Newton's Identities

Milan Mossé

This note introduces Newton's identities, rehearses some proofs of them, and catalogues a few others.[1]

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, observe that

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

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*, since 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 "Newton's formulas" or "Newton's identities," and they first appeared in Isaac Newton's *Arithmetica universalis*, written between 1673 and 1683. In these notes, we outline some proofs of these identities, which can be stated as follows:

Theorem 1.1 (Newton's identities)**.** *Fix some positive integer k. We have*

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

Adopting the convention of stipulating that $s_i = 0$ whenever $i > n$, we arrive at a more concise formulation:

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

[1] The introduction owes much to Reichstein (2000). If you see any typos or have any suggested improvements, please don't hesitate to let me know—you can find my email, along with the most recent version of this document, on my website. Thanks to Ian Baynham, Davin Halim, and Grant Gustafson for helpful suggestions and corrections.

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 use these identities to calculate p_k for any k, using the coefficients of f. For example, we have the Newton identity

$$ks_k + s_0p_k + \sum_{i=1}^{k-1} s_ip_{k-i} = 0,$$

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

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

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

$$\begin{aligned} p_1 &= -s_1 = -(-1)^1(x_1 + x_2 + x_3) = x_1 + x_2 + x_3, \\ p_2 &= -(2s_2 + s_1p_{k-1}) = -(2(x_1x_2 + x_2x_3 + x_1x_3) - (x_1 + x_2 + x_3)p_1) \\ &= (x_1 + x_2 + x_3)^2 - 2(x_1x_2 + x_2x_3 + x_1x_3) = x_1^2 + x_2^2 + x_3^2. \end{aligned}$$

(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.) The above formulas for p_1 and p_2, and the analogous ones for p_i with $i \leq 6$, were obtained by Albert Girard in 1629, over 30 years before Newton's work (though Newton is thought not to have known this). For this reason, Newton's identities are also known as the Newton–Girard formulae.

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, e.g. Eidswick (1968).) 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*

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

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

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

Thus for any j between 1 and n:

$$f(x_j) = s_k + \sum_{i=0}^{k-1} s_ix_j^{k-i} = 0.$$

Summing over all j gives the desired result. □

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

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

Formally, let

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

where the α_i are arbitrary. Then run the earlier argument on g 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

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

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

In this section, we outline a combinatorial proof of Newton's identities, due to Zeilberger (1985):

Theorem 3.1. *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, for $k = 5$ we have

$$w(\{1, 3, 5\}, 4, 2) = (-1)^3 x_1 x_3 x_5 \cdot x_4^2.$$

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 removes j from A if it's in A and adds it to A if it's not in A, 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. □

4 Proofs using calculus

Here, following Eidswick (1968), we outline a proof using some basic calculus. For a similar proof that uses generating functions in the context of coding theory, see page 212 of Berlekamp (1968). For a different calculus proof that uses Laplace transforms, see Cîrnu (2010).

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 polynomial f (with roots $x_1,...,x_n$):*

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

Then the n-reversal of f is the polynomial

$$rev_n(f) = s_nx^n + s_{n-1}x^{n-1} + \cdots + s_1x + s_0.$$

In order to prove Newton's identities, we need the following lemma and corollary:

Lemma 4.2. *We have* $rev_n(f) = x^n f(1/x)$*. The roots of* $rev_n(f)$ *are* $1/x_i$*, for any root* x_i *of* f*.*

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

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

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

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*. We have*

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

Proof. Assume 0 is not a root of f, without loss of generality (why?). Then the n-reversal $v = \text{rev}_n(f)$ is well-defined. Using the above lemma,

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

Looking at the first equality above, note that

$$v^{(k)}(0) = s_k \cdot k!,$$

where $v^{(k)}(0)$ denotes the k-th derivative of v evaluated at 0. As we'll soon see, 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}. We take the logarithmic derivative of v:

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

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

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

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 $[V(x)v(x)]^{(k)}$ 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

