# Frobenius conjugacy classes associated to $q$-linear polynomials over a finite field

Richard Pink[1]

January 4, 2013

Richard Pink ist Professor für Mathematik an der ETH Zürich. Sein Spezialgebiet ist die arithmetische Geometrie, wobei er insbesondere verschiedene wissenschaftliche Beiträge zur Arithmetik von Funktionenkörpern positiver Charakteristik geleistet hat. Seine mathematische Ausbildung erhielt er vor allem an der Universität Bonn.

Ein Grundproblem der Algebra ist die Bestimmung der Galoisgruppe eines separablen Polynoms in einer Variablen. Liegen die Koeffizienten des Polynoms in einem endlichen Körper der Kardinalität $q^n$, so ist diese Galoisgruppe erzeugt von dem Bild des Frobenius-Automorphismus $x \mapsto x^{q^n}$. Hat das Polynom zusätzlich die spezielle Form $a_0X + a_1X^q + \ldots + a_dX^{q^d}$ mit $a_0, a_d \neq 0$, so wird die Operation von Frobenius durch eine Matrix in $\mathrm{GL}_d(\mathbb{F}_q)$ repräsentiert. Der vorliegende Artikel beantwortet die Frage, welche Matrizen auf diese Weise auftreten können für gegebene $q$, $n$ und $d$. In gewissem Sinn löst dies eine Variante des "Umkehrproblems der Galoistheorie" über endlichen Körpern.

*Dept. of Mathematics, ETH Zürich, CH-8092 Zürich, Switzerland, pink@math.ethz.ch

Let $q$ be a power of a prime number $p$. Many of the wonders of algebra in characteristic $p$ are based on the fact that the binomial coefficients $\binom{q}{m}$ are divisible by $p$ for all integers $0 < m < q$. As a consequence, the map $x \mapsto x^q$ on any unitary commutative ring $R$ with $p \cdot 1_R = 0_R$ satisfies not only the multiplicativity relation $(xy)^q = x^q y^q$, but also the additivity relation $(x+y)^q = x^q + y^q$, and is therefore a ring homomorphism. This homomorphism, called *Frobenius*, is an important tool for all questions concerning finite fields of characteristic $p$.

In this short note we answer an elementary question about the action of Frobenius on the zeros of a polynomial over a finite field that seems not to have been raised before. The necessary prerequisites are nothing more than a standard two semester course in algebra.

Throughout this note we fix a finite field $\mathbb{F}_q$ of cardinality $q$, a finite field extension $k/\mathbb{F}_q$ of degree $n$, and an algebraic closure $\bar{k}$ of $k$. Let $\sigma_q\colon x \mapsto x^q$ denote the Frobenius map on $\bar{k}$. Recall that $\sigma_q^n\colon x \mapsto x^{q^n}$ acts trivially on $k$ and that the Galois group $\mathrm{Gal}(\bar{k}/k)$ is the free pro-cyclic group topologically generated by it.

Fix an integer $d \geqslant 0$, and consider a *separable $q$-linear polynomial of degree $q^d$ over $k$*, that is, a polynomial in one variable of the form

$$f(X) = \sum_{i=0}^{d} a_i X^{q^i} = a_0 X + a_1 X^q + \ldots + a_d X^{q^d}$$

with coefficients $a_i \in k$, for which $a_0$ and $a_d$ are non-zero. Since $\sigma_q\colon x \mapsto x^q$ is the identity on $\mathbb{F}_q$, the map $\bar{k} \to \bar{k}$ induced by $f$ is $\mathbb{F}_q$-linear, and so its kernel

$$V_f := \{a \in \bar{k} \mid f(a) = 0\}$$

is an $\mathbb{F}_q$-subspace of $\bar{k}$. On the other hand the formal derivative of $f$ is the non-zero constant polynomial $a_0$; hence $f$ has no multiple roots in $\bar{k}$. Thus $V_f$ has cardinality $q^d$ and therefore dimension $\dim_{\mathbb{F}_q} V_f = d$. Moreover, the fact that $\sigma_q^n$ acts trivially on $k$ implies that $V_f$ is mapped to itself under $\sigma_q^n$. Again the linearity of $\sigma_q^n$ implies that $\sigma_q^n$ induces an automorphism of the $\mathbb{F}_q$-vector space $V_f$. In any basis of $V_f$ over $\mathbb{F}_q$ this automorphism is represented by a matrix $\varphi_f \in \mathrm{GL}_d(\mathbb{F}_q)$, and the conjugacy class of $\varphi_f$ depends only on the data $(q, k, f)$.

