Formulae of Newton and Euler – the formal derivative and trace

K. R. Nagarajan and R. Sridharan*

Chennai Mathematical Institute, 92, G. N. Chetty Road, T. Nagar, Chennai 600 017, India

The well-known formulae of Newton and Euler on roots of polynomials are shown to follow from a general result relating the formal derivative of a polynomial and the trace map.

§1. Let k be a field and $f(X)$ a monic polynomial with coefficients in k and with distinct roots. We denote by $f'(X)$, the formal derivative of $f(X)$. We recall a classical result, due to Newton which relates the sums of powers of the roots of f to the coefficients of the polynomial f'.

Let $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$ and $\alpha_1, \ldots, \alpha_n$ the (distinct) roots of f. We denote by s_i, the sum of the ith powers of the roots for $i \geq 1$, set $s_0 = n$, and $a_n = 1$. The formula of Newton asserts that for all i, with $0 \leq i \leq n$,

$$(n-i)a_{n-i} = \sum_{j=0}^{i} a_{n-i+j}s_j.$$

A result of Euler is closely related to this; namely, under the same conditions as above,

$$\sum_{j=1}^{n} \frac{\alpha_j^i}{f'(\alpha_j)} = 0 \text{ if } 0 \leq i \leq n-2,$$
$$= 1 \text{ if } i = n-1.$$

The above theorems are usually proved by expanding certain rational functions into formal power series and equating coefficients. For instance, if $f(X) = \prod_{1 \leq i \leq n}(X-\alpha_i)$, we have

$$\frac{f'(X)}{f(X)} = \sum_{i=1}^{n} \frac{1}{X-\alpha_i},$$

so that

$$f'(X) = \frac{f(X)}{X} = \sum_{i=1}^{n}\left(1 + \frac{\alpha_i}{X} + \left(\frac{\alpha_i}{X}\right)^2 + \cdots\right).$$

Newton's formula follows by equating the coefficients of X^r, $0 \leq r \leq n-1$. For a proof of Euler's result, see ref. 1 (Lemma 2, p. 56) and ref. 2 (§46).

*For correspondence. (e-mail: rsridhar@cmi.ac.in)

The purpose of this note is to show that both the above formulae follow from a general result which connects the formal derivative of $f(X)$ with the trace map of the regular representation of $k[X]/(f)$, where k is any commutative ring and f any monic polynomial over k (see Proposition 1). This result seems to be of some independent interest. The proof is by a rather quaint inductive argument on the degree of f.

§2. Let A be any commutative ring with identity, and $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$, a monic polynomial of degree n with coefficients in A. Let $B = A[X]/(f)$. Then B is a free A-module with basis $1, x, \ldots, x^{n-1}$, where x denotes the residue class of X modulo (f).

We define an A-linear map $\phi_B = \phi$: $B \rightarrow A$, by setting $\phi(x^i) = 0$ if $0 \leq i \leq n-2$ and $\phi(x^{n-1}) = 1$. This gives rise to the A-bilinear form $\langle , \rangle$: $B \times B \rightarrow A$ defined by $\langle \lambda, \mu \rangle = \phi(\lambda\mu)$. The matrix of this form with respect to the ordered basis $\{1, x, x^2, \ldots, x^{n-1}\}$, being

$$\begin{pmatrix} 0 & \cdots & \cdots & 0 & 1 \\ 0 & \cdots & \cdots & 1 & * \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 1 & * & \cdots & * \\ 1 & * & * & \cdots & * \end{pmatrix},$$

is nonsingular. Hence we can associate an A-linear (in fact B-linear) *isomorphism* θ_B: $B \rightarrow B^*$ by setting for $b, b' \in B$,

$$\theta_B(b)(b') = \langle b, b' \rangle = \phi(bb').$$

We state this as a lemma.

Lemma 1. The map θ_B: $B \rightarrow B^*$ given by $\theta_B = \langle b, b' \rangle = \phi(bb')$ is an isomorphism.