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

Recalling that $v^{(k)}(0) = s_k \cdot k!$, 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

We owe this proof to Mead (1992). Like the approach in Reichstein (2000), and the one we saw from the case $n = k$, this involves adding several equations together. First, 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 $k < n$, we can write $(a_1, ..., a_k)$ instead of $(a_1, ..., a_n)$. Then

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

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

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

which gives a relation among the Newton function (1) and the elementary symmetric functions $(1, 1)$ and $(1, 1, 1)$. We then subtract from it (in order to eliminate the term $(2, 1)$) the equation

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

to obtain the Newton identity:

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

The general proof is the same in spirit. To simplify notation, let $(1_i) = s'_i$ be i ones, and let $(m, 1_i)$ be the number m, followed by i ones. Now, we would like to show that

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

Note that the LHS is the sum of the LHS of the following true equations, if we multiply the ith equation by $(-1)^{i-1}$:

$$\begin{aligned}
(k-1)(1_1) &= (k-0, 1_0) + (k-1, 1_1) \\
(k-2)(1_2) &= (k-1, 1_1) + (k-2, 1_2) \\
(k-3)(1_3) &= (k-2, 1_2) + (k-3, 1_3) \\
&\vdots
\end{aligned}$$

To make sure the LHS sums correctly, we must specify *last* equation in the above sequence. To that end, let $t = \min(k-1, n)$. Then if $n \geq k = t-1$, we define the last equation as follows:

$$(1)(1_{k-1}) = (2, 1_{k-2}) + k(1_k)$$

If instead $k > n = t$, we'll instead make the last equation

$$(k-n)(1_n) = (k-n+1, 1_{n-1}).$$

This is to ensure that the LHS of the equations sum to the LHS of the desired Newton identity.

It remains to show that the RHS sums to 0. Since we're multiplying each equation by increasing powers of (-1), the terms $(k-j, 1_j)$ cancel out. For example, comparing the first and second equations, we have

$$(k-0, 1_0) + (k-1, 1_1) - [(k-1, 1_1) + (k-2, 1_2)],$$

so that the term $(k-1, 1_1)$ cancels out when we sum the first and second equations.

6 Other proofs

Proofs by cases and by induction. The interested reader may consult Reichstein (2000) for a proof by induction and Mináč (2003) for a proof by cases.

Matrix proof. It is well-known that Newton's identities can be applied to compute the characteristic polynomial of a matrix in terms of the traces of the powers of the matrix (see for example the Wikipedia entry on Newton's identities). But one can also derive Newton's identities using a matrix representation Kalman (2000).

Further proofs. There are of course other proofs G. A. Baker (1959); Muirhead (1904) of Newton's identities.

References

Berlekamp, E. (1968). *Algebraic Coding Theory*. McGraw Hill Book Co., New York.

Cîrnu, M. (2010). Newton's identities and the laplace transform. *The American Mathematical Monthly*, 117(1):67–71.

Eidswick, J. A. (1968). A proof of newton's power sum formulas. *The American Mathematical Monthly*, 75(4):396–397.

G. A. Baker, J. (1959). A new derivation of newton's identities and their application to the calculation of the eigenvalues of a matrix. *J. Soc. Indust. Applied Math*, 7:143–148.

Kalman, D. (2000). A matrix proof of newton's identities. *Mathematics Magazine*, 73(4):313–315.

Mead, D. G. (1992). Newton's identities. *The American Mathematical Monthly*, 99(8):749–751.

Mináč, J. (2003). Newton's identities once again! *The American Mathematical Monthly*, 110(3):232–234.

Muirhead, R. F. (1904). Some proofs of newton's theorem on sums of powers of roots. *Proc. Edinb. Math. Soc.*, 23:66–70.

Reichstein, Z. (2000). An inductive proof of newton's identities. *Mat. Prosveshenie (Math. Ed.)*, 3:204–205.

Zeilberger, D. (1985). A combinatorial proof of newton's identities. *Discrete Mathematics*, 49:319.