IC

Copy

Robbin Ernest

From: Document Delivery Services <documentdelivery@duke.edu>
Sent: Wednesday, August 24, 2011 9:19 AM
To: LSC-Requests
Subject: Lending Copy, TN #: 660859

TRAN: 660859 COPY

D01891539-

Journal: Eureka.

Year: 1985, ()
Vol: 45, ()
Pgs: 42-47

ARTICLE
Title: Paul Taylor; Gauss' Second Proof

Author:

Patron: dept; status; Arasu, Krishnasamy
ILL Number: 81220704
MaxCost: 20.00IFM

Copied for:
WSU - Dunbar Library - ILL
Wright State University
Dayton, OH 45435

Fax: (937)775-2356
Ariel: 130.108.121.59

Lending String: YUS,YUS,KUK,MYG,*NDD

NDD Contact Information
Phone: 919-660-5891
Fax: 919-660-5964
Email: illrequests@duke.edu
Ariel: ariel.lib.duke.edu

Gauss' Second Proof

Paul Taylor

The fact that any non-trivial polynomial equation has a root in the field of complex numbers will be familiar to all readers of *Eureka*, just as it became familiar to Mathematicians at some stage between the end of the Dark Ages (ie the publication of the *Ars Magna* in 1540) and the mathematical birth of Gauss in 1796. The best known proof of it occurs in Part IB Complex Variables: for sufficiently large values of the indeterminate, the highest degree term (z^n) dominates and by Rouché's theorem the polynomial has the same number of zeroes as this single term, viz. n. This is not the proof that interests us, however: we want a purely *algebraic* proof for real closed fields. (See the previous article.)

Later I shall give the modern two-line proof of this (due to Artin), using Galois' and Sylow's theorems, which must somehow be a descendant of the one given here. However the connection is difficult to spot, and the precise genealogy could easily form the subject of a PhD thesis in the history and philosophy of science, or indeed the basis of a book on the history of modern algebra since the lemmas I shall state without comment hint at the foundations of most branches of the subject. I should be delighted to hear from any reader who can throw light on some of the connections.

This article is a précis of my translation of Gauss' paper [2]; I should like to thank Bernard Leak for his corrections to my rusty Latin.

Proposition 1 (*Euclid's algorithm for polynomials*) Given two polynomials $Y(x)$ and $Y'(x)$, there are polynomials $Z(x)$ and $Z'(x)$ such that $ZY+Z'Y'$ is the greatest common divisor of Y, Y'.

Corollary Y, Y' have no common factor iff $ZY+Z'Y'=1$ for some Z, Z'.

Proposition 2 Any symmetric polynomial in n variables $a,b,\dots$ may be expressed uniquely as the result of substituting $\sigma_1=\Sigma a$, $\sigma_2=\Sigma ab$, ..., $\sigma_n=\Pi a$ for $s_1,\dots,s_n$ in some polynomial in $s_1,\dots,s_n$.

This result is frequently proved using independent transcendentals in a Galois Theory course; Gauss gave an easier, prettier, constructive and more convincing proof which the reader is invited to find for itself. Gauss' paper consists largely of repeated use of this result, and employs the convention that

Y is the polynomial $x^m - S_{m-1}x^{m-1} + \ldots \pm S_0$ with particular determined coefficients and (possibly) roots A, B, ...,

y is the polynomial $x^m - s_{m-1}x^{m-1} + \ldots \pm s_0$ with indeterminate coefficients and no roots, and

v is the polynomial $(x-a)(x-b)(x-c)\ldots$ with coefficients $\sigma_1, \ldots \sigma_n$ and roots $a, b, c \ldots$

Other corresponding upper case Latin, lower case Latin and lower case Greek letters are used similarly.

The *discriminant* of a polynomial v is the product $\pi = (a-b)(a-c)\ldots(b-a)(b-c)\ldots$ and correspondingly those of Y and y are P and p respectively. Y', y' and v' are used for the formal derivatives of Y, y and v with respect to x.

Proposition 3 Y and Y' have a common factor iff P=0.

