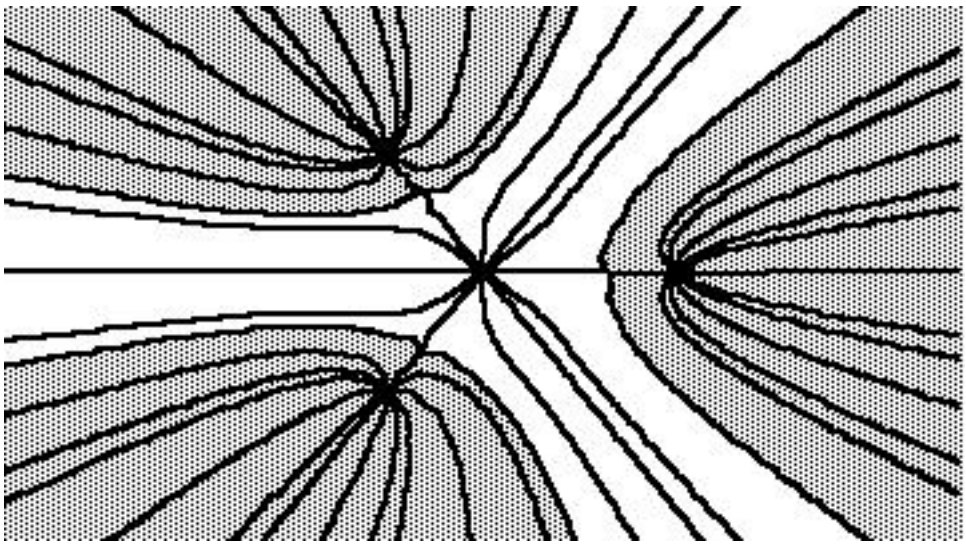
Theory of Equations

Lesson 10

by

Barry H. Dayton
Northeastern Illinois University
Chicago, IL 60625, USA

www.neiu.edu/~bhdayton/theq/



4.6 Newton's Identities

While it was noticed early that the roots of a polynomial are not rational expressions of the coefficients, the mathematician Girard noticed in 1629 that certain expressions in the roots were rational expressions of the coefficients.

Suppose we consider the polynomial equation

$$t^3 + p_1 t^2 + p_2 t + p_3 = 0$$

Let x_1, x_2, x_3 be the roots. Then Girard noted that

$$\begin{aligned} x_1 + x_2 + x_3 &= -p_1 \\ x_1^2 + x_2^2 + x_3^2 &= p_1^2 - 2p_2 \\ x_1^3 + x_2^3 + x_3^3 &= -p_1^3 + 3p_1 p_2 - 3p_3 \\ x_1^4 + x_2^4 + x_3^4 &= p_1^4 - 4p_1^2 p_2 + 4p_1 p_3 + 2p_2^2 \end{aligned}$$

While it seems clear that we can continue, it was not clear to Girard what the general pattern was. In fact, the general pattern is quite complicated but Isaac Newton published in 1683 a simple set of recursive equations. Note that for convenience we are writing the coefficients of the polynomial in a non-standard order.

Theorem 4.6.1 (Newton's Identities) *Let*

$$f(t) = t^n + p_1 t^{n-1} + p_2 t^{n-2} + \cdots + p_{n-1} t + p_n$$

be a polynomial with roots (counted according to multiplicity) $x_1, x_2, \ldots, x_n$*. For* $j = 1, 2, 3, \ldots$ *let*

$$s_j = x_1^j + x_2^j + \cdots + x_n^j$$

Set $p_k = 0$ *for* $k > n$*. Then for all* $j > 0$

$$s_j + p_1 s_{j-1} + p_2 s_{j-2} + \cdots + p_{j-1} s_1 + j p_j = 0$$

What this theorem says is that we have the equations

$$\begin{aligned} s_1 + p_1 &= 0 \\ s_2 + p_1 s_1 + 2p_2 &= 0 \\ s_3 + p_1 s_2 + p_2 s_1 + 3p_3 &= 0 \\ s_4 + p_1 s_3 + p_2 s_2 + p_3 s_1 + 4p_4 &= 0 \end{aligned}$$

The first equation allows us to solve for s_1 in terms of p_1. Then since we know s_1 and the p_k the second equation allows us to solve for s_2. Now we know s_1, s_2 and the p_k so we can solve the third equation for s_3. We can continue in this manner to find s_4, s_5 and so on for as long as we like. The reader should check that solving the four equations gives Girard's formulas above. Newton did not bother to supply a proof for these identities. We will sketch a proof modified from *Uspensky*. We will work with *formal power series*, i.e. expressions of the form $\sum_{i=0}^{\infty} a_i x^i$. By *formal* we mean that we won't worry about convergence, we will only do algebraic operations. The operations of sum and product for formal power series are as defined for polynomials in Chapter 1 and we obtain an integral domain. The main surprise is that formal power series with non-zero constant term have multiplicative inverses. In particular we have the *geometric series*

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 \cdots$$

Proof: We start with the factorization

$$f(t) = t^n + p_1 t^{n-1} + \cdots + p_n = (t - x_1)(t - x_2) \cdots (t - x_n)$$

We then take the *logarithmic derivative*

$$\frac{f'(t)}{f(t)} = \frac{1}{t - x_1} + \frac{1}{t - x_2} + \cdots \frac{1}{t - x_n}$$

Multiplying by t gives

$$t\frac{f'(t)}{f(t)} = \frac{t}{t - x_1} + \frac{t}{t - x_2} + \cdots \frac{t}{t - x_n} \tag{4.10}$$