Next, for any element $\lambda \in B$, multiplication by λ defines an A-linear map λ: $B \rightarrow B$. The trace of this map will be denoted $tr_{B/A}(\lambda)$. We now state the first main result.

Proposition 1. We have $\theta_B(f'(x)) = tr_{B/A}$. Equivalently, for any $\lambda \in B$, $\phi(\lambda f'(x)) = tr_{B/A}(\lambda)$.

The proof is by induction on the degree of f. (One could also argue by considering the minimal polynomial of the 'generic matrix', use the classical theorem of Euler, and deduce the above result by specialization. However, our proof appears to be more direct and elementary). But, first we show that we can reduce the proof to the case where $f(0) = 0$.

Lemma 2. Let $f = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ and let $f_1 = X^n + a_{n-1}X^{n-1} + \cdots + a_1X$.

Let B_1 be the ring $(A[X]/(f_1)) = A[x_1]$, where x_1 is the residue class modulo (f_1) of X. Then for $0 \leq i \leq n-1$, we have (a) $tr_{B/A}(x^i) = tr_{B_1/A}(x_1^i)$ and (b) $\theta_B(f'(x))(x^i) = \theta_{B_1}(f_1'(x_1))(x_1^i)$.

Proof. (a) We know that $1, x, \ldots, x^{n-1}$ is a basis of B as a free A-module. The matrix of the A-linear (multiplication) map x: $B \to B$ with respect to this basis is the 'companion matrix'

$$\begin{pmatrix} 0 & 0 & \cdots & \cdots & -a_0 \\ 1 & 0 & \cdots & \cdots & -a_1 \\ 0 & 1 & 0 & \cdots & -a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & -a_{n-1} \end{pmatrix} = M.$$

Similarly, taking $1, x_1, \ldots, x_1^{n-1}$ as basis of B_1/A, the matrix of x_1: $B_1 \to B_1$ with respect to this basis, is

$$\begin{pmatrix} 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 0 & \cdots & \cdots & -a_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & 1 & -a_{n-1} \end{pmatrix} = M_1.$$

Since trace is independent of bases, in order to prove (a), we need only to show that $tr\ M^r = tr\ M_1^r$ if $0 \le r \le n-1$. This is trivial if $r = 0$. So we assume $r \ge 1$. Let us denote by $E_{i,j}$ then $n \times n$ matrix with 1 at the i, jth place and zeros elsewhere. Then $M_1 = M + a_0E_{1,n}$, so that $M_1^r = M^r + W$, where W is an A-linear combination of words of length r in $E_{i,i-1}$, $i = 2, \ldots, n$, $E_{i,n}$, $i = 1, 2, \ldots, n$, and such that, $E_{1,n}$ occurs in each word. Since tr is additive, $tr(W)$ is the sum of the traces of each word in W. Let w be a word occurring in W. Remembering that $tr(PQ) = tr(QP)$ for any two $n \times n$ matrices P, Q, to prove the result, we may assume that $E_{1,n}$ is the last symbol in w. Let $w = w_1E_{1,n}$. Now $tr\ w \ne 0$ only if $w_1 = E_{n,1}$. This is, however, not possible, since w_1 is a word of length atmost $n-2$ in the symbols $E_{i,n}$, $i \ge 2$ and $E_{i,i-1}$, $2 \le i \le n$. It follows immediately that $tr\ W = 0$ and hence $tr\ M^r = tr\ M_1^r$, for $0 \le r \le n-1$, proving (a).

