TRACE AND NORM, II

KEITH CONRAD

1. Introduction

We continue the study of the trace and norm of a finite extension of fields L/K. The topics we address are the following:

- for $\alpha \in L$ and $g(X) \in K[X]$. expressing $\mathrm{Tr}_{L/K}(g(\alpha))$ and $\mathrm{N}_{L/K}(g(\alpha))$ in terms of the roots of $\chi_{\alpha,L/K}(X)$,
- the trace and norm as polynomial functions in terms of a basis of L/K,
- transitivity of the trace and norm (more subtle for the norm than the trace),
- the trace and norm when L/K is a Galois extension.

2. The Trace and Norm of Polynomial Values

If $\alpha \in L$ has minimal polynomial of degree d over K and that polynomial splits over a large enough field extension of K as $(X-\alpha_1)\cdots(X-\alpha_d)$, then we saw in the first handout on traces and norms that $\mathrm{Tr}_{L/K}(\alpha)$ and $\mathrm{N}_{L/K}(\alpha)$ can be written in terms of the α_i's:

$$\mathrm{Tr}_{L/K}(\alpha) = \frac{n}{d}(\alpha_1 + \cdots + \alpha_d), \quad \mathrm{N}_{L/K}(\alpha) = (\alpha_1 \cdots \alpha_d)^{n/d},$$

where $n = [L : K]$ and $d = [K(\alpha) : K]$.

Every number in $K(\alpha)$ is $g(\alpha)$ for some $g(X) \in K[X]$. What is its trace and norm in the extension L/K? For instance, what is $\mathrm{Tr}_{L/K}(\alpha^3)$ or $\mathrm{N}_{L/K}(\alpha^2 + 5\alpha - 1)$?

Theorem 2.1. *Suppose in a large enough field extension the characteristic polynomial of α for the extension L/K splits completely as*

$$\chi_{\alpha,L/K}(X) = (X - r_1)\cdots(X - r_n).$$

Then for $g(X) \in K[X]$, the characteristic polynomial of $g(\alpha)$ for L/K is

$$\chi_{g(\alpha),L/K}(X) = (X - g(r_1))\cdots(X - g(r_n)),$$

so

$$\mathrm{Tr}_{L/K}(g(\alpha)) = \sum_{i=1}^{n} g(r_i), \quad \mathrm{N}_{L/K}(g(\alpha)) = \prod_{i=1}^{n} g(r_i).$$

In particular, $\chi_{\alpha^m,L/K}(X) = (X - r_1^m)\cdots(X - r_n^m)$, so $\mathrm{Tr}_{L/K}(\alpha^m) = \sum_{i=1}^n r_i^m$.

Proof. The characteristic polynomial $\chi_{\alpha,L/K}(X)$ is a power of the minimal polynomial of α in $K[X]$, so every r_i has the same minimal polynomial over K as α.

Set $f(X) = (X - g(r_1))\cdots(X - g(r_n))$. We want to show this is the characteristic polynomial of $g(\alpha)$. The coefficients of $f(X)$ are symmetric polynomials in $r_1, \ldots, r_n$ with coefficients in K, so by the symmetric function theorem $f(X) \in K[X]$. Let $M(X)$ be the minimal polynomial of $g(\alpha)$ over K, so $M(X)$ is irreducible in $K[X]$. Since α and each r_i have the same minimal polynomial over K, the fields $K(\alpha)$ and $K(r_i)$ are isomorphic over K by sending α to r_i and fixing the elements of K. Applying such an isomorphism to the equation $M(g(\alpha)) = 0$ turns it into $M(g(r_i)) = 0$ (because $M(X)$ and $g(X)$ have coefficients in K), so $M(X)$ is the minimal polynomial for $g(r_i)$ over K since $M(X)$ is monic irreducible in $K[X]$.

We have shown all roots of $f(X)$ have minimal polynomial $M(X)$ in $K[X]$, and $f(X)$ is monic, so $f(X)$ is a power of $M(X)$. Since $\chi_{g(\alpha),L/K}(X)$ is a power of $M(X)$ with degree $[L : K] = n = \deg f$, we have $\chi_{g(\alpha),L/K}(X) = f(X)$. The formulas for $\mathrm{Tr}_{L/K}(g(\alpha))$ and $\mathrm{N}_{L/K}(g(\alpha))$ are obtained by looking at the coefficients in the characteristic polynomial where the trace and norm appear (up to a sign factor). □

Example 2.2. Let γ be a root of $X^3 - X - 1$. The general trace $\mathrm{Tr}_{\mathbf{Q}(\gamma)/\mathbf{Q}}(a + b\gamma + c\gamma^2)$ for $a, b, c \in \mathbf{Q}$ can be computed to be $3a + 2c$ by finding the 3×3 matrix for multiplication by $a + b\gamma + c\gamma^2$ in the basis $\{1, \gamma, \gamma^2\}$. We will now compute this trace without using matrices!