The question we are interested in is whether anything else can be said about $\varphi_f$ if $f$ is arbitrary. In precise terms we mean:

**Question 1** *Which conjugacy classes in* $\mathrm{GL}_d(\mathbb{F}_q)$ *arise as* $\varphi_f$ *for fixed* $\mathbb{F}_q$, $k$, $d$, *and arbitrary* $f$*?*

An answer to this question helps in constructing polynomials with given Galois groups, as in Ziegler's bachelor thesis on the so-called inverse Galois problem [3].

To help the reader develop a feeling for the situation we suggest the following special cases as warmup exercises:

**Exercise 2** *For $k = \mathbb{F}_q$ and $f(X) = X + X^q + X^{q^2}$, show that $V_f$ is contained in an extension of $k$ of degree 3 and that the associated matrix $\varphi_f$ is conjugate to $\left(\begin{smallmatrix} 0 & -1 \\ 1 & -1 \end{smallmatrix}\right)$.*

**Exercise 3** *Show that $f(X) = X^q - aX$ with $a \in k^\times$ has the associated "matrix" $\varphi_f = \alpha \in \mathrm{GL}_1(\mathbb{F}_q) = \mathbb{F}_q^\times$ if and only if $\mathrm{Norm}_{k/\mathbb{F}_q}(a) = \alpha$.*

**Exercise 4** *Show that the identity matrix in $\mathrm{GL}_d(\mathbb{F}_q)$ arises as $\varphi_f$ if and only if $d \leqslant n$.*

(For the last exercise observe that $\varphi_f$ is the identity matrix if and only if $V_f \subset k$, and apply Lemma 13. Note that the last exercise also shows that the question is non-trivial.)

Now we state our general answer to Question 1. For any matrix $\varphi \in \mathrm{GL}_d(\mathbb{F}_q)$ we let $\mathbb{F}_q[\varphi]$ denote the $\mathbb{F}_q$-subalgebra of the ring of $d \times d$-matrices that is generated by $\varphi$.

**Theorem 5** *For any $\varphi \in \mathrm{GL}_d(\mathbb{F}_q)$ and any $k/\mathbb{F}_q$ of degree $n$ the following are equivalent:*

*(a) $\mathbb{F}_q^d$ as a module over $\mathbb{F}_q[\varphi]$ is generated by $\leqslant n$ elements.*

*(b) Every eigenvalue of $\varphi$ in $\bar{k}$ has geometric multiplicity $\leqslant n$.*

*(c) There exists a separable $q$-linear polynomial $f$ over $k$ with $\varphi_f$ conjugate to $\varphi$.*

It may be worthwhile to give yet another equivalent condition in a special case:

**Corollary 6** *If $k = \mathbb{F}_q$, the conditions in Theorem 5 are also equivalent to:*

*(d) $\varphi$ is conjugate to a matrix of the following form:*

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

**Proof.** We prove that (d) is equivalent to condition (a) of Theorem 5. Since $k = \mathbb{F}_q$, we have $n = 1$; hence condition (a) means that $\mathbb{F}_q^d = \sum_{i \geqslant 0} \mathbb{F}_q \cdot \varphi^i(v)$ for some vector $v$. If this holds, let $e$ be the smallest integer $\geqslant 0$ such that $\varphi^e(v)$ is an $\mathbb{F}_q$-linear combination of the vectors $v, \varphi(v), \ldots, \varphi^{e-1}(v)$. Then the subspace $\sum_{i=0}^{e-1} \mathbb{F}_q \cdot \varphi^i(v)$ is mapped to itself under $\varphi$, so it actually contains the elements $\varphi^i(v)$ for all $i \geqslant 0$. On the other hand the vectors $v, \varphi(v), \ldots, \varphi^{e-1}(v)$ are $\mathbb{F}_q$-linearly independent by construction; hence the stated condition is equivalent to saying that these vectors form an $\mathbb{F}_q$-basis of $\mathbb{F}_q^d$. Of course this requires that $e = d$. To show that the condition is equivalent to (d), it remains to observe that the matrix of $\varphi$ associated to any basis of $\mathbb{F}_q^d$ has the indicated form if and only if that basis is $v, \varphi(v), \ldots, \varphi^{d-1}(v)$ for some vector $v$. $\square$

By Theorem 5 the matrices of the form in Corollary 6 (d) actually arise for any value of $n$. Furthermore:

**Corollary 7** *For any $k/\mathbb{F}_q$ of degree $n$ the following are equivalent:*

*(a)* $d \leqslant n$.