To prove (b), first note that the formal derivatives of f and f_1 are equal. We have to show that

$$\theta_B(f'(x))(x^i) = \theta_{B_1}(f_1'(x_1))(x_1^i),$$

or equivalently,

$$\phi_B(f'(x)x^i) = \phi_{B_1}(f'(x_1)x_1^i).$$

Since $f(X)$ is a monic polynomial, we can use the division algorithm and write

$$f'(X)X^i = \lambda(X)f(X) + \mu(X),$$

where $\lambda(X)$ and $\mu(X)$ are polynomials in $A[X]$, with degree $\mu(X) < n$ or $\mu(X) = 0$. This implies that degree λ + degree $f \le \max(\deg \lambda f'(X)X^i, n-1)$. We have degree $f'(X)X^i \le n-1+i$. Therefore $\deg \lambda + \deg f \leqq \max(n-1+i, n-1)$, and hence, $\deg \lambda \le \max(i-1, -1) \leqq n-2$, since $i \le n-1$. Now $\phi(f'(x)x^i)$ = coefficient of the $(n-1)$st degree term of the element $f'(x)x^i$, and this is the coefficient of the $(n-1)$st degree term of $\mu(X)$. Now let us compare this with $\phi(f_1'(x_1)x_1^i)$. Namely, let $f_1'(X)X^i = f'(X)X^i = \lambda_1 f_1 + \mu_1 = \lambda_1(f-a_0) + \mu_1 = \lambda_1 f + (\mu_1 - a_0\lambda)$ so that $\mu_1 = \mu + a_0\lambda$ and $\lambda = \lambda_1$. Hence

$$f_1'(X)X^i = \lambda(X)f_1(X) + (a_0\lambda(X) + \mu(X)).$$

It follows that $\phi_{B_1}(f'(x_1)x_1^i)$ = coefficient of the $(n-1)$st degree term in $a_0\lambda(X) + \mu(X)$ = coefficient of the $(n-1)$st degree term of $\mu(X)$, since degree $\lambda \le n-2$. This proves (b) and completes the proof of the lemma.

The next lemma gives the crucial step in the induction process. By the previous lemma, we are reduced to considering polynomials f such that $f(0) = 0$.

Lemma 3. Let $f = X^n + a_{n-1}X^{n-1} + \cdots + a_1X$ and $\tilde{f} = X^{n-1} + a_{n-1}X^{n-2} + \cdots + a_2X + a_1$. Let $B = A[X]/(f)$ and $\tilde{B} = A[X]/(\tilde{f})$. Let x and $\tilde{x}$ denote the images of X in B and $\tilde{B}$ respectively. Define $\phi_{\tilde{B}}: \tilde{B} \to A$ by $\phi_{\tilde{B}}(\tilde{x}^i) = 0$ if $0 \le i \le n-3$, and $\phi_{\tilde{B}}(\tilde{x}^{n-2}) = 1$. Then for $i \ge 1$,

(a) $tr_{B/A}(x^i) = tr_{\tilde{B}/A}(\tilde{x}^i)$,

(b) $\phi_B(x^i f'(x)) = \phi_{\tilde{B}}(\tilde{x}^i \tilde{f}'(\tilde{x}))$.

Proof. Consider the surjective homomorphism of A-algebras $B \to \tilde{B}$ given by $x \mapsto \tilde{x}$. Its kernel K is the principal ideal of B generated by $\tilde{f}(x)$. For each i with $i \ge 1$, we have the following commutative diagram with exact rows:

$$\begin{array}{ccccccccc} 0 & \to & K & \to & B & \to & \tilde{B} & \to & 0 \\ & & z\downarrow & & \downarrow x^i & & \downarrow \tilde{x}^i & & \\ 0 & \to & K & \to & B & \to & \tilde{B} & \to & 0 \end{array}$$

Note that the map z: $K \to K$ is the zero map since $x\tilde{f}(x) = f(x) = 0$ in B. The arrows are all A-linear. Since $tr_{B/A}(x^i) = tr_{\tilde{B}/A}(\tilde{x}^i) + tr_{K/A}(z)$ and $tr_{K/A}(z) = 0$, part (a) of the lemma follows.

We prove (b).