By linearity of the trace,

$$\mathrm{Tr}(a + b\gamma + c\gamma^2) = a\mathrm{Tr}(1) + b\mathrm{Tr}(\gamma) + c\mathrm{Tr}(\gamma^2),$$

where $\mathrm{Tr} = \mathrm{Tr}_{\mathbf{Q}(\gamma)/\mathbf{Q}}$. The trace of 1 is $[\mathbf{Q}(\gamma) : \mathbf{Q}] = 3$. Since γ generates $\mathbf{Q}(\gamma)/\mathbf{Q}$, $\chi_{\gamma,\mathbf{Q}(\gamma)/\mathbf{Q}}(X) = X^3 - X - 1$, and its X^2-coefficient is 0, so $\mathrm{Tr}(\gamma) = 0$. What is $\mathrm{Tr}(\gamma^2)$?

Let the roots of $X^3 - X - 1$ be r, s, and t. Theorem 2.1 tells us that $\mathrm{Tr}(\gamma^2) = r^2 + s^2 + t^2$, which can be expressed in terms of the coefficients of $X^3 - X - 1$:

$$r^2 + s^2 + t^2 = (r + s + t)^2 - 2(rs + rt + st).$$

From the relations between roots and coefficients of a (monic) polynomial, $r + s + t = 0$ and $rs + rt + st = -1$. Thus $r^2 + s^2 + t^2 = 0^2 - 2(-1) = 2$, so γ^2 has trace 2.

Since $\mathrm{Tr}(1) = 3$, $\mathrm{Tr}(\gamma) = 0$, and $\mathrm{Tr}(\gamma^2) = 2$, we get $\mathrm{Tr}(a + b\gamma + c\gamma^2) = 3a + 2c$.

In comparison with the trace, where we can take advantage of linearity, there is no way to compute the norm of a general element without essentially computing a determinant with variable entries. General norm formulas are often quite unwieldy when $[L : K] > 2$.

An important application of the trace and norm formulas in Theorem 2.1 is a derivation of the following formulas for the discriminant of a power basis.

Theorem 2.3. *Let $n = [K(\alpha) : K]$ and $f(X)$ be the minimal polynomial of α over K. If $f(X) = (X - \alpha_1) \cdots (X - \alpha_n)$ over a splitting field, then*

$$\mathrm{disc}_{K(\alpha)/K}(1, \alpha, \ldots, \alpha^{n-1}) = \prod_{i<j} (\alpha_j - \alpha_i)^2 = (-1)^{n(n-1)/2} \mathrm{N}_{K(\alpha)/K}(f'(\alpha)).$$

Proof. If $n = 1$ all the expressions equal 1 (an empty product is understood to be 1), so we can take $n \geq 2$.

By definition, for each basis $\{e_1, \ldots, e_n\}$ of a finite extension L/K, $\mathrm{disc}_{L/K}(e_1, \ldots, e_n) = \det(\mathrm{Tr}_{L/K}(e_i e_j))$. Using the basis $\{1, \alpha, \ldots, \alpha^{n-1}\}$ of $K(\alpha)/K$, so $e_i = \alpha^{i-1}$,

$$\mathrm{disc}_{K(\alpha)/K}(1, \alpha, \ldots, \alpha^{n-1}) = \det(\mathrm{Tr}_{K(\alpha)/K}(\alpha^{i-1}\alpha^{j-1})).$$

By Theorem 2.1, $\mathrm{Tr}_{K(\alpha)/K}(\alpha^m) = \sum_{k=1}^n \alpha_k^m$, so

$$\mathrm{Tr}_{K(\alpha)/K}(\alpha^{i-1}\alpha^{j-1}) = \sum_{k=1}^n \alpha_k^{i-1}\alpha_k^{j-1} = (\alpha_1^{i-1}, \ldots, \alpha_n^{i-1}) \cdot (\alpha_1^{j-1}, \ldots, \alpha_n^{j-1}),$$

where the row vectors on the right are being combined as a dot product. Let $\mathbf{v}_i = (\alpha_1^{i-1}, \ldots, \alpha_n^{i-1})$, so

$$\mathrm{disc}_{K(\alpha)/K}(1, \alpha, \ldots, \alpha^{n-1}) = \det(\mathbf{v}_i \cdot \mathbf{v}_j).$$

The matrix $(\mathbf{v}_i \cdot \mathbf{v}_j)$ can be written as the product $A^\top A$, where

$$A = \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{pmatrix}.$$

Therefore

$$\det(\mathbf{v}_i \cdot \mathbf{v}_j) = \det(A^\top A) = \det(A^\top)\det(A) = \det(A)\det(A) = \det \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & & | \end{pmatrix}^2,$$

so

$$\mathrm{disc}_{K(\alpha)/K}(1, \alpha, \ldots, \alpha^{n-1}) = \det \begin{pmatrix} | & | & & | \\ 1 & \alpha_i & \cdots & \alpha_i^{n-1} \\ | & | & & | \end{pmatrix}^2 = \det(\alpha_i^{j-1})^2.$$

The matrix (α_j^{i-1}) is called a Vandermonde matrix, and its determinant can be computed by Vandermonde's formula:

$$\det \begin{pmatrix} 1 & \alpha_1 & \cdots & \alpha_1^{n-1} \\ 1 & \alpha_2 & \cdots & \alpha_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \alpha_n & \cdots & \alpha_n^{n-1} \end{pmatrix} = \prod_{i<j} (\alpha_j - \alpha_i).$$

Square this and we have the first formula for the discriminant.