Now if we expand each term on the right using the geometric series we have

$$\frac{t}{t - x_i} = \frac{1}{1 - \frac{x_i}{t}} = 1 + \frac{x_i}{t} + \frac{x_i^2}{t^2} + \frac{x_i^3}{t^3} + \cdots$$

Adding, we see that the right hand side of 4.10 is

$$n + (x_1 + \cdots + x_n)\frac{1}{t} + (x_1^2 + \cdots + x_n^2)\frac{1}{t^2} + (x_1^3 + \cdots + x_n^3)\frac{1}{t^3} + \cdots$$

or, alternatively (writing $s_0 = n$)

$$t\frac{f'(t)}{f(t)} = s_0 + \frac{s_1}{t} + \frac{s_2}{t^2} + \frac{s_3}{t^3} + \cdots \tag{4.11}$$

Multiplying 4.10 by $f(t)$ then gives

$$tf'(t) = (t^n + p_1 t^{n-1} + \cdots + p_n)(s_0 + \frac{s_1}{t} + \frac{s_2}{t^2} + \cdots) \tag{4.12}$$

Multiplying out the right hand side of 4.12 gives a power series

$$b_n t^n + b_{n-1} t^{n-1} + \cdots b_0 + b_{-1} t^{-1} + \cdots$$

where

$$b_{n-j} = s_j + p_1 s_{j-1} + \cdots + p_j s_0$$

On the other hand,

$$tf'(t) = nt^n + (n-1)p_1 t^{n-1} + \cdots + p_{n-1} t$$

This says that the coefficient of t^{n-j} on the left is $(n-j)p_j$ which makes sense for all $j > 0$ since we set $p_j = 0$ for $j > n$.

Finally (since we are using formal power series we may equate the coefficients of t^{n-j} on both sides of 4.12 to get

$$(n-j)p_j = s_j + p_1 s_{j-1} + \cdots + p_j s_0$$

Subtracting $(n-j)p_j$ from both sides of this equation and using the fact that $s_0 = n$ gives Newton's identity.

We give an application of the use of Newton's Identities. This is the application that Newton had in mind and gives a method for finding the largest real root of a polynomial (assuming such a root exists, is not a multiple root, and is actually the root of largest modulus). We note that this method is of no practical value today.

The method is based on the fact that if the real root x_n is is of larger modulus than any other root then for large k, $s_k = x_1^k + \cdots + x_n^k \approx x_n^k$. Thus the sequence $s_1, \sqrt{s_2}, \sqrt[3]{s_3}, \sqrt[4]{s_4}, \ldots$ should converge to x_n.

Example 4.6.2 Let $f(x) = x^3 - 5x^2 + 6x - 1$. Then $p_1 = -5$, $p_2 = 6$ and $p_3 = -1$. From the identities we get

$$\begin{aligned}
s_1 &= -p_1 = 5 \\
s_2 &= -p_1 s_1 - 2p_2 = -(-5)(5) - 2(6) = 13 \\
s_3 &= -p_1 s_2 - p_2 s_1 - 3p_3 = -(-5)(13) - (6)(5) - 3(-1) = 38 \\
s_4 &= -p_1 s_3 - p_2 s_2 - p_3 s_1 = -(-5)(38) - (6)(13) - (-1)(5) = 117
\end{aligned}$$

Thus our sequence is $s_1 = 5, \sqrt{s_2} = 3.6055, \sqrt[3]{s_3} = 3.36197, \sqrt[4]{s_4} = 3.2888, \ldots$ The actual root is 3.2469

Maple Implementation

For actually calculating the ghost coefficients s_j with Maple it is easier to use equation 4.11 directly than Newton's identities. The idea is to expand the left hand side of this equation in an *asymptotic* series, that is a series in negative powers of t. Thus to find the s_j in the above example one might do

```
f:= t^3 - 5*t^2 + 6*t-1;
asympt(t*diff(f,t)/f, t, 5);
```

and the result would look like

$$3 + \frac{5}{t} + \frac{13}{t^2} + \frac{38}{t^3} + \frac{117}{t^4} + O(\frac{1}{t^5})$$

where $s_0, s_1, \ldots, s_4$ are the numerators and the O-term at the end is to remind you that this is an infinite series. Replacing the 5 with a larger number will get you as many terms as you like.

Exercise 23a [10 points] Use the method just described to find the largest real root of $f(x) = x^4 - 2x^3 - 5x^2 + 6x + 3$ correct to 2 significant digits.

4.7 More on Newton's Identities

In this optional section we consider some more applications of Newton's identities. The material in this section will not be needed in the sequel, so you may wish to skip this section to maintain the continuity of our story.

For our next application we note that if we know the s_k we can then calculate the coefficients p_j of our polynomial. For solving Newton's identities for p_j in terms of the s_k we have

$$\begin{aligned} p_1 &= -s_1 \\ p_2 &= \frac{s_1^2 - s_2}{2} \\ p_3 &= \frac{-2s_3 + 2s_1s_2 + s_1^3 - s_1s_2}{6} \end{aligned}$$

and so on. Evidently these formulas get very complicated quickly, but numerically it is easy to solve for the p_j's directly from Newton's Identities.