Let $X^i f'(X) = \lambda(X)f(X) + \mu(X)$ where $\mu = 0$ or $\deg \mu \le n-1$. As remarked earlier, $\phi_B(x^i f'(x))$ = coefficient of the $(n-1)$st degree term in $\mu(X)$. Now $f = X\tilde{f}$ implies $f'(X) = X\tilde{f}'(X) + \tilde{f}(X)$. Therefore

$$\begin{aligned} X^i(X\tilde{f}'(X) + \tilde{f}(X)) &= \lambda(X)f(X) + \mu(X) \\ &= \lambda(X).X\tilde{f}(X) + \mu(X). \end{aligned} \qquad (*)$$

Thus if $i \geq 1$, X divides $\mu(X)$, so that $\mu(X) = X\tilde{\mu}(X)$ with degree $\tilde{\mu}(X) \leq n-2$. We can cancel X in the equation (*) and write

$$X^i \tilde{f}'(X) = \lambda(X)\tilde{f}(X) - X^{i-1}\tilde{f}(X) + \tilde{\mu}(X).$$

Hence

$$\begin{aligned}\phi_{\tilde{B}}(\tilde{x}^i \tilde{f}'(\tilde{x})) &= \phi_{\tilde{B}}(\tilde{\mu}(\tilde{x}))\\ &= \text{coefficient of } X^{n-2} \text{ in } \tilde{\mu}(X)\\ &= \text{coefficient of } X^{n-1} \text{ in } \mu(X)\\ &= \phi_B(x^i f'(x)).\end{aligned}$$

This completes the proof of Lemma 3.

We now turn to the

Proof of Proposition 1. Since $1, x, \ldots, x^{n-1}$ is an A-basis of B, it is enough to show that $\theta_B(f'(x))(x^i) = tr_{B/A}(x^i)$ for $0 \leq i \leq n-1$. If $i = 0$, $\theta_B(f'(x)) = \phi(f'(x)) = n$ and $tr_{B/A}(1) = n$. Suppose $i \geq 1$. Then Lemmas 1 and 2 apply and we use induction on degree of f to complete the proof. □

Corollary: (Euler's Theorem). Let k be a field and f a monic polynomial over k with distinct roots $\alpha_1, \ldots, \alpha_n$. Then

$$\sum_{j=1}^{n} \frac{\alpha_j^i}{f'(\alpha_j)} = 0, \text{ if } 0 \leq i \leq n-2$$

$$= 1, \text{ if } i = n-1.$$

Proof. Let $B = k[X]/(f)$ and x be the image of X in B.

Since f has distinct roots, $f(X)$ and $f'(X)$ are coprime in $k[X]$, so that $f'(x)$ is an invertible element of B. By Proposition 1, we have $tr(\lambda) = \phi(\lambda f'(x))$ for all λ in B. Let $\lambda = x^i/f'(x)$. Then

$$tr_{B/A}\left(\frac{x^i}{f'(x)}\right) = \phi(x^i) = 0 \text{ if } 0 \leq i \leq n-2$$

$$\text{and} = 1 \text{ if } i = n-1.$$

Now, $tr(x^i/(f'(x)))$ = sum of the eigenvalues of the k-linear map $B \to B$ defined by multiplication by $x^i/(f'(x))$. Since the eigenvalues of x: $B \to B$ are $\alpha_1, \ldots, \alpha_n$, the eigenvalues of $x^i/f'(x) : B \to B$ are precisely $\alpha_j^i / f'(\alpha_j)$, $j = 1, \ldots, n$ and we have,

$$tr\left(\frac{x^i}{f'(x)}\right) = \sum_{j=1}^{n} \alpha_j^i / f'(\alpha_j).$$

The corollary now follows.

The next proposition is a reformulation of Newton's formula in the general set up.

