowns $x,a,b,c,\ldots,\mathcal{L}',\mathcal{L}'',\ldots$, symmetric in the $a,b,c,\ldots$, can be expressed in terms of the unknowns

$x, \lambda', \lambda'', \ldots, \mathcal{l}', \mathcal{l}'', \ldots$ and denoted by

$$\varphi(x, \lambda', \lambda'', \ldots, \mathcal{l}', \mathcal{l}'', \ldots)$$

Thus

$$\varphi(x, \lambda', \lambda'', \ldots, \lambda', \lambda'', \ldots)$$

becomes the product of the factors

$$f(x-a, \lambda', \lambda'', \ldots), f(x-b, \lambda', \lambda'', \ldots), f(x-c, \lambda', \lambda'', \ldots), \ldots$$

and is exactly divisible by ρ, since as is easily seen each factor of ρ is contained in one of these factors. We will therefore set

$$\varphi(x, \lambda', \lambda'', \ldots, \lambda', \lambda'', \ldots) = \rho\psi(x, \lambda', \lambda'', \ldots),$$

ψ indicating an integral function. From this follows at once the identity

$$\varphi(x, L', L'', \ldots, L', L'', \ldots) = R\psi(x, L', L'', \ldots).$$

We have proved above, however, that the product of the factors $f(x-a, L', L'', \ldots)$, $f(x-b, L', L'', \ldots)$, $f(x-c, L', L'', \ldots)$, ..., which is $\varphi(x, \lambda', \lambda'', \ldots, L', L'', \ldots)$ has $N^m x^{mv}$ as its highest term; hence the function $\varphi(x, L', L'', \ldots, L', L'', \ldots)$ will have the same highest term and accordingly, will not be identically zero. Therefore R, and likewise the discriminant of the function Z, cannot be identically zero.

<u>Theorem</u>: If $\varphi(u,x)$ denotes the product of an arbitrary number of factors which are linear in u and x and so of the form

$$\alpha + \beta u + \quad x, \alpha' + \beta' u + \quad ' x, \alpha'' + \beta'' u + \quad '' x, \ldots,$$

and if w is another unknown, the function

$$\left(u + w\frac{d\varphi(u,x)}{dx}, x - w\frac{d\varphi(u,x)}{du}\right) = \Omega$$

will be exactly divisible by $\varphi(u,x)$.

Proof: If we set

$$\varphi(u,x) = (\alpha + \beta u + \gamma x)\, Q = (\alpha' + \beta' u + \gamma' x)\, Q' = \ldots,$$

then $Q, Q', \ldots$ will be integral functions of the unknowns $u, x, \alpha, \beta, \gamma, \alpha', \beta', \gamma', \ldots$ and we shall have

$$\frac{d\varphi(u,x)}{dx} = \gamma Q + (\alpha + \beta u + \gamma x)\frac{dQ}{dx}$$

$$= \gamma' Q' + (\alpha' + \beta' u + \gamma' x)$$

$$\frac{d\varphi(u,x)}{du} = \beta Q + (\alpha + \beta u + \gamma x)\frac{dQ}{du}$$

$$= \beta' Q' + (\alpha' + \beta' u + \gamma' x)$$

If we introduce these values into the factors of the product , that is, into

$$\alpha + \beta u + \gamma x + \beta w\frac{d\varphi(u,x)}{dx} - \gamma w\frac{d\varphi(u,x)}{du},$$

$$\alpha' + \beta' u + \gamma' x + \beta' w\frac{d\varphi(u,x)}{dx} - \gamma' w\frac{d\varphi(u,x)}{du}, \ldots,$$

we obtain the expressions

$$(\alpha + bu + \gamma x)\left(1 + \beta w\frac{dQ}{dx} - \quad w\frac{dQ}{du}\right),$$

$$(\alpha' + \beta' u + \gamma' x)\left(1 + \beta' w\frac{dQ'}{dx} - \gamma' w\frac{dQ'}{du}\right), \ldots,$$

so that Ω becomes the product of $\varphi(u,x)$ and the factors

$$1 + \beta w\frac{dQ}{dx} - \gamma w\frac{dQ}{du},\ 1 + \beta' w\frac{dQ'}{dx} - \gamma' w\frac{dQ'}{du}, \ldots,$$

that is, of $\varphi(u,x)$ and an integral function of the unknowns $u, x, w, \alpha, \beta, \gamma, \alpha', \beta', \gamma', \ldots$