This technique has been used to calculate the eigenvalues of a matrix. Recall that the eigenvalues of an $n \times n$ matrix A are the roots of the characteristic polynomial $f(\lambda) = \det(\lambda I - A) = \lambda^n + p_1\lambda^{n-1} + p_2\lambda^{n-2} + \cdots + p_{n-1}\lambda + p_n$. (Note that we are using $\det(\lambda I - A)$ rather than the formula $\det(A - \lambda I)$ found in some linear algebra books in order that our polynomial be monic, these formulas differ by a factor of $(-1)^n$.) If the eigenvalues (complex and counted according to multiplicities) are $x_1, x_2, \ldots, x_n$ then it is not too difficult to see that the sum $s_1 = x_1 + \cdots + x_n$ is the sum of the diagonal entries of A, in fact both are equal to $-p_1$. The sum of the diagonal entries of a matrix A is called the *trace* of the matrix, trace(A).

It is a bit harder to see that the eigenvalues of A^2 are $x_1^2, x_2^2, \ldots, x_n^2$. When there are no repeated eigenvalues this follows from the fact that λ is an eigenvalue of A if there exists a vector $v \neq 0$ so that $Av = \lambda v$. Then $A^2v = A(Av) = A(\lambda v) = \lambda Av = \lambda^2 v$ so λ^2 is an eigenvalue of A^2. It follows that the trace of A^2 is then $x_1^2 + x_2^2 + \cdots + x_n^2 = s_2$. More generally the argument above suggests that $s_k = x_1^k + \cdots + x_n^k$ is the trace of A^k for each $k > 0$. In fact this is true. Thus the method is to calculate A^k for $k = 1, 2, \ldots, n$ and set s_k to be the trace of A^k. Then working backwards using Newton's identities we can find the coefficients $p_1, p_2, \ldots, p_n$ of the characteristic polynomial.

Example 4.7.1 We wish to find the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 3 & -2 & 1 \\ 2 & 1 & 4 \end{bmatrix}$$

We first calculate

$$A^2 = \begin{bmatrix} 7 & -2 & 2 \\ -1 & 11 & 2 \\ 13 & 6 & 17 \end{bmatrix}, \quad A^3 = \begin{bmatrix} 5 & 20 & 6 \\ 36 & -22 & 19 \\ 65 & 31 & 74 \end{bmatrix}$$

We then see that $s_1 = \text{trace}(A) = 3$, $s_2 = \text{trace}(A^2) = 35$ and $s_3 = \text{trace}(A^3) = 57$. Now from the identity

$$s_1 + p_1 = 0$$

we calculate $p_1 = -3$. From the identity

$$s_2 + p_1 s_1 + 2p_2 = 0$$

we have $35 + (-3)(3) + 2p_2 = 0$ so $p_2 = -13$. Finally from

$$s_3 + p_1 s_2 + p_2 s_1 + 3p_3 = 0$$

we obtain $57 + (-3)(35) + (-13)(3) - 3p_3 = 0$ so $p_3 = 29$. We thus conclude that the characteristic polynomial of A is

$$f(\lambda) = \lambda^3 - 3\lambda^2 - 13\lambda + 29$$

Using any method from Chapter 2 we see that the roots of $f(\lambda)$ are -3.3813, 1.9244 and 4.4569, i.e. these are the eigenvalues of A.

Exercise 23b [20 points] Find the characteristic polynomial and the eigenvalues of the matrix

$$A = \begin{bmatrix} 1 & 2 & 0 & -1 \\ 2 & 1 & -2 & 0 \\ 0 & -2 & 3 & 3 \\ -1 & 0 & 3 & 1 \end{bmatrix}$$

Hint: The eigenvalues are real.

There are similar identities for finding sums of negative powers of the roots of a polynomial.

Theorem 4.7.2 *Let $f(x) = p_0t^n + p_1t^{n-1} + \cdots + p_n$ be a polynomial with roots (counted according to multiplicity) $x_1, x_2, \ldots, x_n$. Assume that $p_n \neq 0$, i.e. that no $x_i = 0$ and set $p_k = 0$ for $k < 0$. Define for $k \geq 0$ $s_{-k} = x_1^{-k} + x_2^{-k} + \cdots + x_n^{-k} = \frac{1}{x_1^k} + \cdots + \frac{1}{x_n^k}$. Then for all $j \leq n$*

$$(n-j)p_j + p_{j+1}s_{-1} + \cdots + p_n s_{j-n} = 0$$

The proof is similar to that of Theorem 4.6.1 by noting that expansion as a formal power series in positive powers of t gives

$$-t\frac{f'(t)}{f(t)} = s_{-1}t + s_{-2}t^2 + s_{-3}t^3 + \cdots \tag{4.13}$$

Maple Implementation

Again the most efficient way to calculate the ghost coefficients s_{-j} is to expand the left hand side of equation 4.13 as a series, eg.

```
series(-t*diff(f,t)/f,t,8);
```

would give s_{-j} as the coefficient of t^j for $j = 1, 2, \ldots, 7$.

Often in the literature the Newton's identity for positive and negative powers is combined by adding the two identities as follows:

Corollary 4.7.3 *With hypotheses as in the previous theorem, setting $s_0 = n$ we have for all integers (positive, negative and zero) j*

$$p_0 s_j + p_1 s_{j-1} + p_2 s_{j-2} + \cdots + p_n s_{j-n} = 0$$

Actually, the Newton's Identity for negative powers looks nicer if we use our more standard notation for polynomials:

Corollary 4.7.4 *Let $f(x) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$ be a polynomial of degree n with $a_0 \neq 0$. Let $x_1, \ldots, x_n$ be the roots counted according to multiplicity and let $s_{-k} = x_1^{-k} + \cdots + x_n^{-k}$ for all $k > 0$. Then for all $j > 0$*

$$j a_j + a_{j-1} s_{-1} + a_{j-2} s_{-2} + \cdots + a_0 s_{-j} = 0$$

It should be noted from this result that alghough the number of roots and the actual roots depends on the degree and all the coeficients, the *sum* of the j^{th} powers of the reciprocals of the roots depends only on the coefficients $a_0, a_1, \ldots, a_j$ when $j < n$ and is independent of the degree. This observation motivated the mathematician Euler to apply this last corollary to power series. Most modern mathematicians will say that Euler's argument is wrong, but his results are correct.