Proposition 2. Let A be any commutative ring and $f(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0$ any monic polynomial. Let $B = A[X]/(f)$ and x the image of X in B. We set $a_n = 1$. Then, for $0 \leq i \leq n$,

$$tr_{B/A}(x^i + a_{n-1}x^{i-1} + \cdots + a_{n-i+1}x + a_{n-i}) = (n-i)a_{n-i}.$$

Proof. The result is trivial for $i = 0$. Also, since $x^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0 = 0$, the statement of the proposition is clear, if $i = n$. Thus if degree $f = 1$, the result follows. Let $n \geq 2$. We apply induction on degree of f.

Define $\tilde{f}$ and $\tilde{B}$ as in Lemma 3. Then for $1 \leq i \leq n-1$,

$$\begin{aligned}tr_{B/A}&(x^i + a_{n-1}x^{i-1} + \cdots + a_{n-i+1}x)\\ &= tr_{\tilde{B}/A}(\tilde{x}^i + a_{n-1}\tilde{x}^{i-1} + \cdots + a_{n-i+1}\tilde{x})\\ &= (n-1-i)a_{n-i} - (n-1)a_{n-i}\\ &= -ia_{n-i}\end{aligned}$$

by induction hypothesis, so that $tr_{B/A}(x^i + a_{n-1}x^{(i)} + \ldots + a_{n-i}) = (n-i)a_{n-i}$.

The completes the proof of Proposition 2.

Remark. We note that the formula of Proposition 2 gives an inductive procedure for the explicit computation of the traces of powers of a linear transformation. For example, we obtain,

$$tr_{B/A}x = -a_{n-1}, \quad tr_{B/A}(x^2) = a_{n-1}^2 - 2a_{n-2},$$

$$tr_{B/A}(x^3) = -a_{n-1}^3 + 3a_{n-1}a_{n-2} - 3a_{n-3}, \ldots.$$

We may also note the following. The statement of the proposition can be displayed in matrix form as

$$\begin{pmatrix} 1 & a_{n-1} & a_{n-2} & \cdots & \cdots & a_1 \\ 0 & 1 & a_{n-1} & \cdots & \cdots & a_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & \cdots & \cdots & \cdots & 1 & a_{n-1} \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix} \begin{pmatrix} tr\, x^{n-1} \\ tr\, x^{n-2} \\ \vdots \\ tr\, 1 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ \vdots \\ na_n \end{pmatrix}.$$

The inverse of the $n \times n$ matrix M in the above equation is a matrix of the same form so that it is determined by its first row, $(1, c_{n-1}, c_{n-2}, \ldots, c_1)$, where $c_{n-1} = -a_{n-1}$ and for $2 \leq i \leq n-1$, $c_{n-i} = -a_{n-1}c_{n-i+1} - a_{n-2}c_{n-i+2} - \cdots - a_{n-i+1}c_{n-1} - a_{n-i}$. (This can be directly verified or proved by writing the matrix M as $I + N$, so that $M^{-1} = I - N + N^2 - \cdots + (-1)^{n-1}N^{n-1}$. The successive powers of N can be computed easily).

Now

$$\begin{pmatrix} tr\ x^{n-1} \\ tr\ x^{n-2} \\ \vdots \\ tr\ 1 \end{pmatrix} = M^{-1} \begin{pmatrix} a_1 \\ 2a_2 \\ \vdots \\ na_n \end{pmatrix}$$

and hence we have, for $1 \le i \le n$,

$$tr\ x^{n-i} = ia_i + (i+1)a_{i+1}c_{n-1} + (i+2)a_{i+2}c_{n-2} + \cdots + (n-1)a_{n-1}c_{n-i+1} + nc_{n-i}.$$

It can also be verified that c_{n-i} is a polynomial in a_{n-1} of degree i with leading coefficient $(-1)^i$, so that $tr\ x^i$ is a polynomial in a_{n-1} of degree i, with the same property. This yields a fairly explicit formula for the traces of powers of a linear transformation in terms of the entries of the companion matrix.

1. Serre, J. P., *Local Fields*, GTM 67, Springer Verlag, New York, 1979.
2. Weber, H., *Lehr Buch Der Algebra*, Erster Band, Chelsea Publishing Co.