To show

$$\prod_{i<j} (\alpha_j - \alpha_i)^2 = (-1)^{n(n-1)/2} \mathrm{N}_{K(\alpha)/K}(f'(\alpha))$$

we will rearrange the terms on the left. From the product rule for derivatives,

$$f(X) = (X - \alpha_1) \cdots (X - \alpha_n) \Longrightarrow f'(\alpha_i) = \prod_{j \neq i} (\alpha_j - \alpha_i).$$

Multiplying these over all i,

$$\prod_{i=1}^n \prod_{j \neq i} (\alpha_j - \alpha_i) = \prod_{i=1}^n f'(\alpha_i).$$

The product of $\alpha_j - \alpha_i$ runs over sets of distinct indices i and j. To rewrite this product over index pairs where $i < j$, collect $\alpha_j - \alpha_i$ and $\alpha_i - \alpha_j$ together as $-(\alpha_j - \alpha_i)^2$. There are $\binom{n}{2} = \frac{n(n-1)}{2}$ such pairs, so

$$\prod_{i<j} (\alpha_j - \alpha_i)^2 = (-1)^{n(n-1)/2} \prod_{i=1}^n f'(\alpha_i).$$

The product of derivatives is $\mathrm{N}_{K(\alpha)/K}(f'(\alpha))$ by Theorem 2.1. □

Example 2.4. To compute the discriminant of the basis $\{1, \sqrt[3]{2}, \sqrt[3]{4}\}$ for $\mathbf{Q}(\sqrt[3]{2})/\mathbf{Q}$, we use the norm formula with the derivative of $f(X) = X^3 - 2$: writing N for $\mathrm{N}_{\mathbf{Q}(\sqrt[3]{2})/\mathbf{Q}}$,

$$(-1)^{3(3-1)/2} \mathrm{N}\left(3\sqrt[3]{2}^2\right) = -\mathrm{N}(3)(\mathrm{N}(\sqrt[3]{2}))^2 = -27 \cdot 4 = -108.$$

3. The Trace and Norm as Multivariable Polynomial Values

In calculations of the trace and norm, their formulas are polynomials in the coefficients of the basis that is used. For instance, $\mathrm{Tr}_{\mathbf{C}/\mathbf{R}}(a+bi) = 2a$ and $\mathrm{N}_{\mathbf{C}/\mathbf{R}}(a+bi) = a^2+b^2$, which are polynomials in the coefficients a and b. Quite generally, a *polynomial function* on a finite extension L/K is a function $f\colon L \to K$ such that for some K-basis $\{e_1,\dots,e_n\}$ of L and some polynomial $P(x_1,\dots,x_n) \in K[x_1,\dots,x_n]$, $f(\sum_{i=1}^n c_ie_i) = P(c_1,\dots,c_n)$ for all $c_i \in K$. When this holds for one K-basis of L it holds for every K-basis of L, with P usually changing when the basis changes, so the concept of a polynomial function is independent of the choice of basis.

Theorem 3.1. *The trace and norm of α are both polynomial functions in coordinates of α: if we pick a K-basis $e_1,\dots,e_n$ of L then there are polynomials P and Q in $K[x_1,\dots,x_n]$ such that*

$$\mathrm{Tr}_{L/K}(c_1e_1+\cdots+c_ne_n) = P(c_1,\dots,c_n), \quad \mathrm{N}_{L/K}(c_1e_1+\cdots+c_ne_n) = Q(c_1,\dots,c_n)$$

for all $c_i \in K$. More specifically, P is a homogeneous polynomial of degree 1 or is identically 0 and Q is a homogeneous polynomial of degree n.

Proof. Since the trace is K-linear,

$$\mathrm{Tr}_{L/K}(c_1e_1+\cdots+c_ne_n) = c_1\mathrm{Tr}_{L/K}(e_1)+\cdots+c_n\mathrm{Tr}_{L/K}(e_n) = P(c_1,\dots,c_n),$$

where $P(x_1,\dots,x_n) = \sum_{i=1}^n \mathrm{Tr}_{L/K}(e_i)x_i$. This either has degree 1 or (if each $\mathrm{Tr}_{L/K}(e_i)$ is 0) is identically zero.