The theorem of the foregoing paragraph is clearly applicable to the function f, which from now on we will denote by

$$f(u, x, \lambda', \lambda'', \ldots),$$

so that

$$f\left(u + w\frac{d\int}{dx},\ x - w\frac{d\int}{du},\ \lambda',\ \lambda'',\ \ldots\right)$$

is exactly divisible by $\int$; the quotient, which is an integral function of the unknowns u,x,w,a,b,c,... and is symmetric in a,b, c,..., we will denote by

$$\psi(u,x,w,\lambda',\lambda'',\ldots).$$

From this follow the identities

$$f\left(u + w\frac{dz}{dx},\ x - w\frac{dz}{du},\ \ell',\ \ell'',\ldots\right) = z\psi(u,x,w,\ell',\ell'',\ldots),$$

$$f\left(u + w\frac{dZ}{dx},\ x - w\frac{dZ}{du},\ L',\ L'',\ldots\right) = Z\psi(u,x,w,L',L'',\ldots).$$

If then we indicate the function Z simply by F(u,x), that is, set

$$f(u,x,L',L'',\ldots) = F(u,x),$$

we shall have the identity

$$F\left(u + w\frac{dZ}{dx},\ x - w\frac{dZ}{du}\right) = Z\psi(u,x,w,L',L'',\ldots).$$

Assuming that particular values of u and x, say u = U and x = X, give

$$\frac{dZ}{dx} = X',\ \frac{dZ}{du} = U',$$

we have identically

$$F(U + wX',\ X - wU') = F(U,X)\psi(U,X,w,L',L'',\ldots).$$

Whenever U' does not vanish we can set

$$w = \frac{X - x}{U'}$$

and obtain

$$F\left(U + \frac{XX'}{U'} - \frac{X'x}{U'},\ x\right) = F(U,X)\,\psi\left(U,\ X,\ \frac{X - x}{U'},\ L',L'',\ldots\right).$$

If we set $u = U + \frac{XX'}{U'} - \frac{X'x}{U'}$, the function Z therefore becomes

$$F(U,X)\psi\left(U, X, \frac{X-x}{U'}, L', L'', \ldots\right).$$

Since in case P is not zero the discriminant of the function Z is a function of the unknown x that is not identically zero, the number of particular values of x for which this discriminant can vanish is finite; accordingly an infinite number of values of the unknown x can be assigned which give this discriminant a value different from zero. Let X be such a value of x (which moreover we may assume real). Then the discriminant of the function F(u,X) will not be zero and it follows by Theorem II that the functions

$$F(u,X) \text{ and } \frac{dF(u,X)}{du}$$

can have no common divisor. We will further assume that there is a particular value U of u, which may be real or imaginary, that is, of the form $g + h\sqrt{-1}$, and which makes $F(u,X) = 0$, so that $F(U,X) = 0$. Then $u - U$ will be a factor of the function $F(u, X)$ and hence the function $\frac{dF(u,X)}{du}$ is certainly not divisible by u-U. If then we assume that this function $\frac{dF(u,X)}{du}$ takes on the value U' for $u = U$, surely U' cannot be zero. Clearly, however, U' is the value of the partial derivative $\frac{dZ}{du}$ for $u = U$, $x = X$; if then we denote by X' the value of the partial derivative $\frac{dZ}{dx}$ for the same values of u and x, it is clear from the proof in the foregoing section that by the substitution

$$u = U + \frac{XX'}{U'} - \frac{X'x}{U'}$$

the function Z vanishes identically and so is exactly divisible by the factor

$$u + \frac{X'}{U'}x - \left(U + \frac{XX'}{U'}\right).$$

If we set $u = x^2$, clearly $F(x^2, x)$ is divisible by

$$x^2 + \frac{X'}{U'}x - \left(U + \frac{XX'}{U'}\right)$$

and thus takes on the value zero if for x we take a root of the equation

$$x^2 + \frac{X'}{U'}x - U + \frac{XX'}{U'} = 0$$

that is,

$$x = \frac{-X' \pm \sqrt{(4UU'U' + 4XX'U' + X'X')}}{2U'}$$

These values are either real or of the form $g + h\sqrt{-1}$.