Gauss takes the opportunity here of being rude to his inferior contemporaries. Of course the result is obvious for v, ie if the polynomial factorises, but that is begging the question (*petitio principii*). He spends several pages on the proof (bringing various clever formulae out of a hat), but the result is easy for us with abstract algebra (we do not suffer from the handicap of *believing* in the real numbers) since it's a triviality to construct field extensions containing the required roots. We need only that Y and Y' factorise in *some* field containing **R**, not necessarily **C** itself.

Given a polynomial (say of degree $2^\mu k$ with k odd) with zero discriminant (which is really trying to say that it has a repeated root), we may extract from it the common factor with its derivative. Continuing this process inductively, any polynomial may be split (rationally) into factors each of which has a nonzero discriminant. Moreover the degree of at least one of these factors will be $2^\nu l$ with l odd and $\nu \le \mu$. The proof will proceed by induction on the multiplicity of 2 as a prime factor of the degree, whilst the degree itself will increase astronomically.

Gauss now introduces a little elementary group theory (in the form of a discussion of symmetric polynomials) before proceeding in his usual fashion with a brilliant unmotivated proof using another indeterminate u. Subscripts $_u$ and $_x$ will be used for partial derivatives.

Let ζ denote the product of all $u-(a+b)x+ab$ excluding repetitions, and Z, z the polynomials corresponding in the usual way. By similar methods as before he proves

Proposition 4 If P≠0 then the discriminant of Z with respect to u cannot be identically zero as a function of x.

For those who are at this point wondering where we are going, I have to tell you that there are many more complicated calculations still to do. Moreover it is not until the very last section that the real-closed hypotheses make their appearance.

Lemma Let $\phi(u,x)$ denote a product of any number of factors, into each of which the variables u,x enter only linearly, ie which are of the form $\alpha+\beta u+\gamma x$, $\alpha'+\beta' u+\gamma' x$, $\alpha''+\beta'' u+\gamma'' x$, ..., and w be another variable. Then the polynomial $\Omega = \phi(u+w\phi_x(u,x),\ x-w\phi_u(u,x))$ is divisible by $\phi(u,x)$.

This lemma is clearly applicable to the polynomial ζ, which we may write as $f(u,x,\sigma_1,\sigma_2,...)$ so that

$$f(u+w\ \frac{\partial\zeta}{\partial x},\ x-w\frac{\partial\zeta}{\partial u},\ \sigma_1,\sigma_2,...)$$

is exactly divisible by ζ: the quotient, which will be a polynomial in $u,x,w,a,b,c,...$ symmetric in $a,b,c,...$, we may write as $\psi(u,x,a,b,c,...)$. Hence

$$f(u+w\frac{\partial z}{\partial x},\ x-w\frac{\partial z}{\partial u},\ s_1,s_2,\ ...) \;=\; z\ \psi(u,x,w,s_1,s_2,...)$$

and

$$f(u+w\frac{\partial Z}{\partial x},\ x-w\frac{\partial Z}{\partial u},\ S_1,S_2,...) \;=\; Z\ \psi(u,x,w,S_1,S_2,...)$$

Then we may more simply write the polynomial Z as $F(u,x)$ so that

$$F(u+w\frac{\partial Z}{\partial x},\ x-w\frac{\partial Z}{\partial u}) \;=\; F(u,x)\ \psi(u,x,w,S_1,S_2,...)$$

If we put $u=U$, $x=X$, so that, say, $\partial Z/\partial x=X'$, $\partial Z/\partial u=U'$, then we shall have

$$F(U+wX',\ X-wU') \;=\; F(U,X)\ \psi(U,X,w,S_1,S_2,...)$$

Then as long as U' doesnt vanish, we may set $w=(X-x)/U'$ to get

$$F(U+\frac{X\ X'}{U'} - \frac{x\ X'}{U'},\ x) \;=\; F(U,X)\ \psi(U,X,\frac{X-x}{U'},S_1,S_2,...)$$