Received 11 April 2003; accepted 17 April 2003

A new plasma wave over low latitude ionosphere during Leonid meteor storm

S. P. Gupta, R. Sekar* and Y. B. Acharya

Physical Research Laboratory, Navrangpura, Ahmedabad 380 009, India

Leonid meteor storm is a unique astronomical event that occurs once in 33 years. In order to investigate the effect of Leonid meteor storm over low latitude ionosphere, rocket measurements of plasma parameters were carried out on 18 and 20 November 1999 from Sriharikota, India. The meteoric activity was at its peak on 18 November 1999. Results obtained on plasma waves using a high frequency Langmuir probe revealed for the first time, an experimental evidence for the presence of sub-meter scale size plasma wave over low latitude E-region. The peak amplitude of the plasma wave occurs at 105 km with a magnitude of ~4% of ambient electron density. The ambient plasma conditions during these measurements imply that the causative mechanism for the generation of this plasma wave is different from well known gradient drift waves.

LEONID meteor shower associated with comet Tempel-Tuttle occurs every year during 17–18 November with a typical meteor zenith hourly rate (ZHR) of around five[1]. However, Leonid meteor storm occurs once in about 33 years with intense ZHR of a few thousands[2,3] when the above comet approaches inner solar system. In the literature[3,4], *in situ* measurements during an intense meteor storm is not available from a low latitude region prior to the present rocket measurements during Leonid storm 1999. Two RH-300 Mark-II rockets (F.1 and F.2) carrying high frequency Langmuir probe sensor mounted on booms perpendicular to spin axes of the rockets were launched from Sriharikota (13.7°N, 80.2°E, dip lat. 6.0°N) on 18 and 20 November 1999 respectively at 7.25 and 7.03 IST (Indian Standard Time = UT + 5.30). The launch of F.1 coincided with the peak activity of Leonid meteor storm, while the launch of F.2 happens to be when the activity reduced to one-third. The ambient electron densities (n_e) and fluctuations (Δn_e) in them which represent the plasma waves were measured along the trajectories of the rockets. Earlier studies[5–7] on plasma waves during normal days revealed only gradient drift waves over a low latitude station like Sriharikota. However, over magnetic equator, other plasma waves known as type I waves[5,8,9] associated with equatorial electrojet have also been observed. The above two types of plasma waves have a cut-off scale size of a few meters[5,9].

Figure 1 depicts the telemetry raw data from a high frequency (100 Hz to 3 kHz) channel representing the fluctuations in the electron densities corresponding to about 105 km altitude obtained on 18 November 1999. The expanded portion of the diagram corresponding to the time interval when high frequency fluctuations are observed in Figure 1 *a*, is given in the bottom panel of the diagram as Figure 1 *b*. During a time interval of 0.02 s, forty peaks are seen in Figure 1 *b* which correspond to ~ 2 kHz wave frequency in the rocket frame of reference. Similar features were observed on 20 November 1999. It can be noticed from Figure 1 *a* that these high frequency fluctuations are found only on certain durations revealing that these fluctuations are geophysical and anisotropic in nature. Taking into consideration measured vertical velocity of the rocket (~ 1 km/s) at 105 km altitude, the observed 2 kHz plasma waves (in Figure 1 *b*) correspond to a scale size of about 50 cm. Thus, an evidence for the presence of sub-meter scale size of the plasma waves is reported here for the first time from a low latitude E-region of the ionosphere.

Figure 2 depicts the altitude profiles of electron density along with the average values of the amplitudes of the plasma waves observed on 18 November 1999. Considering the errors in the measurements, wave amplitudes greater than ~ 0.5% have physical significance. The plasma waves are observed to confine in the altitude region of 100 to 120 km with a maximum amplitude at an altitude of 105 km. From Figure 2 it is clear that the absence of plasma waves in a steep electron density gradient region (90–94 km) and a presence of maximum

*For correspondence. (e-mail: rsekar@prl.ernet.in)