For the norm,

$$\mathrm{N}_{L/K}(c_1e_1+\cdots+c_ne_n) = \det(m_{c_1e_1+\cdots+c_ne_n}) = \det(c_1m_{e_1}+\cdots+c_nm_{e_n}).$$

For indeterminates $x_1,\dots,x_n$, the determinant

$$Q(x_1,\dots,x_n) := \det(x_1[m_{e_1}]+\cdots+x_n[m_{e_n}])$$

is a homogeneous polynomial in $K[x_1,\dots,x_n]$ of degree n from the expansion formula for determinants as a sum of products, since each entry of the matrix $x_1[m_{e_1}]+\cdots+x_n[m_{e_n}]$ is a K-linear combination of $x_1,\dots,x_n$, hence is a homogeneous polynomial of degree 1. A product of n homogeneous polynomials of degree 1 is a homogeneous polynomial of degree n and a sum of homogeneous polynomials of degree n is a homogeneous polynomial of degree n or is 0. The polynomial $Q(x_1,\dots,x_n)$ is not 0 since $Q(a_1,\dots,a_n)=1$ where $\sum_{i=1}^n a_ie_i = 1$, so Q is homogeneous of degree n. Substituting c_i for x_i shows $\mathrm{N}_{L/K}(\sum_{i=1}^n c_ie_i) = Q(c_1,\dots,c_n)$ for all $c_i \in K$. □

Example 3.2. For the extension $\mathbf{Q}(\gamma)/\mathbf{Q}$ where $\gamma^3-\gamma-1=0$, using basis $\{1,\gamma,\gamma^2\}$,

$$x_1[m_1]+x_2[m_\gamma]+x_3[m_{\gamma^2}] = \begin{pmatrix} x_1 & x_3 & x_2 \\ x_2 & x_1+x_3 & x_2+x_3 \\ x_3 & x_2 & x_1+x_3 \end{pmatrix}.$$

Each entry of this matrix is a homogeneous polynomial of degree 1 in the x_i's. The trace is the sum of the terms along the main diagonal, and is homogeneous of degree 1 in the x_i's: it is $3x_1+2x_3$. The determinant is, up to signs, a sum of products of one term from each row and column, such as $x_1(x_1+x_3)^2$ using the main diagonal, and these terms are all homogeneous of degree 3. In full the norm is

$$x_1^3+x_2^3+x_3^3+2x_1^2x_3+x_1x_3^2-x_1x_2^2-x_2x_3^2-3x_1x_2x_3.$$

4. Transitivity of the Trace and Norm

If $L/F/K$ is a tower of finite extensions then a basis for L/K can be computed using bases for L/F and F/K: if $\{e_i\}$ is a basis of L/F and $\{f_j\}$ is a basis of F/K, then $\{e_i f_j\}$ is a basis of L/K. In a similar spirit, traces and norms can be calculated for L/K as a composition of traces and norms, respectively, for L/F and F/K.

Theorem 4.1. *Let $L/F/K$ be finite extensions. For $\alpha \in L$, $\mathrm{Tr}_{L/K}(\alpha) = \mathrm{Tr}_{F/K}(\mathrm{Tr}_{L/F}(\alpha))$.*

Remark 4.2. Don't write the right side as $\mathrm{Tr}_{L/F}(\mathrm{Tr}_{F/K}(\alpha))$, which *makes no sense*: the first map applied to α should go from L down to F and the second map should go from F down to K, not the other way around.

Proof. Let $m = [L : F]$ and $d = [F : K]$, as in the field diagram below.

$$\begin{array}{c} L \\ \Big|\,{\scriptstyle m} \\ F \\ \Big|\,{\scriptstyle d} \\ K \end{array}$$

To prove transitivity of the trace, let $\{e_1, \ldots, e_m\}$ be an F-basis of L and $\{f_1, \ldots, f_d\}$ be a K-basis of F. Then a K-basis of L is

$$\{e_1 f_1, \ldots, e_1 f_d, \ldots, e_m f_1, \ldots, e_m f_d\}.$$

For $\alpha \in L$, let

$$\alpha e_j = \sum_{i=1}^{m} c_{ij} e_i, \quad c_{ij} f_s = \sum_{r=1}^{d} b_{ijrs} f_r,$$

for $c_{ij} \in F$ and $b_{ijrs} \in K$. Thus $\alpha(e_j f_s) = \sum_i \sum_r b_{ijrs} e_i f_r$. Using the above bases for L/F, F/K, and L/K, we have

$$[m_\alpha]_{L/F} = (c_{ij}), \quad [m_{c_{ij}}]_{F/K} = (b_{ijrs}), \quad [m_\alpha]_{L/K} = ([m_{c_{ij}}]_{F/K}),$$

where the field extension in the subscript indicates what extension is being used for that matrix. The last matrix is a block matrix. Using these matrices,

$$\begin{aligned} \mathrm{Tr}_{F/K}(\mathrm{Tr}_{L/F}(\alpha)) &= \mathrm{Tr}_{F/K}\left(\sum_i c_{ii}\right) \\ &= \sum_i \mathrm{Tr}_{F/K}(c_{ii}) \\ &= \sum_i \sum_r b_{iirr} \\ &= \mathrm{Tr}_{L/K}(\alpha). \end{aligned}$$

□

Theorem 4.3. *Let $L/F/K$ be finite extensions. For $\alpha \in L$, $\mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{F/K}(\mathrm{N}_{L/F}(\alpha))$.*

Proof. The proof of transitivity of the trace was essentially a straightforward calculation. By comparison, the proof of transitivity of the norm is more delicate.[1]

[1]When L/K is Galois, we'll see another proof of Theorem 4.3 in Section 5.