Now it can be easily shown that for these same values of x the function Y also must vanish. For it is clear that $f(xx, x, \lambda', \lambda'', \ldots)$ is the product of all $(x-a)(x-b)$ without repetitions and so equals v^{m-1}. From this follow immediately

$$f(xx, x, \ell', \ell'', \ldots) = y^{m-1},$$

$$f(xx, x, L', L'', \ldots) = Y^{m-1},$$

or $F(xx, x) = Y^{m-1}$; accordingly, a particular value of this function F cannot be zero unless at the same time the value of Y is zero.

By the above investigations the solution of the equation $Y = 0$, that is the determination of a particular value of x which satisfies the equation and is either real or of the form $g + h\sqrt{-1}$, is made to depend upon the solution of the equation $F(u, X) = 0$, provided the discriminant of the function Y is not zero. It may be remarked that if all the coefficients in Y, that is, the numbers

L',L", ... , are real, and if as is permissible we take a real value for X, all the coefficients in F(u,X) are also real. The degree of the auxiliary equation F(u,X) = 0 is expressed by the number $1/2\,m(m-1)$; if then m is an even number of the form $2^{\mu}k$, k designating an odd number, the degree of the second equation is expressed by a number of the form $2^{\mu-1}k$.

In case the discriminant of the function Y is zero, it will be possible to find another function Φ which is a divisor of Y, whose discriminant is not zero, and whose degree is expressed by a number $2^{\nu}k$ with $\nu < \mu$. Every solution of the equation $\Phi = 0$ will also satisfy the equation Y = 0; the solution of the equation $\Phi = 0$ is again made to depend upon the solution of another equation whose degree is expressed by a number of the form $2^{\nu-1}k$.

From this we conclude that in general the solution of every equation whose degree is expressed by an even number of the form $2^{\mu}k$ can be made to depend upon the solution of another equation whose degree is expressed by a number of the form $2^{\mu'}k$ with $\mu' < \mu$. In case this number also is even, —that is, if μ' is not zero, — this method can be applied again, and so we proceed until we come to an equation whose degree is expressed by an odd number; the coefficients of this equation are all real if all the coefficients of the original equation are real. It is known, however, that such an equation of odd degree is solvable and indeed has a real root. Hence each of the preceding equations is solvable, having either real roots or roots of the form $g + h\sqrt{-1}$.

Thus it has been proved that every function Y of the form $x^m - L'x^{m-1} + L''x^{m-2} - \dots$, in which L',L", ... are particular real numbers, has a factor $x-A$ where A is real or of the form $g + h\sqrt{-1}$. In the second case it is easily seen that Y is also zero for $x = g-h\sqrt{-1}$ and therefore divisible by $x-(g-h\sqrt{-1})$ and so by the product $xx - 2gx + gg + hh$. Consequently every function Y certainly has a real factor of the first or second degree. Since the same is true of the quotient [of Y by this factor], it is clear that Y can be reduced to real factors of the first or second degree. To prove this fact was the object of Gauss' second proof of the fundamental theorem.

