Netwon's Identities

Milan Mossé

January 2019

The introduction to Newton's identities owes much to [7]. The rest of this document rehearses some proofs of Newton's identities and catalogues a few others. (Eventually, I hope to turn the sections that merely catalogue proofs into ones that rehearse them.) If you see any typos or have any suggested improvements, please let me know! You can click on my name (above) to send me an email.

1 Introduction

Consider a field F and a polynomial f in $F[x]$ of degree n with roots $x_1, ..., x_n$. Let us assume that f is *monic*, i.e., that the coefficient of x^n is 1. Express

$$\begin{aligned} f(x) &= s_0 x^n + s_1 x^{n-1} + \cdots + s_{n-1} x + s_n \\ &= \prod_{i=1}^{n} (x - x_i). \end{aligned}$$

Expanding the above product, and considering an arbitrary coefficient s_i (with i between 1 and n), we have

$$s_i = (-1)^i \sum_{j_1 < \ldots < j_i} x_{j_1} \cdots x_{j_i}.$$

The polynomial s_i in $x_1, .., x_n$ is symmetric (it does not change if we renumber the roots x_i) and homogenous (all terms have the same degree). The polynomials $s_i' = s_i \cdot (-1)^i$ are called elementary symmetric polynomials because every symmetric polynomial in $x_1, ..., x_n$ can be uniquely written as a polynomial in $s_1', ..., s_n'$. We say that the s_i' form a *basis* for all such symmetric polynomials. *Another* such basis is given by $p_1, ..., p_n$, where

$$p_i(x_1, ..., x_n) = x_1^i + \cdots x_n^i.$$

The polynomials p_i are called *power sums*. The transition formulas between these two bases are known as "Netwon's formulas" or "Netwon's identities," and they first appeared in Isaac Newton's *Arithmetica universalis*, written between 1673 and 1683. In these notes, we outline some proof of these identities, but before we do that, it will help to consider how these identities can be formulated.[1] Here is a natural formulation:

Theorem 1.1. *Fix some positive integer k. We have*

$$ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} = 0 \text{ if } k \leq n$$

$$\sum_{i=0}^{n} s_i p_{k-i} = 0 \text{ if } k > n$$

Note that there are *infinitely many* identities: one for each choice of k. This is why a lot of people call the above theorem "Newton's identities" and not "Newton's identity." We can arrive at a more concise formulation, if we adopt the natural convention of making "coefficients" s_i equal to 0 when i is greater than n:

Theorem 1.2. *Fix some positive integer k and define $s_i = 0$ for all $i > n$. We have*

$$ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} = 0.$$

If one stares long enough, one can see that these two formulations are indeed equivalent. We can use these identities to calculate p_k for any k, using the coefficients of f. Indeed, using the second formulation, we find that

$$ks_k + s_0 p_k + \sum_{i=1}^{k-1} s_i p_{k-i} = 0,$$

and we can rearrange to solve for p_k. Recalling that $s_0 = 1$, we have

$$p_k = (-1)(ks_k + \sum_{i=1}^{k-1} s_i p_{k-i}).$$

For example, suppose $n = 3$. Then, using the roots x_1, x_2, x_3 of f, we have

$$\begin{aligned}
p_1 &= -s_1 = -(-1)^1(x_1 + x_2 + x_3) = x_1 + x_2 + x_3, \\
p_2 &= -(2s_2 + s_1 p_{k-1}) = -(2(x_1x_2 + x_2x_3 + x_1x_3) - (x_1 + x_2 + x_3)p_1) \\
&= (x_1 + x_2 + x_3)^2 - 2(x_1x_2 + x_2x_3 + x_1x_3) = x_1^2 + x_2^2 + x_3^2.
\end{aligned}$$

[1] The fact that there are different formulations of Newton's identities is made more complicated by the fact that many people use very different labels for relevant values. For example, the coefficients s_i are sometimes α_i, the power sums p_i are sometimes S_i; the polynomial $f(x)$ is sometimes $P(z)$; the indices on the s_i sometimes run 1 to n instead of n to 1; and sometimes one drops the assumption that f is monic. Here, I have chosen to consistently use the notation that (I hope) makes things clearest, and I've tried to be very careful about translating between other's notation and my own, so that all (relevantly) different formulations of the identities are explored in the introduction. One thing unusual about my notation is that elementary symmetric polynomials are written s_i', and not s_i, but I am not sure that I could change that notation and keep things as clear as (I hope) they are.