The argument we will give is due to Scholl [7].

Case 1: $\alpha \in F$.

If the transitivity formula $\mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{F/K}(\mathrm{N}_{L/F}(\alpha))$ were already known, then when $\alpha \in F$ the right side is $\mathrm{N}_{F/K}(\alpha^{[L:F]}) = \mathrm{N}_{F/K}(\alpha)^{[L:F]}$, so our goal is to prove $\mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{F/K}(\alpha)^{[L:F]}$ when $\alpha \in F$.

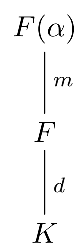
We will first show that $\chi_{\alpha,L/K}(X) = \chi_{\alpha,F/K}(X)^{[L:F]}$ when $\alpha \in F$. Both $\chi_{\alpha,L/K}(X)$ and $\chi_{\alpha,F/K}(X)$ are powers of $\pi_{\alpha,K}(X)$, where the first has degree $[L : K]$ and the second has degree $[F : K]$. Since $\chi_{\alpha,F/K}(X)^{[L:F]}$ is the power of $\pi_{\alpha,K}(X)$ with degree $[L : F][F : K] = [L : K]$, this power must be $\chi_{\alpha,L/K}(X)$. Looking at the constant terms on both sides of $\chi_{\alpha,L/K}(X) = \chi_{\alpha,F/K}(X)^{[L:F]}$, we have

$$(-1)^{[L:K]}\mathrm{N}_{L/K}(\alpha) = ((-1)^{[F:K]}\mathrm{N}_{F/K}(\alpha))^{[L:F]}.$$

The power of -1 on the right is $(-1)^{[L:F][F:K]} = (-1)^{[L:K]}$, and canceling this common power on both sides settles Case 1.

Case 2: $L = F(\alpha)$.

Let $m = [F(\alpha) : F]$ and $d = [F : K]$. The field diagram is as follows.



Let $h(X)$ be the minimal polynomial of α over F, so $h(X)$ is monic of degree m, say $h(X) = X^m + c_{m-1}X^{m-1} + \cdots + c_1X + c_0$. Then

$$\mathrm{N}_{F/K}(\mathrm{N}_{F(\alpha)/F}(\alpha)) = \mathrm{N}_{F/K}((-1)^m c_0) = (-1)^{dm}\mathrm{N}_{F/K}(c_0). \tag{4.1}$$

We will now compute $\mathrm{N}_{F(\alpha)/K}(\alpha)$ and arrive at the same value as in (4.1). Let $\{f_1, \ldots, f_d\}$ be a K-basis of F, so a K-basis of $F(\alpha)$ is

$$\{f_1, \ldots, f_d, \alpha f_1, \ldots, \alpha f_d, \ldots, \alpha^{m-1}f_1, \ldots, \alpha^{m-1}f_d\}.$$

The number $\mathrm{N}_{F(\alpha)/K}(\alpha)$ is the determinant of a matrix for multiplication by α on $F(\alpha)$ as a K-linear map. Using the above basis, the matrix is

$$\begin{pmatrix} O & O & \cdots & O & -C_0 \\ I_d & O & \cdots & O & -C_1 \\ O & I_d & \cdots & O & -C_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ O & O & \cdots & I_d & -C_{m-1} \end{pmatrix}, \tag{4.2}$$

where C_i is the $d \times d$ matrix for multiplication by c_i on F relative to the K-basis $\{f_1, \ldots, f_d\}$. This is because

$$\alpha \cdot \alpha^i f_j = \alpha^{i+1} f_j \text{ for } 0 \leq i < m-1,$$
$$\alpha \cdot \alpha^{m-1} f_j = \alpha^m f_j = -\alpha^{m-1}(c_{m-1}f_j) - \cdots - \alpha(c_1 f_j) - c_0 f_j,$$

and expressing $c_i f_j$ as a K-linear combination of $f_1, \ldots, f_d$ involves the matrix C_i.

In condensed form, the square block matrix (4.2) is

$$\begin{pmatrix} O & -C_0 \\ I_{(m-1)d} & B \end{pmatrix}, \tag{4.3}$$

where O is $d \times (m-1)d$ and B is $(m-1)d \times d$. It is left as an exercise to prove that a square block matrix of the form

$$\begin{pmatrix} O & A \\ I_N & B \end{pmatrix},$$

where O is an $M \times N$ zero matrix, A is $M \times M$ and B is $N \times M$, has determinant $(-1)^{MN} \det A$. (Use induction on N and compute the determinant by expansion along the first column.) Therefore the determinant of (4.2), by viewing it as (4.3), equals

$$(-1)^{(m-1)d^2} \det(-C_0) = (-1)^{(m-1)d}(-1)^d \det(C_0) = (-1)^{dm} \det(C_0).$$

Since C_0 is a matrix for multiplication by c_0 on F as a K-vector space, its determinant is $\mathrm{N}_{F/K}(c_0)$, so (4.3) has determinant $(-1)^{dm}\mathrm{N}_{F/K}(c_0)$. Thus $\mathrm{N}_{F(\alpha)/K}(\alpha) = (-1)^{dm}\mathrm{N}_{F/K}(c_0)$. We previously found that $\mathrm{N}_{F/K}(\mathrm{N}_{F(\alpha)/F}(\alpha))$ also equals $(-1)^{dm}\mathrm{N}_{F/K}(c_0)$, so $\mathrm{N}_{F(\alpha)/K}(\alpha) = \mathrm{N}_{F/K}(\mathrm{N}_{F(\alpha)/F}(\alpha))$.