CHAPTER III

THE THIRD PROOF OF THE FUNDAMENTAL THEOREM

The translation of Gauss' third proof by Professor Maxime Bocher[5] is as follows:

Let $f(z) = 0$ be the equation (of the nth degree) for which we wish to prove the existence of a root, and suppose that in the polynomial $f(z)$ the coefficient of z^n is 1. The idea which underlies the proof we shall give is, that if we can prove that $\phi(z)/f(z)$, where $\phi(z)$ is a polynomial, does not remain finite for all values of z, $f(z) = 0$ must have a root. Let

$$\phi(z) = zf'(z) = \frac{zdf(z)}{dz}$$

Write

$$F(z) = \frac{zf'(z)}{f(z)} = u(x,y) + iv(x,y),$$

where

$$z = x + yi$$

Note that: (1) $u(0,0) = 0$; (2) if we describe a circle of radius a about the origin, $u(x,y)$ can be made positive at all points on the circumference of this circle by taking a sufficiently large, since $F(\infty) = n$.

Let

$$f(z) = z^n + (a_1 + b_1 i)z^{n-1} + \dots + (a_{n-1} + b_{n-1} i)z + a_n + bi = \sigma + \tau i,$$

and let

$$zf'(z) = \sigma' + \tau' i.$$

Then letting $z = r(\cos\phi + i\sin\phi)$, we have

$$\sigma = r^n \cos n\phi + a_1 r^{n-1} \cos(n-1)\phi + \ldots$$
$$- b_1 r^{n-1} \sin(n-1)\phi - \ldots$$

$$\tau = r^n \sin n\phi + a_1 r^{n-1} \sin(n-1)\phi + \ldots$$
$$+ b_1 r^{n-1} \cos(n-1)\phi + \ldots$$

$$\sigma' = nr^n \cos n\phi + (n-1)\, a_1 r^{n-1} \cos(n-1)\phi + \ldots$$
$$- (n-1)\, b_1 r^{n-1} \sin(n-1)\phi - \ldots$$

$$\tau' = nr^n \sin n\phi + (n-1)\, a_1 r^{n-1} \sin(n-1)\phi + \ldots$$
$$+ (n-1)\, b_1 r^{n-1} \cos(n-1)\phi + \ldots,$$

$$F(z) = \frac{\sigma' + \tau' i}{\sigma + \tau i} = \frac{\sigma\sigma' + \tau\tau'}{\sigma^2 + \tau^2} + \frac{\sigma\tau' - \tau\sigma'}{\sigma^2 + \tau^2} i = u + vi$$

We wish now to find the derivatives of u and v with regard to r and ϕ. For this purpose we note the following relations:

$$\frac{\delta\sigma}{\delta r} = \frac{\sigma'}{r}, \quad \frac{\delta\sigma}{\delta\phi} = -\tau',$$

$$\frac{\delta\tau}{\delta r} = \frac{\tau'}{r}, \quad \frac{\delta\tau}{\delta\phi} = \sigma'$$

We also have formulae of precisely the same sort for expressing the derivatives of σ' and τ' with regard to r and ϕ in terms of σ'' and τ'' where

$$\sigma'' = n^2 r^n \cos n\phi + (n-1)^2 a_1 r^{n-1} \cos(n-1)\phi + \ldots$$
$$- (n-1)^2 b_1 r^{n-1} \sin(n-1)\phi - \ldots,$$

$$\tau'' = n^2 r^n \sin n\phi + (n-1)^2 a_1 r^{n-1} \sin(n-1)\phi + \ldots$$
$$+ (n-1)^2 b_1 r^{n-1} \cos(n-1)\phi + \ldots$$

We get then by direct differentiation

$$\frac{\delta u}{\delta r} = \frac{1}{r}\frac{\delta v}{\delta \phi}$$

$$= \frac{(\sigma^2+\tau^2)(\sigma\sigma''+\tau\tau'') + (\sigma\tau' - \tau'\sigma')^2 - (\sigma\sigma' + \tau\tau')^2}{r(\sigma^2 + \tau^2)^2} = T.$$

Now form the double integral

$$\Omega = \int_0^a \int_0^{2\pi} T d\phi dr \ .$$

If here we integrate first with regard to ϕ and then with regard to r, we obviously get $\Omega = 0$. If, however, we integrate first with regard to r and then with regard to ϕ, we get, remembering that u vanishes at the origin,

$$\Omega = \int_0^{2\pi} u d\phi \ ,$$

the integral being taken around the circumference of a circle with radius <u>a</u> and center at the origin, so that Ω will be positive if <u>a</u> is sufficiently large. The fact that we get different values for Ω according to the order of integration shows that T cannot be everywhere finite, continuous, and single valued, and this can be explained only by the vanishing of $\sigma^2 + \tau^2$ (since r, which also occurs in the denominator of T is a factor of each term of the numerator). A point where $\sigma^2 + \tau^2$ vanishes is a root of $f(z) = 0$.