Euler wanted to calculate the number

$$\zeta(n) = \sum_{k=1}^{\infty} \frac{1}{k^n}$$

for n a positive even integer. To this end he started with the series

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$$

and noted that $\sin(x)$ has roots $k\pi$ for all integers k. Dividing by x eliminates the root at 0 and replacing x by $\sqrt{x}$ eliminates the negative roots according to Euler. Thus Euler argued that

$$f(x) = \frac{\sin(\sqrt{x})}{\sqrt{x}} = 1 - \frac{x}{3!} + \frac{x^2}{5!} - \frac{x^3}{7!} + \cdots$$

has roots at $(k\pi)^2$ for $k = 1, 2, 3, \ldots$. This is certainly correct, however (here is the questionable step!) Euler then set

$$s_{-j} = \sum_{k=1}^{\infty} \frac{1}{((k\pi)^2)^j} = \frac{1}{\pi^{2j}} \sum_{k=1}^{\infty} \frac{1}{k^{2j}} \tag{4.14}$$

and calculated the s_{-j} using the Newton's Identities of the last corollary. For $j = 1$ he had $1a_1 + a_0 s_{-1} = 0$ so $s_{-1} = -a_1 = \frac{1}{3!} = 1/6$. Multiplying equation (4.14) by π^2 gives

$$\zeta(2) = \sum_{k=1}^{\infty} \frac{1}{k^2} = \pi^2 s_{-1} = \frac{\pi^2}{6}$$

Next using $2a_2 + a_1 s_{-1} + a_0 s_{-2} = 0$ gives $s_{-2} = (\frac{1}{3!})^2 - \frac{2}{5!} = \frac{1}{90}$. Multiplying (4.14) by π^4 gives

$$\zeta(4) = \sum_{k=1}^{\infty} \frac{1}{k^4} = \pi^4 s_{-2} = \frac{\pi^4}{90}$$

It should be mentioned that there is still no simple formula for odd values of ζ, for example $\zeta(3) = \sum_{k=1}^{\infty} \frac{1}{k^3}$ is known only numerically and only in the last few years has even been shown to be irrational.

Exercise 23c [10 points] Assuming that Euler's calculation is correct, find $\zeta(6)$ and $\zeta(8)$.

4.8 Symmetric Polynomials

The expressions $s_k = x_1^k + \cdots + x_n^k$ are examples of *symmetric polynomials*. Symmetric polynomials played a large role in the development of the modern theory of solvability of polynomials. The 18th century mathmatician Edward Waring is often credited for much of the development of the theory of symmetric polynomials.

More generally we start with a polynomial $f(x_1, x_2, \ldots, x_n)$ in n **variables** x_j. This means that $f(x_1, x_2, \ldots, x_n)$ is a sum of terms each one is a constant times positive powers of some or all of the variables x_j. $f(x_1, x_2, \ldots, x_n)$ is *symmetric* if any permutation of the variables leaves the result unchanged. More precisely, $f(x_1, x_2, \ldots, x_n)$ is symmetric if for all $1 \leq j < k \leq n$ $f(\ldots, x_j, \ldots, x_k, \ldots) = f(\ldots, x_k, \ldots, x_j, \ldots)$. For example

$$f(x_1, x_2, x_3) = x_1^2 x_2 x_3 + x_1 x_2^2 x_3 + x_1 x_2 x_3^2$$

is a symmetric polynomial but

$$g(x_1, x_2, x_3) = x_1 x_2^2 x_3^3$$

is not for $g(x_2, x_1, x_3) = x_1^2 x_2 x_3^3 \neq g(x_1, x_2, x_3)$.

Consider a polynomial $f(t) = t^3 + p_1 t^2 + p_2 t + p^3$ of degree three with roots x_1, x_2, x_3. Then $f(t)$ factors as $f(t) = (t - x_1)(t - x_2)(t - x_3)$. Multiplying this last expression back out we get $f(t) = t^3 - (x_1 + x_2 + x_3)t^2 + (x_1x_2 + x_1x_3 + x_2x_3)t - x_1x_2x_3$. Thus we conclude that

$$\begin{aligned} p_1 &= -(x_1 + x_2 + x_3) \\ p_2 &= x_1x_2 + x_1x_3 + x_2x_3 \\ p_3 &= -x_1x_2x_3 \end{aligned}$$

We note that the coefficients p_1, p_2, p_3 are symmetric functions of the roots x_1, x_2, x_3 which should not be suprising since the order in which we listed the roots was clearly irrelevant.

This pattern holds also for polynomials of higher degree and was well known to mathematicians of the 17^{th} and 18^{th} century.