Case 3: General situation.

When α is an element of L, insert $F(\alpha)$ into the tower of field extensions as in the following diagram.

$$\begin{array}{c} L \\ | \\ F(\alpha) \\ | \\ F \\ | \\ K \end{array}$$

Then

$$\begin{aligned} \mathrm{N}_{L/K}(\alpha) &= \mathrm{N}_{F(\alpha)/K}(\alpha)^{[L:F(\alpha)]} \quad \text{by Case 1 for } L/F(\alpha)/K \\ &= (\mathrm{N}_{F/K}(\mathrm{N}_{F(\alpha)/F}(\alpha)))^{[L:F(\alpha)]} \quad \text{by Case 2 for } F(\alpha)/F/K \\ &= \mathrm{N}_{F/K}(\mathrm{N}_{F(\alpha)/F}(\alpha)^{[L:F(\alpha)]}) \quad \text{by multiplicativity of the norm} \\ &= \mathrm{N}_{F/K}(\mathrm{N}_{L/F}(\alpha)) \quad \text{by Case 1 for } L/F(\alpha)/F. \end{aligned}$$

This completes the proof that the norm map is transitive. □

Here are references to some other proofs of the transitivity of the norm.

- Bourbaki [1, p. 548] and Jacobson [5, Sect. 7.4] prove it as a special case of a transitivity formula for determinants of block matrices with commuting blocks.
- Lang [6, Chap. VI, Sect. 5] proves it using field embeddings and inseparable degrees. See B. Conrad's handout [2] for a similar argument with more details included.
- Flanders [3, Theorem 3, Theorem 5], [4, Theorem 3] proves transitivity of the norm by a *characterization* of the norm: for finite extensions L/K with degree n, the norm map $\mathrm{N}_{L/K}$ is the unique function $f\colon L \to K$ such that (i) $f(\alpha\beta) = f(\alpha)f(\beta)$

for all α and β in L, (ii) $f(c) = c^n$ for all $c \in K$ and (iii) f is a polynomial function over K of degree at most n. (We saw the norm is such a polynomial function in Theorem 3.1.) The transitivity of the norm is an immediate consequence of these properties: for $K \subset F \subset L$, the composite map $\mathrm{N}_{F/K} \circ \mathrm{N}_{L/F}$ satisfies the three conditions that characterize $\mathrm{N}_{L/K}$ as a map from L to K.

5. The Trace and Norm for a Galois extension

Let L/K be a finite Galois extension, with Galois group $G = \mathrm{Gal}(L/K)$. We can express characteristic polynomials, traces, and norms for the extension L/K in terms of G.

Theorem 5.1. *When L/K is a finite Galois extension with Galois group G and $\alpha \in L$,*

$$\chi_{\alpha,L/K}(X) = \prod_{\sigma \in G}(X - \sigma(\alpha)).$$

In particular,

$$\mathrm{Tr}_{L/K}(\alpha) = \sum_{\sigma \in G} \sigma(\alpha), \quad \mathrm{N}_{L/K}(\alpha) = \prod_{\sigma \in G} \sigma(\alpha).$$

Proof. Let $\pi_{\alpha,K}(X)$ be the minimal polynomial of α over K, so $\chi_{\alpha,L/K}(X) = \pi_{\alpha,K}(X)^{n/d}$, where $n = [L : K]$ and $d = [K(\alpha) : K] = \deg \pi_{\alpha,K}$. From Galois theory,

$$\pi_{\alpha,K}(X) = \prod_{i=1}^{d}(X - \sigma_i(\alpha)),$$

where $\sigma_1(\alpha), \ldots, \sigma_d(\alpha)$ are all the distinct values of $\sigma(\alpha)$ as σ runs over the Galois group. For each $\sigma \in G$, $\sigma(\alpha) = \sigma_i(\alpha)$ for a unique i from 1 to d. Moreover, $\sigma(\alpha) = \sigma_i(\alpha)$ if and only if $\sigma \in \sigma_i H$, where $H = \{\tau \in G : \tau(\alpha) = \alpha\} = \mathrm{Gal}(L/K(\alpha))$. Therefore as σ runs over G, the number $\sigma_i(\alpha)$ appears as $\sigma(\alpha)$ whenever σ is in the left coset $\sigma_i H$, so $\sigma_i(\alpha)$ occurs $|H|$ times, and $|H| = [L : K(\alpha)] = [L : K]/[K(\alpha) : K] = n/d$. Therefore