CHAPTER IV

THE FOURTH PROOF OF THE FUNDAMENTAL THEOREM[6]

Let $z = x + iy$, then the variable represents points in a plane, and the function $f(z)$ has a definite value at each point in the plane. We may write $f(z) = P + iQ$, where P and Q are functions of x and y with real coefficients. To find expressions for P and Q, let $x = r\cos\phi$, $y = r\sin\phi$. By De Moivre's Theorem,

$$z^m = r^m(\cos\phi + i\sin\phi)^m = r^m(\cos m\phi + i\sin m\phi).$$

Substituting for z in $f(z)$, we get,

$$P = r^n\cos n\phi + a_1r^{n-1}\cos(n-1)\phi + a_2r^{n-2}\cos(n-2)\phi + \dots + a_n,$$

$$Q = r^n\sin n\phi + a_1r^{n-1}\sin(n-1)\phi + a_2r^{n-2}\sin(n-2)\phi + \dots + a_{n-1}r\sin\phi.$$

A second expression for P and Q is obtained by letting $t = \tan 1/2\phi$. We obtain,

$$\cos\phi = \frac{1-t^2}{1+t^2}, \sin\phi = \frac{2t}{1+t^2}, z = r\frac{(1+it)^2}{1+t^2}$$

This gives,

$$(1+t^2)^n(P+iQ) = r^n(1+it)^{2n} + a_1r^{n-1}(1+it)^{2n-2}(1+t^2) + \dots + a_n(1+t^2)^n.$$

If we expand the binomials by the binomial formula, and arrange the result according to the powers of t, we get

$$P = \frac{g(t)}{(1+t^2)^n}, \quad Q = \frac{h(t)}{(1+t^2)^n},$$

where $g(t)$ and $h(t)$ are rational integral functions of t, the degrees of which do not exceed $2n$.

All points in the plane having the same value for r lie upon a circle of radius r, the center of which is at the origin of coordinates. To determine the points on this circle for which P and Q vanish, we must solve the equation $g(t) = 0$ and $h(t) = 0$, for the given value of r. But we know that if $h(t) = 0$ and $g(t) = 0$ have roots at all, they cannot have more than $2n$. From this it follows that neither P nor Q can be equal to zero at all points of an area in the plane, for in that event we could select r such that the circle would pass through that area, and P and Q would vanish at an infinite number of points on this circle.

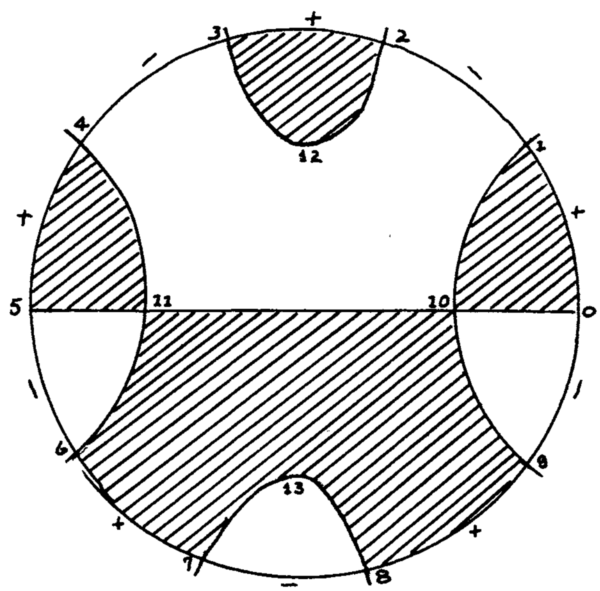
The value of Q may be written

$$Q = r^n(\sin n\phi + \frac{a_1}{r} \sin (n-1)\phi + \frac{a_2}{r^2} \sin (n-2)\phi + \dots).$$