The above formulas, and the analgous ones for p_i with i at most six, were obtained by Albert Girard in 1629, over 30 years before Newton's work (but Newton is thought to have been ignorant of this). For this reason, Newton's identities are also known as the Newton–Girard formulae. Note that we don't actually need to know what the roots are in order to use the formulae to solve for p_k; we just need the coefficients s_i of f.

We will turn shortly to our first proof of Newton's identities, but first, a brief remark. The assumption that f is monic is not strictly necessary: we could allow $a_0 \neq 0$, and then we would find that

$$\begin{aligned} f(x) &= s_0x^n + s_1x^{n-1} + \cdots + s_{n-1}x + s_n \\ &= s_0 \prod_{i=1}^{n} (x - x_i), \end{aligned}$$

and then we would have

$$s_i = \left(\frac{1}{s_0}\right)(-1)^i \sum_{j_1 < \ldots < j_i} x_{j_1} \cdots x_{j_i}.$$

The identities, with s_i defined in this new way, would then hold. (Indeed, as we will see, they are sometimes stated and proven that way; see [3].) But this makes the expression for s_i messier than it already is, so for readability, we'll often assume f is monic. Now, let us consider our first proof of Newton's identities.

2 Proof from the Case $n = k$

We prove the special case $n = k$ and derive the general identities from this case.

Theorem 2.1. *Let $k = n$. We claim that*

$$ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} = 0.$$

Proof. Let f be as above, with roots $x_1, ..., x_n$. Recall that

$$f(x) = s_0x^n + s_1x^{n-1} + \cdots + s_{n-1}x + s_n$$

Consider $f(x_j)$ for any j between 1 and n:

$$\sum_{i=0}^{k} s_{k-i}x_j^i = 0.$$

Summing over all j gives

$$ks_k + \sum_{i=1}^{k} s_{k-i}p_i = 0,$$

which is what we wanted to show (to see this, one fiddles with indices). □

The general identities follow from this one. Indeed, suppose first that $k > n$. Informally, we can throw in an extra $k-n$ roots by adding them to f, and then set them equal to 0 to obtain the identity

$$\sum_{i=0}^{n} s_i p_{k-i} = 0.$$

Formally, let

$$f'(x) = f(x) \prod_{i=k-n+1}^{k} (x - \alpha_i),$$

where the α_i are arbitrary. Then run the earlier argument on f' instead of f, and set the α_i to 0. Since

$$s_i = (-1)^i \sum_{j_1 < \ldots < j_i} x_{j_1} \cdots x_{j_i},$$

any term in which an α_i appears will be equal to 0, and the desired identity holds.

Now, suppose instead that $k < n$. We would like to show that

$$k s_k + \sum_{i=1}^{k} s_{k-i} p_i = 0.$$

If we combine like terms, it will suffice to show that the coefficient of any term

$$x_1^{a_1} \cdots x_n^{a_n}, \text{ with each } a_i \text{ a nonnegative integer,}$$

is 0. Since at most k of the a_i are nonzero, we can delete at least $n-k$ roots x_i from the monomial and not change its value. But then we know that the coefficient of the monomial *must* be 0. For we have, in effect, set $n-k$ of the roots x_i to 0, and are dealing with the case where f is a polynomial of degree k; and we know from this case that the identity holds, i.e., that the coefficients of the combined terms are 0.

3 Combinatorial Proof (1983)

In this section, we give a combinatorial proof of Newton's identities. A combinatorial proof is usually either (a) a proof that shows that two quantities are equal by giving a bijection between them, or (b) a proof that counts the same quantity in two different ways. Before we discuss Newton's identities, the following example may be help to clarify what (b) comes to. If you've seen this example before or would like to get straight to the identities, feel free to skip it.

Theorem 3.1. *Let $k \leq n$ be positive nonnegative integers. Then*

$$k! \binom{n}{k} = n(n-1)(n-2) \cdots (n-k+1)$$

Proof. Consider any set S with n elements. Let us count the number N_k of k-element sequences of elements of S, where we do not repeat elements.