$$\prod_{\sigma \in G}(X - \sigma(\alpha)) = \prod_{i=1}^{d}(X - \sigma_i(\alpha))^{n/d} = \left(\prod_{i=1}^{d}(X - \sigma_i(\alpha))\right)^{n/d} = \pi_{\alpha,K}(X)^{n/d},$$

and that power of the minimal polynomial is the characteristic polynomial. □

Example 5.2. In $\mathbf{Q}(\sqrt{d})/\mathbf{Q}$, where d is a nonsquare rational number, the two elements of the Galois group are $\sigma(a + b\sqrt{d}) = a + b\sqrt{d}$ and $\sigma(a + b\sqrt{d}) = a - b\sqrt{d}$. Then

$$\mathrm{Tr}_{\mathbf{Q}(\sqrt{d})/\mathbf{Q}}(a + b\sqrt{d}) = (a + b\sqrt{d}) + (a - b\sqrt{d}) = 2a,$$

$$\mathrm{N}_{\mathbf{Q}(\sqrt{d})/\mathbf{Q}}(a + b\sqrt{d}) = (a + b\sqrt{d})(a - b\sqrt{d}) = a^2 - db^2,$$

$$\chi_{a+b\sqrt{d},\mathbf{Q}(\sqrt{d})/\mathbf{Q}}(X) = (X - (a + b\sqrt{d}))(X - (a - b\sqrt{d})) = X^2 - 2aX + (a^2 - db^2).$$

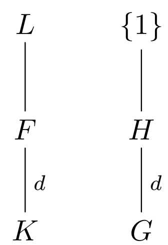
Example 5.3. For $\alpha \in \mathbf{F}_{p^n}$,

$$\mathrm{Tr}_{\mathbf{F}_{p^n}/\mathbf{F}_p}(\alpha) = \alpha + \alpha^p + \cdots + \alpha^{p^{n-1}} \text{ and } \mathrm{N}_{\mathbf{F}_{p^n}/\mathbf{F}_p}(\alpha) = \alpha\alpha^p \cdots \alpha^{p^{n-1}} = \alpha^{(p^n-1)/(p-1)}.$$

Corollary 5.4. *Let $L/F/K$ be finite extensions where L/K is Galois. For all $\alpha \in L$, $\mathrm{Tr}_{L/K}(\alpha) = \mathrm{Tr}_{F/K}(\mathrm{Tr}_{L/F}(\alpha))$ and $\mathrm{N}_{L/K}(\alpha) = \mathrm{N}_{F/K}(\mathrm{Tr}_{L/F}(\alpha))$.*

Corollary was proved already in Theorems 4.1 and 4.3 without using a Galois hypothesis. We give a second proof in this special case to give another application of Theorem 5.1. The proof of transitivity for the trace and norm this time will be by the same method, so you could say Corollary 5.4 makes transitivity of the trace look just as hard as transitivity of the norm when it really is easier.

Proof. Let $G = \mathrm{Gal}(L/K)$ and $H = \mathrm{Gal}(L/F)$.



Since L/F is Galois with Galois group H, the formulas for $\mathrm{Tr}_{L/F}(\alpha)$ and $\mathrm{N}_{L/F}(\alpha)$ from Theorem 5.1 imply

$$(5.1)\quad \mathrm{Tr}_{F/K}(\mathrm{Tr}_{L/F}(\alpha)) = \mathrm{Tr}_{F/K}\left(\sum_{\tau\in H}\tau(\alpha)\right) \text{ and } \mathrm{N}_{F/K}(\mathrm{N}_{L/F}(\alpha)) = \mathrm{N}_{F/K}\left(\prod_{\tau\in H}\tau(\alpha)\right).$$

(Be careful: the trace and norm on the right sides are *not* a sum of all $\mathrm{Tr}_{F/K}(\tau(\alpha))$ or a product of all $\mathrm{N}_{F/K}(\tau(\alpha))$ when $\alpha \notin F$, since $\tau(\alpha) \notin F$ for $\tau \in H$.)

Let $d = [F : K] = [G : H]$ and $\sigma_1, \ldots, \sigma_d$ be left coset representatives for H in G, so each element of G is $\sigma_i\tau$ for a unique coset representative σ_i and $\tau \in H$. By Theorem 5.1 and the left H-coset decomposition of G,

$$(5.2)\qquad \mathrm{Tr}_{L/K}(\alpha) = \sum_{\sigma\in G}\sigma(\alpha) = \sum_{\substack{1\le i\le d\\ \tau\in H}}\sigma_i(\tau(\alpha)) = \sum_{i=1}^{d}\sum_{\tau\in H}\sigma_i(\tau(\alpha)) = \sum_{i=1}^{d}\sigma_i\left(\sum_{\tau\in H}\tau(\alpha)\right)$$

and

$$(5.3)\qquad \mathrm{N}_{L/K}(\alpha) = \prod_{\sigma\in G}\sigma(\alpha) = \prod_{\substack{1\le i\le d\\ \tau\in H}}\sigma_i(\tau(\alpha)) = \prod_{i=1}^{d}\prod_{\tau\in G}\sigma_i(\tau(\alpha)) = \prod_{i=1}^{d}\sigma_i\left(\prod_{\tau\in H}\tau(\alpha)\right).$$