Hence if in Z we put $u = U + XX'/U' - xX'/U'$ it becomes

$$F(U,X)\ \psi(U,X,\frac{X-x}{U'},S_1,S_2,...)$$

When, in the case that $P\neq 0$, the discriminant (wrt u) of the polynomial $Z=F(u,x)$ is a nonvanishing function of x, clearly the number of definite values of x for which the discriminant of Z can be 0 will be finite, so that there are infinitely many choices for x giving a nonvanishing discriminant. Let X be such a (real) value, so that the discriminant of the polynomial $F(u,X)$ will be nonzero and so $F(u,X)$ and $F_u(u,X)$ have no common factor. Now let us suppose that there is some definite (complex) value, say U, of u satisfying $F(u,X)=0$, ie such that $F(U,X)=0$, so that $(u-U)$ will be a factor of the polynomial $F(u,X)$ and not of $F_u(u,X)$. Let the latter take the value U' for $u=U$, so $U'\neq 0$. Let X' be the value of $F_x(u,x)$ for $u=U$, $x=X$. Then by the above result Z will vanish identically by the substitution $u = U + XX'/U' - X'x/U'$ and so Z is divisible by the factor $u + X'x/U' - (U + XX'/U')$.

Hence $F(x^2,x)$ is divisible by $x^2 + xX'/U' - (U + XX'/U')$ and so has roots

$$\frac{-X' \pm\sqrt{4UU'^2 + 4XX'U' + X'^2}}{2U'}$$

in C. Moreover *it may easily be shown that* for the same values of x the polynomial Y must also vanish; for clearly $f(x^2,x,\sigma_1,\sigma_2,...)$ is the product of all $(x-a)(x-b)$ excluding repetitions and so equal to v^{m-1}. Hence it immediately follows that $F(x^2,x)=Y^{m-1}$, which cannot vanish unless Y itself vanishes.

With the help of the preceding discussion, the solution of the equation Y=0 of degree n (where the discriminant of Y is nonzero) is reduced to that of $F(u,X)=0$. It is appropriate to observe that if all of the coefficients in Y are real quantities then so are those in $F(u,X)$, since it is possible to make X real. The degree of the secondary equation $F(u,X)=0$ is $n(n-1)/2$, in which 2 occurs as a prime factor once less often than in n (assuming n even).

If the discriminant of Y is zero, then as remarked above it may be split into factors whose discriminants do not vanish, and it suffices to find a root of any of these.

We thus obtain a sequence of polynomial equations of degrees $2^{\mu_0}k_0$, $2^{\mu_1}k_1,...$ with $\mu_0>\mu_1>\mu_2>...$, and hence ultimately one of odd degree which may be solved by hypothesis (or, *obviously* in **R**). Moreover any solution to the last yields one for the previous ones and hence the original equation, as required.

Now let us attempt to understand how this proof works in modern terms. First, one may easily show from the Euclidean algorithm in the same way as for Z that

Proposition 5 Any polynomial in one variable over an arbitrary field factorises into irreducibles, uniquely up to permutation and multiplication by nonzero scalars.

Of course Gauss will have been well aware of this, and it is a little surprising thet he doesnt use the word irreducible. The result may now be restated as follows:

Theorem 6 (Gauss) A non-trivial irreducible polynomial over **R** is a scalar multiple of $x^2 + 2ax + (a^2+b^2)$ and hence factorises over C.

At this point I shall make use of our 170-year head start on the master, which is the means by which I have already reduced propositions 3 and 4 to trivialities and hence cut down Gauss' paper by three quarters. Let k be any field and $f(x)$ an irreducible polynomial over k. Let α be a new indeterminate and denote by $k(\alpha)$ the vector space of polynomials in α over k of degree less than that of f; this has a multiplicative structure given by setting multiples of $f(\alpha)$ to zero.

Proposition 7 $k(\alpha)$ is a field containing k in which $f(x)=0$ has a root α. Moreover if L is another field containing k and a root β then there's a unique embedding of $k(\alpha)$ into L which preserves k and identifies α with β.