Theorem 4.8.1 *Let $f(t) = t^n + p_1 t^{n-1} + \cdots + p_{n-1}t + p_n$ be a monic polynomial of degree n with roots $x_1, x_2, \ldots, x_n$. Then*

$$\begin{aligned} p_1 &= -(x_1 + x_2 + \cdots + x_n) = -\sum_{j=1}^{n} x_j \\ p_2 &= x_1x_2 + \cdots + x_{n-1}x_n = \sum_{1 \leq j < k \leq n} x_j x_k \\ p_3 &= -(x_1x_2x_3 + \cdots + x_{n-2}x_{n-1}x_n) = -\sum_{1 \leq j < k < \ell \leq n} x_j x_k x_\ell \\ &\vdots \\ p_n &= (-1)^n x_1 x_2 \cdots x_n \end{aligned}$$

Example 4.8.2 Find a polynomial with roots 1,2,3,4. By the Theorem

$$\begin{aligned} p_1 &= -(1 + 2 + 3 + 4) = -10 \\ p_2 &= 1*2 + 1*3 + 1*4 + 2*3 + 2*4 + 3*4 = 35 \\ p_3 &= -(1*2*3 + 1*2*4 + 1*3*4 + 2*3*4) = -50 \\ p_4 &= 1*2*3*4 = 24 \end{aligned}$$

so $f(t) = t^4 - 10t^3 + 35t^2 - 50t + 24$.

Example 4.8.3 Suppose we know for some reason (eg. we started Graeffe's method) that $f(x) = x^3 - x^2 - x - 15$ has an imaginary root of modulus $\sqrt{5}$. We wish to find the roots. Call this root $x_1 = a + bi$. Then $x_2 = a - bi$ is also a root so $|x_1|^2 = x_1x_2 = 5$. Now $p_3 = -15 = -x_1x_2x_3 = -5x_3$ so we see easily that $x_3 = 3$. But $p_1 = -x_1 - x_2 - x_3 = -2a - 3 = -1$ so $-2a = 2$ or $a = -1$. Since $|x_1|^2 = 5 = a^2 + b^2$ it follows that $b^2 = 4$ so $b = \pm 2$. Thus the roots are $-1 + 2i, -1 - 2i$ and 3.

Since Theorem 4.8.1 gives the close relation between the roots and the coefficients one might hope that these formulas might give a way to find the roots given the coefficients. Unfortunately this will not work, but to some extent these formulas are the basis for all attempts after the days of Cardano.

In Theorem 4.8.1 we can view $p_1, p_2, \ldots$ as symmetric polynomials in the variables $x_1, \ldots, x_n$. These symmetric polynomials are often known in the literature as the *elementary symmetric polynomials*. Thus we can view Theorem 4.6.1 as saying that the symmetric functions $s_k = x_1^k + \cdots + x_n^k$ are polynomials in the elementary symmetric polynomials, eg. as Girard noted $s_3 = -p_1^3 + 3p_1p_2 - 3p_3$. What is much more suprising is the following:

Theorem 4.8.4 (Fundamental Theorem on Symmetric Polynomials) *Let $f(x_1, x_2, \ldots, x_n)$ be a symmetric polynomial in n variables with coefficients in an integral domain R. Then $f(x_1, x_2, \ldots, x_n)$ can be expressed as a polynomial in the n elementary symmetric functions $p_1, p_2, \ldots, p_n$ with coefficients in R.*

The coefficient ring R will generally be $\boldsymbol{Z}$ or $\boldsymbol{Q}$. Proofs of this theorem are contained in the books by *Uspenski* and *Lang*, in addition *Adams and Loustanau* sketch a modern proof as an exercise and details of this proof can be found in *Fine and Rosenberger*. We will simply illustrate by some examples.

Example 4.8.5 In the two variable case the elementary symmetric functions are $p_1 = -x_1 - x_2$ and $p_2 = x_1x_2$. Consider the symmetric polynomial $D = (x_1 - x_2)^2$. Expanding, $D = x_1^2 - 2x_1x_2 + x_2^2 = s_2 - 2p_2$ where $s_2 = x_1^2 + x_2^2$ as in §4.6. By Newton's Identities $s_2 = p_1^2 - 2p_2$ so $D = p_1^2 - 2p_2 - 2p_2 = p_1^2 - 4p_2$ Note that this is just the discriminant of the quadratic polynomial $f(t) = t^2 + p_1t + p_2$.

Example 4.8.6 For a real cubic polynomial $f(t) = t^3 + p_1t^2 + p_2t + p_3$ the discriminant is $D = (x_1 - x_2)^2(x_1 - x_3)^2(x_2 - x_3)^2$ where $x_1, x_2, \ldots, x_3$ are the roots. It is a bit much work to put down here but it has been shown that

$$D = 18p_1p_2p_3 - 4p_1^3p_3 + p_1^2p_2^2 - 4p_2^3 - 27p_3^2$$

If $p_1 = 0$ as in §4 then we simply have $D = -4p_2^3 - 27p_3^2$ as in Theorem 4.4.1.

More generally given a polynomial $f(t) = t^n + p_1 t^{n-1} + \cdots + p_n$ of degree n with roots $x_1, \ldots, x_n$ the *discriminant* is defined by

$$D = (x_1 - x_2)^2 (x_1 - x_3)^2 \cdots (x_{n-1} - x_n)^2$$

i.e. the product of all differences $(x_j - x_k)^2$ for $j < k$. The general formula, which requires some advanced techniques to justify, is

$$D = \det \begin{bmatrix} s_0 & s_1 & \cdots & s_{n-1} \\ s_1 & s_2 & \cdots & s_n \\ s_2 & s_3 & \cdots & s_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ s_{n-1} & s_n & \cdots & s_{2n-2} \end{bmatrix}$$

where as in §6 $s_k = x_1^k + \cdots + x_n^k$. Newton's Identities then can be used to express D in terms of the p_j's. Note in particular that if the coefficients p_j are real then D is real, if the p_j are all integers then D is an integer. The generalization of Theorem 4.4.1 is

Theorem 4.8.7 *Let D be the discriminant of $f(t)$ as above, where the coefficients p_j are real. Then $f(t)$ has multiple roots if and only if $D = 0$. Otherwise if $D > 0$ then $f(t)$ has an even number of* **pairs** *of imaginary roots, if $D < 0$ then $f(t)$ has an odd number of* **pairs** *of imaginary roots.*

It should go without saying that except for $n = 2, 3$ calculating the discriminant is a terrible way to tell if $f(t)$ has multiple roots or to count real roots.