Comparing (5.1) with the trace and norm formulas in (5.2) and (5.3), which are expressed in terms of $\sum_{\tau\in H}\tau(\alpha)$ and $\prod_{\tau\in H}\tau(\alpha)$, the transitivity formulas for $\mathrm{Tr}_{L/K}(\alpha)$ and $\mathrm{N}_{L/K}(\alpha)$ will be established by showing for all $\beta \in F$ that

$$(5.4)\qquad \mathrm{Tr}_{F/K}(\beta) = \sum_{i=1}^{d}\sigma_i(\beta) \quad\text{and}\quad \mathrm{N}_{F/K}(\beta) = \prod_{i=1}^{d}\sigma_i(\beta)$$

and then using $\beta = \sum_{\tau\in H}\tau(\alpha)$ to make (5.2) and (5.3) agree with the trace and norm formulas in (5.1). To prove the formulas in (5.4) for all $\beta \in F$ we'll show

$$(5.5)\qquad \chi_{\beta,F/K}(X) = \prod_{i=1}^{d}(X - \sigma_i(\beta))$$

for all $\beta \in F$. Then comparing second-leading coefficients and constant terms on both sides produces (5.4).

To prove (5.5), let's recall what it means to say $\sigma_1, \ldots, \sigma_d$ are left coset representatives of H in G. By Galois theory, each left H-coset σH for $\sigma \in G$ consists of the elements of G that restrict in the same way to F (the fixed field of H). Therefore $\sigma_1(\beta), \ldots, \sigma_d(\beta)$ on the right side of (5.5) are *all the possible values where* β *can be sent by* G (allow some repetitions if $[F(\beta) : F] < d$), and those values are precisely the K-conjugates of β by Galois theory. At the same time, $\chi_{\beta,F/K}(X)$ on the left side of (5.5) is a power of the minimal polynomial of β over K, so its roots are also the K-conjugates of β. Therefore the roots of both sides of (5.5) are the same, so we need to show the multiplicities of each root on both sides match to get equality. To take care of that, we'll focus on a special case first.

While F/K may not be Galois, it is separable, so $F = K(\gamma)$ for some γ by the primitive element theorem. Consider (5.5) when $\beta = \gamma$. On the left side, $\chi_{\gamma,F/K}(X)$ is in $K[X]$ with degree $d = [F : K] = [K(\gamma) : K]$ and root γ, so it is the minimal polynomial of γ over K. Since it has the root γ in the Galois extension L/K, it splits completely over L, say

$$\chi_{\gamma,F/K}(X) = \prod_{i=1}^{d}(X - \gamma_i)$$

with $\gamma = \gamma_1$. The numbers $\gamma_1, \ldots, \gamma_d$ are all different ($\chi_{\gamma,F/K}(X)$ is separable) and each is a K-conjugate of γ, so these are $\sigma_1(\gamma), \ldots, \sigma_d(\gamma)$. Therefore (5.5) is true when $\beta = \gamma$, as the roots on both sides are the same numbers and they all have multiplicity 1.

For general β in F, we can write $\beta = g(\gamma)$ for some $g(X) \in K[X]$. Now use Theorem 2.1 to describe $\chi_{\beta,F/K}(X) = \chi_{g(\gamma),F/K}(X)$ in terms of $g(X)$ evaluated at roots of $\chi_{\gamma,F/K}(X)$:

$$\chi_{\beta,F/K}(X) = \chi_{g(\gamma),F/K}(X) = \prod_{i=1}^{d}(X - g(\gamma_i)) = \prod_{i=1}^{d}(X - g(\sigma_i(\gamma))).$$

Since $g(\sigma_i(\gamma)) = \sigma_i(g(\gamma)) = \sigma_i(\beta)$, $\chi_{\beta,F/K}(X) = \prod_{i=1}^{d}(X - \sigma_i(\beta))$, which proves (5.5). $\square$

References

[1] N. Bourbaki, "Algebra I," Springer-Verlag, New York, 1989.
[2] B. Conrad, Norm and Trace, `math.stanford.edu/ conrad/210BPage/handouts/normtrace.pdf`.
[3] H. Flanders, *The Norm Function of an Algebraic Field Extension*, Pacific J. Math **3** (1953), 103–113.
[4] H. Flanders, *The Norm Function of an Algebraic Field Extension, II*, Pacific J. Math **5** (1955), 519–528.
[5] N. Jacobson, "Basic Algebra I," 2nd ed., W. H. Freeman & Co., New York, 1985.
[6] S. Lang, "Algebra," 3rd ed., Addison-Wesley, New York, 1993.
[7] A. J. Scholl, Transitivity of Trace and Norm, `https://www.dpmms.cam.ac.uk/study/II/Galois/handout-3.pdf`.