From this expression it is readily seen that r may be taken so large that Q has the same sign as $\sin n\phi$ on all points of the circle where $\sin n\phi$ is numerically larger than some ε, which may be as small as we please, but not zero. Mark on the circle the points

$$0, \frac{\pi}{n}, \frac{2\pi}{n}, \dots, \frac{(2n-1)\pi}{n},$$

and designate them, respectively, by $0, 1, 2, \dots, 2n-1$. Thus the circle is divided into $2n$ arcs, (01), (12), (23), ..., $(2n-1, 0)$, in which $\sin n\phi$ is alternately $+$ and $-$. The figure shows the division for $n = 5$. In passing from arc (01) to arc (12), the function Q, for sufficiently large values of r, changes from $+$ to $-$. Since Q is a continuous function having real values, in going along the



circle from + to − , it must at the point 1 pass through zero. Similarly, Q must pass through zero also at the points 2,3,... , (2n−1), but it does this at no other points of the circle.

Similar remarks apply to P. It is readily seen that, for sufficiently large values of r, P and cos $n\phi$ have always equal signs; that P is positive at the points 0,2,...,(2n−2), and in their vicinity, and negative at the points 1,3,5,...,(2n−1), and in their vicinity.

We have seen that Q cannot vanish at all points of an area. Consequently, the area within the circle can be divided into districts so that in some districts Q is everywhere positive, while in others it is everywhere negative. These districts are marked off by boundary lines along which Q vanishes. To aid the eye, the positive districts are shaded.

An arc (2h,2h+1) of the circle, along which Q is positive, lies in a positive district. This district lies partly inside and partly outside the circle. Designate by I the part of it that is inside. Several cases may arise. The area I may terminate inside, as does (2,12,3), in which case (2h, 2h+1) is the only arc of the circle on its boundary. Or, the area I may run into another positive arc (2k, 2k+1), or it may divide into two or more branches, each of which terminates in a positive arc $(2\ell, 2\ell+1)$. If there could be within I an area, like an island, in which Q were negative, then the conclusions which we are about to draw would still follow.

Consider the boundary line within the circle, passing from $2h+1$ to $2k$. Along this line $Q = 0$. But P is negative at the point $2h+1$ and positive at the point $2k$. Since P is continuous and represents real values, P must pass through zero in at least one point along the boundary line connecting $2h+1$ and $2k$. Thus, at that point, we have not only $Q = 0$ but also $P = 0$; that is, $f(z) = P + iQ = 0$. Thus the existence of at least one root of $f(z) = 0$ is demonstrated.

A Note on Gauss' Fourth Proof

The theorem has been proved for the special case in which the coefficients of the given equation are all real. The general case, in which some or all of the coefficients are complex, easily follows, For, if $f_1(z)$ is a function of z, whose coefficients are, respectively, the conjugate imaginaries of the coefficients of a second function $f_2(z)$, then we may write $f_1(z) \equiv A + iB$ and $f_2(z) \equiv A - iB$, and $f_1(z)\, f_2(z) \equiv A^2 + B^2 = f(z)$, where $f(z)$ has only real coefficients. Now, if $f(z) = 0$ can be shown to have a root α_1, then we must have either $f_1(\alpha_1) = 0$ or $f_2(\alpha_1) = 0$. Suppose $f_1(\alpha_1) = 0$, then it follows that $f_2(\alpha_2) = 0$, where α_2 is the conjugate of α_1. Hence, $f_1(z) = 0$ and $f_2(z) = 0$ have each at least one root.

CHAPTER V

A MODERN PROOF OF THE FUNDAMENTAL THEOREM

The proofs of Gauss are indeed remarkable and are of considerable historical interest. However, as we mentioned in the first chapter, this theorem in its classic form is no longer regarded as belonging to algebra. The basic ideas of the modern treatment probably go back to Galois (1811-1832).