We will need a few more calculations for the next section:

Example 4.8.8 Let $f(t) = t^4 + p_1 t^3 + p_2 t^2 + p_3 t + p_4$ be a biquadratic polynomial with roots x_1, x_2, x_3, x_4. Consider

$$\begin{aligned} A &= x_1 + x_2 - x_3 - x_4 \\ B &= x_1 - x_2 + x_3 - x_4 \\ C &= x_1 - x_2 - x_3 + x_4 \end{aligned}$$

Let $a = A^2, b = B^2$ and $c = C^2$. Then $a + b + c, ab + ac + bc$ and ABC are all seen to be symmetric polynomials in $x_1, \ldots, x_4$. We can calculate

$$a + b + c = 3p_1^2 - 8p_2 \tag{4.15}$$

$$ab + ac + bc = 3p_1^4 - 16p_1^2 p_2 + 16p_1 p_3 + 16p_2^2 - 64p_4 \tag{4.16}$$

$$ABC = p_1^3 - 4p_1 p_2 + 8p_3 \tag{4.17}$$

We leave the first two calculations as rather hard exercises for the reader (see *Uspensky* for example) and tackle only the third. We note that multiplying out we can have 3 types of terms, $x_j^3, x_j^2x_k$ and $x_jx_kx_\ell$ where in each term j, k, ℓ are different. Note that x_1 occurs only with "$+$" signs, but each of x_2, x_3, x_4 occur exactly twice with $-$ signs in the expressions A, B, C. Thus the expansion of ABC will contain the terms x_j^3 for each j so it will contain $s_3 = x_1^3 + x_2^3 + x_3^3 + x_4^3$.

To get a term of the form $x_j^2x_k$ we must pick two of A, B, C to pick out the j and the k comes from the third. There are 3 ways to do this. It can be seen that two of the ways give a "$-$" sign but the third gives a "$+$" so each of the 12 terms $x_j^2x_k$ occurs with a coefficient of -1 in the expansion. Now note, for example when $j = 1$, $x_1^2x_2 + x_1^2x_3 + x_1^2x_4 = x_1^2x_1 + x_1^2x_2 + x_1^2x_3 + x_1^2x_4 - x_1^3 = x_1^2(-p_1) - x_1^3$. Repeating this for $j = 2, 3, 4$, adding and multiplying by the -1, we see that the contribution of the $x_j^2x_k$ in ABC is $(x_1^2 + x_2^2 + x_3^2 + x_4^2)p_1 + (x_1^3 + x_2^3 + x_3^3 + x_4^3) = s_2p_1 + s_3$.

Finally we have terms of the form $x_jx_kx_\ell$ where we can take $j < k < \ell$. Note that each such term can be generated 6 ways, i.e. we can choose the j from either factor A, B or C, then we have only two factors from which to choose the k and we must choose the ℓ from the remaining factor. By careful inspection, we see that 4 ways give "$+$" signs and 2 ways give "$-$" signs, thus the contribution of the $x_jx_kx_\ell$ terms in the product ABC is $2x_1x_2x_3 + 2x_1x_2x_4 + 2x_1x_3x_4 + 2x_2x_3x_4 = -2p_3$.

Thus we can conclude that $ABC = s_3 + (s_2p_1 + s_3) - 2p_3 = 2s_3 + s_2p_1 - 2p_3 = 2(-p_1^3 + 3p_1p_2 - 3p_3) + (p_1^2 - 2p_2)p_1 - 2p_3 = -p_1^3 + 4p_1p_2 - 8p_3$ as claimed.

Exercise 23d[10 points] Write the symmetric polynomial $f(x_1, x_2, x_3) = x_1^3x_2x_3 + x_1x_2^3x_3 + x_1x_2x_3^3$ in three variables in terms of the elementary symmetric functions p_1, p_2, p_3 of three variables.

4.9 Lagrange's Solution of the Biquadratic

After Cardano's publication of del Ferro, Tartaglia and Ferrari's solution of the cubic and biquadratic, many mathematicians tried to find similar methods for solving the quintic (5th degree) and higher degree equations. They failed, as we now know they must, and so, for the most part, history has not recorded their efforts. Several of the attempts were more noteworthy than the others, for example Vandermonde's almost correct solution (in 1770) of the cyclotomic equation $x^n - 1$ in radicals of degree less than n (Gauss filled in the details in 1801). However the most significant attempt was made by Lagrange. In order to attack higher degree polynomials he started with a detailed analysis of the solution of the cubic and biquadratic. The information he gained

was, of course, not enough to help him solve the quintic, but it laid the foundation for the proofs by Ruffini, Abel and Galois that the quintic and higher degree polynomials could not, in general, be solved by radicals. A good discussion of the history of these ideas can be found in *B.L van der Waerden's* "A History of Algebra". Lagrange devised solution methods for the cubic and the biquadratic, these are given in Chapter XI of *Uspensky*. We give his solution of the biquadratic since, unlike the cubic where Lagrange re-derives Cardano's equations (see Exercise 23e), the solution is given in a different, and more elegant, form than that of Ferrari.