On the one hand, first we can choose an unordered set of k elements of S, and *then* put them in order. There are $\binom{n}{k}$ choices of elements, and $k!$ ways of ordering each choice: this shows that the quantity on the left-hand side (LHS) is equal to N_k.

On the other hand, we can construct the sequence *as* we choose elements, so that we pick one element to go first in the sequence, and another to go second, and so on. There are n choices for the first element, then $n-1$, and so on. This shows that the quantity on the RHS is also equal to N_k. $\square$

Now, following [9], we give a combinatorial proof of the concise formulation of Newton's identities:

Theorem 3.2. *Fix some positive integer k. We have*

$$ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} = 0.$$

Proof. Consider the set $\mathscr{A}(n,k)$ of tuples (A, j, ℓ) where

(i) A is a subset of $[n]$, with $|A|$ at most k. (Recall that $[n]$ is the set of whole numbers $\{1, ..., n\}$.)

(ii) j is a member of $[n]$.

(iii) $\ell = k - |A|$

(iv) If ℓ is 0, then j is in A.

Define the *weight* of (A, j, ℓ) by

$$w(A, j, \ell) = (-1)^{|A|}\big(\prod_{a \in A} x_a\big) \cdot x_j^{\ell}.$$

For example,

$$w(\{1, 3, 5\}, 2, 3) = (-1)^3 x_1 x_3 x_5 \cdot x_2^3.$$

To show the theorem, it will suffice to show that the sum in the theorem is the sum of the weights of all elements of $\mathscr{A}(n,k)$, and that this sum is 0.

First, we show that that the sum in the theorem is the sum of the weights of all elements of $\mathscr{A}(n,k)$. Using the identities in the introduction, we have

$$\begin{aligned} ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} + = \\ &= k(-1)^k \sum_{j_1 < \ldots < j_k} x_{j_1} \cdots x_{j_k} + \\ &\sum_{i=0}^{k-1} (-1)^i (x_1^{k-i} + \cdots x_n^{k-i}) \sum_{j_1 < \ldots < j_i} x_{j_1} \cdots x_{j_i} \quad (*) \end{aligned}$$

There are two summands on the RHS. First, we consider the first summand:

$$k(-1)^k \sum_{j_1<\ldots<j_k} x_{j_1} \cdots x_{j_k} \cdot 1.$$

Set $A = \{j_1, ..., j_k\}$. Then as we range over choices of indices, we range over all choices of A. Since ℓ is 0, $x_j^\ell = 1$ contributes nothing to the product (but is written above on the RHS, for clarity). By (iv), j is in A. This gives k choices of j:

$$\sum_{|A|=k; j\in A; \ell=0} w(A, j, \ell) = k \cdot \sum_{|A|=k; \ell=0} w(a, j', \ell),$$

where j' is an arbitrary element of A. This shows that the first summand in $(*)$ can be written as the sum of all weights of elements of $\mathscr{A}(n, k)$ with $|A| = k$. To see that all other elements make up the other summand in $(*)$, one can multiply it out.

It remains for us to show that the sum of the weights of all elements of $\mathscr{A}(n, k)$ is 0. Define a map $T\colon \mathscr{A} \to \mathscr{A}$ by

$$T(A, j, \ell) = \begin{cases} (A - \{j\}, j, \ell+1) \text{ if } j \text{ is in } A \\ (A \cup \{j\}, j, \ell-1) \text{ if } j \text{ is not in } A \end{cases}$$

(Intuitively, T adds j to A if it's in A, and removes it if it isn't, adjusting $\ell = k - |A|$ as required.) Then applying T takes us to a distinct set with opposite weight:

$$w(T(A, j, \ell)) = -w(A, j, \ell),$$

and T^2 is the identity (i.e., T is an *involution*). Thus, all the weights can be arranged in mutually cancelling pairs, and their sum is 0. $\square$

4 Proof Using Calculus (1968)

Here, following [3], we give a proof using some basic calculus. For a similar proof that uses generating functions, in the context of coding theory, see page 212 of [1].