In 1851, Aquinas J. Liouville published the following famous theorem bearing his name: If $f(z)$ is analytic for all values of z, and $|f(z)|$ is bounded, then $f(z)$ is a constant. The proof of this theorem can be found in any modern textbook on theory of functions of a complex variable. Armed with this theorem, we can establish the Fundamental Theorem of Algebra in its classic form immediately.

Let $G(z) = a_0 z^n + a_1 z^{n-1} + \ldots + a_n$, $a_0 \neq 0$, $0 < n$, be an arbitrary polynomial. The Fundamental Theorem states that the equation:

$$G(z) = 0,$$

has a root.

To prove this, we will employ an indirect proof, using Liouville's theorem to arrive at an absurdity.

Suppose $G(z)$ had no root. Form the function

$$F(z) = \frac{1}{G(z)} .$$

Then $F(z)$ is analytic in the z plane. To show that it is bounded, consider its value on a circle,

$$z = R e^{\varphi i}.$$

Since, for all such points,

$$G(z) = z^n \left[a_o + \frac{a_1}{z} + \ldots + \frac{a_n}{z^n} \right]$$

it follows that

$$\left| G(z) \right| \geq |z|^n \left[|a_o| - \frac{|a_1|}{|z|} - \ldots - \frac{|a_n|}{|z|^n} \right].$$

Now, choose R so that

$$\frac{|a_1|}{R} + \frac{|a_2|}{R^2} + \ldots + \frac{|a_n|}{R^n} < \frac{1}{2} \left| a_o \right|$$

Then

$$\left| G(z) \right| \geq \frac{1}{2} |a_o| R^n, \quad R \leq |z|.$$

Therefore,

$$\left| F(z) \right| \leq \frac{2}{|a_o| R^n}, \quad R \leq |z|,$$

and so F(z) is bounded. By Liouville's Theorem, F(z) is a constant. But this is absurd. Therefore G(z) has a root.

SELECTED BIBLIOGRAPHY

1. Adams, C. Raymond. "Second Proof of the Fundamental Theorem of Algebra", A Source Book in Mathematics. New York and London: McGraw-Hill Book Company, 1929, pp. 292-306.

2. Ball, W. W. A Short Account of the History of Mathematics. London: MacMillan and Company, 1922.

3. Bell, E. T. Development of Mathematics. New York: McGraw-Hill Book Company, 1940.

4. Bell, G. T. Men of Mathematics. New York: Simon Schuster, 1937.

5. Bocher, Maxime. "Gauss' Third Proof", American Mathematical Society Bulletin (May, 1895) p. 205.

6. Cajori, Florian. Theory of Equations. London: MacMillan and Company, 1924, p. 26.

7. Heath, T. L. Diophantus of Alexander. Cambridge: Cambridge University Press, 1910, p. 58.

8. van der Waerden, B. L. Modern Algebra. New York: Unger Publishing Company, 1949, p.224.

VITA

Richard Moore Lotspeich

Candidate for the Degree of

Master of Science

Report: THE FUNDAMENTAL THEOREM OF ALGEBRA - THE PROOFS OF GAUSS WITH NOTES

Major Field: Mathematics

Biographical:

Personal: Born in Nashville, Tennessee, December 15, 1925, the son of Edgar S. and Ruth M. Lotspeich.

Education: Attended elementary school in Louisville, Kentucky; graduated from Fortier High School, New Orleans, Louisiana in 1943; received the Bachelor of Arts degree from Union College, Barbourville, Kentucky, with a major in Mathematics, in May, 1948; received the Bachelor of Science in Education degree from the same institution in May, 1951; completed requirements for the Master of Science degree in May, 1959.

Professional experience: Served in the United States Navy from 1944 to 1946; taught mathematics at Union College, Barbourville, Kentucky, 1948-1949; taught mathematics and physical science at Charleston High School, Charleston, Missouri, 1951-1958; is presently attending Oklahoma State University as a National Science Foundation fellow.