The reader is invited to formulate and prove the uniqueness of $k(\alpha)$. It is an easy matter to show (in the same sense) that

Proposition 8 Let k be any field and $f(x)$ any polynomial over it. Then there is a smallest field K containing k in which f splits into linear factors.

K is called the *splitting field* of f over k; a field extension (ie an inclusion of one field in another) is said to be *normal* if it is of this form. A field extension is normal iff every polynomial which is irreducible over the smaller field but has a root in the larger actually splits in the larger (*"one out - all out"*).

Suppose then we have a non-trivial irreducible polynomial $Y(x)$ over **R** with splitting field K; we aim to show that K≅**C**, ie the dimension of K as a real vector space (which is called the *degree* of the extension K:R) is 2.

Now consider Gauss' secondary polynomial $F(u,X)$. This clearly splits in K, so its splitting field L is contained in K. On the other hand the roots of the original polynomial are obtained from those of $F(u,X)$ by solving some quadratics, which is the same as saying that K is obtained by adjoining some square roots to L. Hence, whilst the secondary equation may have much larger degree, it is in some sense no more difficult to solve, and indeed possibly easier.

Repeating Gauss' construction, we obtain a descending sequence of field extensions contained in K:R which is such that each is obtained from the next by adjoining square roots and the last is the splitting field for an odd-degree polynomial.

Gauss' construction is a little stronger than this. In order to construct K we do not need the splitting field of the whole of the secondary polynomial $F(u,X)$, but only of a non-trivial irreducible factor. This is because Gauss only asks for a single root of the secondary polynomial in order to get a root of the orininal one, and since K is normal this is all we need. Hence at the last stage it is sufficient to consider the linear factor which an odd-degree polynomial is guaranteed to have; the splitting field of this is of course just R.

Hence K, the splitting field of the original polynomial, is obtained by adjoining square roots to R itself. But we can only do this once because the existence of square roots in **C** is an easy exercise. Hence K is indeed just **C**.

Now I shall give Artin's proof, quoting major theorems from two Part II courses; the word *separable* is included for purely legal reasons, the condition being automatic for fields containing **Q**.

Theorem 9 (Galois) [4] Let K:k be a normal separable extension of finite degree and let G be the group of field automorphisms of K which fix each element of k. Then there is an order reversing bijection beteen the subgroups H of G and the fields L lying between k and K; moreover the degree of L:k is equal to the index G:H.

Theorem 10 (Sylow) [3] Let G be a finite group of order $p^a m$ where p is a prime not dividing m. Then G has subgroups of order p^b for each $0 \leq b \leq a$.

Applying this to the case in hand with $p=2$, K:R being the splitting field of an irreducible polynomial Y, there is a field L lying between R and K such that the degrees of K:L and L:R are respectively a power of 2 and odd. L must be obtained from R by adjoining roots of irreducible odd-degree polynomials (which is impossible) and K from L by solving quadratics. Hence L=R and K=L(i)=R(i)=C.

Two problems I shall leave to the reader are spotting the theorem of the primitive element and the proof of Sylow's theorem for $p=2$ (I believe one can generalise Gauss' method to a proof of Sylow's theorem in general). Of course Gauss does not prove Galois' theorem because in Artin's proof this serves merely to translate Sylow's theorem from (permutation) groups to fields (and hence polynomials), whereas if one speaks the local language fluently oneself one does not need an interpreter.

References

The first of these is an addendum to my previous article [5], kindly supplied by Walter Ledermann.

[1] Frobenius, G., Über lineare Substitutionen und bilineare Formen, *Journal für die reine und angewendte Mathematik* (Crelle's Journal) **84** (1878) 1-63

[2] Gauss, C.F., Demonstratio nova altera theorematis omnem functionem algebraicam rationalem unius variabilis in factores reales primi vel secundi gradûs resolvi posse. *Gauss Werke* **3** 33-56

[3] Johnstone, P.T., *Algebra I example sheet 3*, DPMMS, 1982

[4] Stewart, I.N., *Galois Theory* Chapman & Hall

[5] Taylor, P., The Uniqueness of the Quaternions, *Eureka* **44** (1984) 63-68