To present this proof, we need to introduce and briefly discuss the n-reversal of a polynomial f, which is just the result of arranging the coefficients of f in reverse order.

Definition 4.1. *Consider a poynomial f (with roots $x_1, ..., x_n$):*

$$\begin{aligned} f(x) &= s_0 x^n + s_1 x^{n-1} + \cdots + s_{n-1} x + s_n \\ &= \prod_{i=1}^{n} (x - x_i). \end{aligned}$$

Then the n-reversal of f is the polynomial

$$rev_n(f) = s_n x^n + s_{n-1} x^{n-1} + \cdots + s_1 x + s_0.$$

In order to prove Newton's identities, we need the following lemma.

Lemma 4.2. *Let f be as above. Then $rev_n(f) = x^n f(1/x)$, and the roots of $rev_n(f)$ are $1/x_i$, for any x_i a root of f.*

Proof. By squinting, one intuits that $\mathrm{rev}_n(f) = x^n f(1/x)$. Note then that if x_i is a root of f, it follows that $1/x_i$ is a root of $\mathrm{rev}_n(f)$. Since $\mathrm{rev}_n(f)$ is of degree n, *all* of its roots are of this form. $\square$

As an aside, we mention that the interested reader may like to try using the above lemma to establish the following corollary.

Corollary 4.3. *Let f and g be polynomials of $n \geq m$, respectively, with g monic. Using the Euclidean algorithm, express $f = qg + r$ for some polynomials q, r. Then the reversal identity holds:*

$$rev_n(f) = rev_{n-m}(q) \cdot rev_m(g) + x^{n-m+1} \cdot rev_{m-1}(r).$$

For the solution, see [2]. Now, we can prove Newton's identities. For reasons that will become apparent, the less concise formulation is more useful here:

Theorem 4.4. *Fix some positive integer k. Assume that 0 is not a root of f (i.e., that $s_n \neq 0$). We have*

$$\begin{aligned} ks_k + \sum_{i=0}^{k-1} s_i p_{k-i} &= 0 \text{ if } k \leq n \\ \sum_{i=0}^{n} s_i p_{k-i} &= 0 \text{ if } k > n \end{aligned}$$

Proof. Let f be as above. Let v denote $\mathrm{rev}_n(f)$. Then using the above lemma,

$$\begin{aligned} v(x) &= s_n x^n + s_{n-1} x^{n-1} + \dots + s_0 \\ &= s_n \prod_{i=1}^{n} (x - x_i^{-1}). \end{aligned}$$

Looking at the first equality above, note that if we evaluate the k-th derivative of v at 0, we obtain the coefficient s_k:

$$v^{(k)}(0) = s_k$$

It turns out that the logarithmic derivative of v, when evaluated at 0, is a multiple of p_{k+1}. This proof proceeds by turning the relation between v and its logarithmic derivative into a relation between the polynomials s_k and p_{k+1}. Since 0 is not a root of f, it is not a root of v, and we can take its logarithmic derivative:

$$V(x) = \frac{v'(x)}{v(x)} = \sum_{i=1}^{n} (x - x_i^{-1})^{-1}.$$

To see why the equality on the RHS holds, use the generalized product rule on the factorization of v above, noting that the derivative of each factor is 1. Now, we claim that

$$V^{(k)}(0) = - \cdot k! \cdot p_{k+1}, \quad (*)$$

where $V^{(k)}$ is the k-th derivative of V. This holds because we have

$$V^1(x) = -1 \sum_{i=1}^{n} (x - x_i^{-1})^{-2}$$
$$V^2(x) = 2 \sum_{i=1}^{n} (x - x_i^{-1})^{-3},$$

and so on, and plugging in $x = 0$ gives, for example, $V^2(0) = -1 \cdot 2! \cdot p_3$. When k is even, the negative sign comes from the fact that $(1/ - x_i^{-1})^{k+1}$ is negative. When k is odd, the negative sign comes from our application of the power rule. This shows that $(*)$ is true. To complete the proof, we establish another equality, and we apply $(*)$ to derive Newton's identities.