*(b) For every $\varphi \in \mathrm{GL}_d(\mathbb{F}_q)$ there exists a separable $q$-linear polynomial $f$ over $k$ with $\varphi_f$ conjugate to $\varphi$.*

**Proof.** By Theorem 5 the condition $d \leqslant n$ is sufficient for (b). As the identity matrix in $\mathrm{GL}_d(\mathbb{F}_q)$ satisfies condition 5 (a) if and only if $d \leqslant n$, the condition is also necessary. $\square$

Now we begin with the preparations for the proof of Theorem 5. For any positive integer $r$ we let $k_r$ denote the finite subextension of $\bar{k}$ of degree $r$ over $k$. Then $k_r/k$ is Galois, and its Galois group $\Gamma_r := \mathrm{Gal}(k_r/k)$ is cyclic of order $r$ with generator $\gamma_r := \sigma_q^n|k_r$. We are interested in the structure of $k_r$ as a representation of $\Gamma_r$ over $\mathbb{F}_q$. By general principles this is equivalent to describing $k_r$ as a module over the group ring $\mathbb{F}_q[\Gamma_r]$.

**Lemma 8** *As an $\mathbb{F}_q[\Gamma_r]$-module $k_r$ is free of rank $n$.*

**Proof.** Since $k_r/k$ is a finite Galois extension, it possesses a normal basis, i.e., there exists an element $y \in k_r$ such that the elements $\gamma(y)$ for all $\gamma \in \Gamma_r$ form a basis of $k_r$ over $k$. Let $x_1, \ldots, x_n$ be a basis of $k$ over $\mathbb{F}_q$. Then the elements $\gamma(y) \cdot x_i$ for all $\gamma \in \Gamma_r$ and $1 \leqslant i \leqslant n$ form a basis of $k_r$ over $\mathbb{F}_q$. Since the elements $\gamma \in \Gamma_r$ form a basis of $\mathbb{F}_q[\Gamma_r]$ over $\mathbb{F}_q$, it follows that $x_1, \ldots, x_n$ is a basis of $k_r$ as a free module over $\mathbb{F}_q[\Gamma_r]$. $\square$

Next, for any finite dimensional representation $W$ of $\Gamma_r$ over $\mathbb{F}_q$ let $W^* := \mathrm{Hom}_{\mathbb{F}_q}(W, \mathbb{F}_q)$ denote the dual vector space endowed with the contragredient representation of $\Gamma_r$ defined by $\Gamma_r \times W^* \to W^*$, $(\gamma, \ell) \mapsto \ell \circ \gamma^{-1}$. In the special case of the regular representation $\mathbb{F}_q[\Gamma_r]$ we obtain:

**Lemma 9** *The dual representation $\mathbb{F}_q[\Gamma_r]^*$ is isomorphic to $\mathbb{F}_q[\Gamma_r]$.*

**Proof.** This is a general fact about group rings of finite groups. Indeed, by direct calculation one can show that the element $\ell \in \mathbb{F}_q[\Gamma_r]^*$ defined by $\sum_\gamma \alpha_\gamma \gamma \mapsto \alpha_1$ is a basis of $\mathbb{F}_q[\Gamma_r]^*$ as a free module of rank 1 over $\mathbb{F}_q[\Gamma_r]$. $\square$

**Lemma 10** *For any finite dimensional $\mathbb{F}_q[\Gamma_r]$-module $W$ the following are equivalent:*

*(a) $W$ is generated by $\leqslant n$ elements.*

*(b) Every eigenvalue of $\gamma_r$ on $W \otimes_k \bar{k}$ has geometric multiplicity $\leqslant n$.*

*(c) Every eigenvalue of $\gamma_r$ on $W^* \otimes_k \bar{k}$ has geometric multiplicity $\leqslant n$.*

*(d) $W^*$ is generated by $\leqslant n$ elements.*

**Proof.** These equivalences are special properties of representations of cyclic groups. We deduce them from properties of the Jordan normal form in the guise of modules over the polynomial ring $\mathbb{F}_q[X]$.