We start with a polynomial $f(t) = t^4 + pt^3 + qt^2 + rt + s$ where for notational simplicity we are using the letters p, q, r, s instead of p_1, p_2, p_3, p_4 respectively. Let x_1, x_2, x_3, x_4 be the roots of $f(t)$. As in example 4.8.8 we let

$$\begin{aligned} A &= x_1 + x_2 - x_3 - x_4 \\ B &= x_1 - x_2 + x_3 - x_4 \\ C &= x_1 - x_2 - x_3 + x_4 \end{aligned} \tag{4.18}$$

and $a = A^2, b = B^2$ and $c = C^2$. We then have

$$a + b + c = 3p^2 - 8q = u \tag{4.19}$$

$$ab + ac + bc = 3p^4 - 16p^2q + 16pr + 16q^2 - 64s = v \tag{4.20}$$

$$abc = (p^3 - 4pq + 8r)^2 = w \tag{4.21}$$

From Theorem 4.8.1 it follows that a, b, c are the roots of the *resolvent cubic*

$$g(t) = t^3 + ut^2 + vt + w$$

The cubic equation $g(t) = 0$ can be solved by Cardano's method (or Lagrange's method in Exercise 23e) so a, b, c can be calculated. Then A, B, C can be found by taking square roots of a, b, c, being careful only to select signs so that

$$ABC = -p^3 + 4pq - 8r$$

as required by Example 4.8.8. We then have equations (4.18) together with the equation

$$-p = x_1 + x_2 + x_3 + x_4$$

Thus we have a system of 4 linear equations in the 4 unknowns x_1, x_2, x_3, x_4 which can

be solved once and for all by

$$\begin{aligned} x_1 &= \frac{-p+A+B+C}{4} \\ x_2 &= \frac{-p+A-B-C}{4} \\ x_3 &= \frac{-p-A+B-C}{4} \\ x_4 &= \frac{-p-A-B+C}{4} \end{aligned}$$

While this method is very elegant in theory, we warn the reader that in practice we may have $u \neq 0$ in the resolvant cubic, and in any case the resolvant cubic may have a messy solution, the square roots of which must then be calculated!

We end this section with an analysis of Lagrange's method. It is based on the following theorem, which is a special case of a theorem proved by Lagrange. This theorem deals with what we might call *somewhat symmetric polynomials* in that some permutations, but perhaps not all, may leave the value unchanged.

Theorem 4.9.1 *Let $g(x_1, x_2, \ldots, x_n)$ be a polynomial in n variables. Suppose that under all permutations of the variables the polynomial takes on exactly m different values, $g_1 = g, g_2, \ldots, g_m$ then there is a polynomial $f(t)$ of degree m whose coefficients are polynomials in the elementary symmetric functions of $x_1, \ldots, x_n$ so that the roots of $f(t)$ are $g_1, g_2, \ldots, g_m$.*

The idea of the proof is to construct the elementary symmetric functions $P_1, \ldots, P_m$ of the g_j, i.e. $P_1 = g_1 + g_2 + \cdots + g_m$, $P_m = g_1 g_2 \cdots g_m$ etc. It can be seen that the P_j are actual symmetric functions and hence by Theorem 4.8.4 expressible in the elementary symmetric functions on $x_1, \ldots, x_n$. But by Theorem 4.8.1 $g_1, \ldots, g_m$ are roots of $f(t) = t^m + P_1 t^{m-1} + \cdots + P_m$.

In Lagrange's solution of the biquadratic we took for $g(x_1, \ldots, x_n)$ the somewhat symmetric function $a = (x_1 + x_2 - x_3 - x_4)^2$. It should be noted that permuting the variables gives exactly three different values, mainly a, b and c. Thus a, b, c are roots of a polynomial whose coefficients are polynomials in p, q, r and s, mainly the resolvent polynomial.

In connection with Theorem 4.9.1, Lagrange noted that if one multiplied the number of different values taken by $g(x_1, \ldots, x_n)$ by the number of permutations which left $g(x_1, \ldots, x_n)$ unchanged, the product will always be $n!$, the number of all permutations

of the n variables. For example, in the paragraph above, a takes on 3 distinct values under permutations, but there are 6 ways to permute the variables so that a remains unchanged. $4 * 6 = 24 = 4!$. At the time Lagrange lived, group theory had not yet been invented. When group theory was invented 100 years later the mathematician Camile Jordan named a now famous theorem of group theory after Lagrange because Lagrange's observation was simply a special case of this general theorem.

Exercise 23e [30 points] Derive Lagrange's solution of the cubic. Let $f(t) = t^3 + pt^2 + qt + r$ have roots x_1, x_2, x_3 and set $A = x_1 + \omega x_2 + \omega^2 x_3$ and $B = x_1 + \omega^2 x_2 + \omega x_3$ where $\omega = \frac{-1+\sqrt{3}i}{2}$ is a cube root of 1. Show that $a = A^3$ takes on only 2 values under permutations of the roots, mainly a and $b = B^3$ and thus these two values are roots of the resolvent quadratic. Find the coefficients of the resolvent quadratic in terms of p, q and r. Note also that AB is a symmetric function of the roots so calculate AB in terms of p, q, r. Then solving the resolvent quadratic, taking appropriate cube roots of a, b to get the correct AB, the formulas for A, B and $-p = x_1 + x_2 + x_3$ give three linear equations in 3 unknowns which can be solved to obtain Cardano's Equations.