Now, let $[V(x)v(x)]^{(k-1)}$ be the k-th derivative of $V(x)v(x)$. We have

$$\begin{aligned} v^{(k)}(x) &= [V(x)v(x)]^{(k-1)} \\ &= \sum_{i=0}^{k-1} \binom{k-1}{i} V^{(i)}(x) v^{(k-1-i)}(x), \end{aligned}$$

where the first equality comes from the definition of the logarithmic derivative V, and the second equality comes the product rule and the binomial theorem. Now, we replace $V^{(r)}$ with the expression in $(*)$ and rearrange to obtain

$$\frac{v^{(k)}(0)}{k!} = -\frac{1}{k} \sum_{i=0}^{k-1} \frac{v^{(k-1-i)}(0)}{(k-1-i)!} p_{i+1}.$$

Recalling that $v^{(k)}(0) = s_k$, we have

$$\begin{aligned} -ks_k &= \sum_{i=0}^{k-1} s_{k-(i+1)} p_{i+1} && \text{if } k \leq n, \\ 0 &= \sum_{i=k-n-1}^{k-1} s_{k-(i+1)} p_{i+1} && \text{if } k > n, \end{aligned}$$

which are Newton's identities, if one fiddles with the indices. □

5 Proof with Clever Notation (1992)

We owe this proof to [5]. Like the approach in [7], and the one we saw from the case $n = k$, this involves adding several equations together. Let us just introduce the notation, to get a feel for the approach. Let f be as above, of degree n with roots $x_1, .., x_n$. Let $(a_1, ..., a_n)$, where the a_i are nonnegative integers and nonincreasing from left to right, represent

$$\sum_{i_1 < \ldots < i_n} x_{i_1}^{a_1} \cdot x_{i_2}^{a_2} \cdots x_{i_n}^{a_n}.$$

If $a_i = 0$ for i greater than k, we can write $(a_1, ..., a_k)$ instead of $(a_1, ..., a_n)$. Then

$$p_i = (i),$$
$$s_i' = (1, ..., 1), \text{where 1 is repeated } i \text{ times.}$$

This notation makes statements (and proofs) of Newton's identities easier on the eyes. For example, if $n \geq k = 3$, we can subtract this equation:

$$(2)(1) = (3) + (2, 1)$$

from this one:

$$(1)(1, 1) = (2, 1) + 3(1, 1, 1),$$

to obtain the Newton identity:

$$p_3 - p_2 s_1' + p_1 s_2' - 3s_3' = 0 \implies 3s_3 + \sum_{i=0}^{k-1} s_i p_{k-i} = 0.$$

For the complete proof, we refer the reader to [5].

6 Proofs by Cases and by Induction (2003)

The interested reader may consult [7] for a proof by induction and [6] for a proof by cases. (As [5] remarks, these kinds of proofs can sometimes feel unmotivated and difficult to follow, but it usually does more good than harm to try to understand another perspective on things.)

7 Matrix Proof (2000)

Newton's identities provide a means of computing the characteristic polynomial of a matrix in terms of the traces of the powers of the matrix, and one can *derive* Newton's identities using matrices. I have a hard time thinking about matrices in general, but [4] provides an interesting and well-written proof.

References

[1] Berlekamp, Elwyn. *Algebraic Coding Theory.* New York: McGraw Hill Book Co. 1968.

[2] dxiv. Polynomial reversal. Retrieved January 20, 2017.

[3] Eidswick, J.A. "A Proof of Newton's Power Sum Formulas." *The American Mathematical Monthly,* Vol. 75, no. 4 (April, 1968): 396-397.

[4] Kalman, Dan. "A Matrix Proof of Newton's Identities." *Mathematics Magazine*, Vol. 73, No. 4 (2000): 313-315.

[5] Mead, D.G. "Netwon's Identities." *The American Mathematical Monthly,* Vol. 99, No. 8 (1992): 749-751.

[6] Mináč, Ján. "Newton's Identities Once Again!" *The American Mathematical Monthly,* Vol. 110, No. 3 (2003): 232-234.

[7] Reichstein, Zinovy."An inductive proof of Newton's identities."

[8] Wikipedia contributors. "Newton's identities." *Wikipedia, The Free Encyclopedia.*

[9] Zeilberger, Doron. "A Combinatorial Proof of Newton's Identities." *Discrete Mathematics*, Vo. 49 (1985): 319.