First, we view $W$ as a module over the polynomial ring $R := \mathbb{F}_q[X]$ such that $\sum_i a_i X^i$ acts as $\sum_i a_i \gamma_r^i$. By the elementary divisor theorem there exist a non-negative integer $m$ and non-constant monic polynomials $P_i \in R$ for all $1 \leqslant i \leqslant m$ such that $P_i$ divides $P_{i+1}$ for all $1 \leqslant i < m$ and that $W \cong \bigoplus_{i=1}^m R/RP_i$. Clearly $W$ is then generated by $m$ elements. Conversely, any irreducible factor $P$ of $P_1$ divides every $P_i$; hence there exists a surjection $W \twoheadrightarrow \bigoplus_{i=1}^m R/RP \cong (R/RP)^m$. The latter is a vector space of dimension $m$ over the residue field $R/RP$; hence it cannot be generated by fewer than $m$ elements. Together it follows that $m$ is the minimal number of generators of $W$ as an $R$-module, or equivalently as a module over $\mathbb{F}_q[\Gamma_r]$. Thus (a) is equivalent to $m \leqslant n$.

Secondly, every $P_i$ divides $P_m$; hence the minimal polynomial of $\gamma_r$ as an endomorphism of $W$ is $P_m$; and so the eigenvalues of $\gamma_r$ on $W \otimes_k \bar{k}$ are precisely the roots of $P_m$. Write $P_m(X) = \prod_{j=1}^s (X - \alpha_j)^{\mu_{m,j}}$ with distinct $\alpha_1, \ldots, \alpha_s \in \bar{k}$ and multiplicities $\mu_{m,j} \geqslant 1$. Since each $P_i$ divides $P_m$, we can also write $P_i(X) = \prod_{j=1}^s (X - \alpha_j)^{\mu_{i,j}}$ with multiplicities $\mu_{i,j} \geqslant 0$. By the Chinese remainder theorem we then have

$$W \otimes_k \bar{k} \;\cong\; \bigoplus_{i=1}^m \bar{k}[X]/\bar{k}[X]P_i \;\cong\; \bigoplus_{i=1}^m \bigoplus_{j=1}^s \bar{k}[X]/\bar{k}[X](X - \alpha_j)^{\mu_{i,j}}$$

as a module over $\bar{k}[X]$. For any $1 \leqslant j \leqslant s$, the geometric multiplicity of the eigenvalue $\alpha_j$ on $\bar{k}[X]/\bar{k}[X](X - \alpha_j)^{\mu_{i,j}}$ is 1 if $\mu_{i,j} \geqslant 1$, and 0 otherwise. The geometric multiplicity of $\alpha_j$ on $W \otimes_k \bar{k}$ is therefore the number of indices $1 \leqslant i \leqslant m$ with $\mu_{i,j} > 0$. Of course this number is always $\leqslant m$. Conversely, at least one of the eigenvalues is a root of the non-constant polynomial $P_1$ and hence of every $P_i$. The geometric multiplicity of this eigenvalue is therefore equal to $m$, and together it follows that $m$ is the maximum of the geometric multiplicities of all eigenvalues of $\gamma_r$ on $W \otimes_k \bar{k}$. Thus (b) is equivalent to $m \leqslant n$.

Thirdly, the above decomposition of $W \otimes_k \bar{k}$ induces a decomposition

$$W^* \otimes_k \bar{k} \;\cong\; \bigoplus_{i=1}^m (\bar{k}[X]/\bar{k}[X]P_i)^* \;\cong\; \bigoplus_{i=1}^m \bigoplus_{j=1}^s (\bar{k}[X]/\bar{k}[X](X - \alpha_j)^{\mu_{i,j}})^*,$$