4.10 Insolvability of the Quintic

Lagrange had hoped that his study of solution methods of the cubic and biquadratic would lead to a solution of the quintic (5th degree polynomial equation). After all, by his Theorem, all one needed to find was a suitable somewhat symmetric function of 5 variables that took on exactly 4 different values under permutation of the variables. If this function was a power of a linear function of the variables, then with the additional equation $-p_1 = x_1 + \cdots + x_5$ he could then solve the "resolvent biquadratic" and take roots to obtain 5 equations in 5 unknowns which could be solved for the roots of the original polynomial. Unfortunately, such a function does not exist.

While Lagrange was optimistic that the solution method would be found, the Italian mathematian Paulo Ruffini realized that Lagrange's analysis could lead instead to a proof that no solution could exist. Ruffini claimed that he had such a proof in 1798 but many mathematicians were skeptical. What Ruffini actually proved was that Lagrange's method would not lead to a solution, but there was a gap in his argument that if Lagrange's method did not work, no method would work. In 1824 the 22 year old Niels Henrik Abel filled the gap in Ruffini's proof.

It is important to understand exactly what Ruffini and Abel proved. They started with **variables** $x_1, x_2, \ldots, x_5$ and defined $p_1, p_2, \ldots, p_5$ as in 4.8.1. The polynomial

$f(t) = t^5 + p_1 t^4 + \cdots + p_5$ is called the general quintic. The goal was to solve this in terms of $p_1, \ldots, p_5$ using only the algebraic operations of addition, subtraction, multiplication, division and the taking of roots (square roots, cube roots, etc.). In other words, the goal is to recover the variables $x_1, \ldots, x_5$ from the polynomials $p_1, \ldots, p_5$ using only algebraic operations. Of course if this were possible then by replacing the p_j by the coefficients of an actual polynomial and then doing these algebraic operations then the numbers $x_1, \ldots, x_5$ obtained would be the actual roots.

For example, in the case of the quadratic, the *general quadratic* is $f(t) = t^2 + pt + q$ where $p = -(x_1 + x_2)$ and $q = x_1 x_2$. The *general quadratic formula* $x = \frac{-p \pm \sqrt{p^2 - 4q}}{2}$ then recovers x_1, x_2 from p, q as follows:

$$\begin{aligned}
\frac{-p \pm \sqrt{p^2 - 4q}}{2} &= \frac{x_1 + x_2 \pm \sqrt{(x_1 + x_2)^2 - 4x_1 x_2}}{2} \\
&= \frac{x_1 + x_2 \pm \sqrt{x_1^2 - 2x_1 x_2 + x_2^2}}{2} \\
&= \frac{x_1 + x_2 \pm \sqrt{(x_1 - x_2)^2}}{2} \\
&= \frac{x_1 + x_2 \pm (x_1 - x_2)}{2} \\
&= x_1 \text{ or } x_2
\end{aligned}$$

A solution of this type is called a *solution by radicals* for the general equation. What Ruffini and Abel proved is that a solution by radicals did not exist for the general quintic. In particular, there is no single solution method which works for all quintic equations.

Several questions still remain. First, perhaps while no one single solution method works for all quintics, maybe there are several methods one of which would work for any given quintic. For instance, Cardano thought, wrongly as we now know, that 13 different methods were necessary to solve all cubics. Perhaps we need 13 methods to solve all quintics? There certainly are some types of quintics which can be solved, for example the cyclotomic equation $x^5 - 1 = 0$ was solved by radicals by Vandermonde. Perhaps different methods would solve other types of quintics. The first question then is whether this is actually true.

Even if it was not possible to have a finite list of solution methods covering all quintics, one would surely expect that for any given quintic with rational coefficients the roots would be algebraic expressions involving rational numbers, sums, differences, products, quotients and roots of various orders. So the second question is: “is this true?”

In 1831 Evariste Galois showed that in fact the answer to both questions is no! For example, not only is there no algebraic method to find the roots of $x^5 - 6x + 3 = 0$ but the roots cannot be expressed in terms of radicals. Galois went much further than this, by showing that this negative result applies also to polynomials of degree higher than 5. More importantly, he gave a method for determining whether or not a given polynomial could be solved by radicals (at least in principle, if not in practice). Galois' method, like the method of Ruffini and Abel following Lagrange, involves permuting the roots of the polynomial. But unlike Lagrange, Galois does not allow all permutations of the roots, only a *group* of permutations which somehow preserve the algebra of numbers which can be built up from the roots (the set of such numbers is called the *root field* of the polynomial). Thus Galois replaces the polynomial by its root field and then replaces the root field by the abstract algebraic object now known as the *Galois group* and shows that solvability of the polynomial is equivalent to some facts about the structure of the group.

Galois' proof method, now known as *Galois Theory* remains this day as one of the most elegant theories in mathematics. As there are many good accounts in the mathematical literature we will not pursue this any further here. For the reader who wants a reasonably elementary introduction we recommend the account in *Birkhoff and Mac Lane* or the one in *Fine and Rosenberger*.

The technique of replacing one mathematical object (eg. polynomials) by others easier to analyze (eg. fields, groups) has become central in modern mathematics. In addition there are many important direct applications of Galois theory. In a modern research journal such as the *Bulletin of the American Mathematical Society* the name "Galois" appears as often as the name of any other single mathematician. Yet Galois' work did not bring him any fame, or even recognition, in his lifetime. Galois sent two papers to Cauchy, who lost them. He sent one paper to Fourier who promptly died, and this paper is also lost. His most important paper (1831) was given to the mathematicians Poisson and Lacroix to review, but they couldn't understand it. A year later Galois was shot in a duel, not yet 21 and not yet known. Finally in 1846 the paper was published by Liouville in his journal. However the importance of Galois' work did not become apparent to the mathematical public until 1870 when Jordan published his full account of Galois theory.