where the dual vector spaces in the middle and on the right hand side are taken over $\bar{k}$. This decomposition is invariant under the natural endomorphism induced by $\gamma_r^* \colon W^* \to W^*$, $\ell \mapsto \ell \circ \gamma_r$. But each non-zero summand $\bar{k}[X]/\bar{k}[X](X - \alpha_j)^{\mu_{i,j}}$ corresponds to a single indecomposable Jordan block of $\gamma_r$ on $W \otimes_k \bar{k}$ with eigenvalue $\alpha_j$; hence its dual corresponds to an indecomposable Jordan block of $\gamma_r^*$ on $W^* \otimes_k \bar{k}$ with the same eigenvalue $\alpha_j$. Moreover, since the contragredient representation on $W^*$ is defined by letting $\gamma_r$ act through $(\gamma_r^*)^{-1}$, it follows that each non-zero $(\bar{k}[X]/\bar{k}[X](X - \alpha_j)^{\mu_{i,j}})^*$ corresponds to an indecomposable Jordan block of the contragredient action of $\gamma_r$ on $W^* \otimes_k \bar{k}$ with the eigenvalue $\alpha_j^{-1}$. Thus $m$ is also the maximum of the geometric multiplicities of all eigenvalues of $\gamma_r$ in its contragredient action on $W^* \otimes_k \bar{k}$. Thus (c) is equivalent to $m \leqslant n$.

The above three characterizations of $m$ already prove the equivalences (a)$\Leftrightarrow$(b)$\Leftrightarrow$(c). Applying the equivalence (a)$\Leftrightarrow$(b) to $W^*$ in place of $W$ also shows (c)$\Leftrightarrow$(d). This finishes the proof of Lemma 10. $\square$

**Lemma 11** *The conditions in Lemma 10 are also equivalent to:*

*(e) There exists an injective homomorphism of $\mathbb{F}_q[\Gamma_r]$-modules $W \hookrightarrow k_r$.*

**Proof.** The condition (d) of Lemma 10 is equivalent to saying that there exists a surjective homomorphism of $\mathbb{F}_q[\Gamma_r]$-modules $\mathbb{F}_q[\Gamma_r]^n \twoheadrightarrow W^*$. Since Lemmas 8 and 9 provide isomorphisms of $\mathbb{F}_q[\Gamma_r]$-modules

$$k_r^* \;\cong\; (\mathbb{F}_q[\Gamma_r]^n)^* \;\cong\; (\mathbb{F}_q[\Gamma_r]^*)^n \;\cong\; \mathbb{F}_q[\Gamma_r]^n,$$

this amounts to giving a surjective homomorphism of $\mathbb{F}_q[\Gamma_r]$-modules $k_r^* \twoheadrightarrow W^*$. By duality any such homomorphism corresponds to an injective homomorphism of $\mathbb{F}_q[\Gamma_r]$-modules $W \hookrightarrow k_r$, and vice versa. Thus (d) is equivalent to (e), as desired. $\square$

To prove Theorem 5 we will apply the above results in the special case that $r$ is the order of the finite group $\mathrm{GL}_d(\mathbb{F}_q)$. With this choice we have:

**Lemma 12** *Any $\sigma_q^n$-invariant $\mathbb{F}_q$-subspace $U \subset \bar{k}$ of dimension $d$ is contained in $k_r$.*

**Proof.** By Lagrange the $r$-th power of any element of $\mathrm{GL}_d(\mathbb{F}_q)$ is the identity matrix. Thus the power $\sigma_q^{nr}$ acts trivially on $U$. But by Galois theory the field of fixed points of $\sigma_q^{nr}$ on $\bar{k}$ is just $k_r$; hence we have $U \subset k_r$, as desired. $\square$

As a final ingredient, the following lemma concerns the passage back from $V_f$ to $f$:

**Lemma 13** *For every finite dimensional $\sigma_q^n$-invariant $\mathbb{F}_q$-subspace $U \subset \bar{k}$ there exists a separable $q$-linear polynomial $f$ over $k$ with $V_f = U$.*

**Proof.** Since $U$ is a finite set, we can form the polynomial $f(X) := \prod_{u\in U}(X-u) \in \bar{k}[X]$, which by construction is separable with set of zeros $U$. Moreover, as $U$ is invariant under $\sigma_q^n$, so is $f$; hence $f$ already lies in $k[X]$. That $f$ is $q$-linear follows from its explicit description in terms of the Moore determinant from [2, Statement III] or [1, Lemma 1.3.6]. $\square$

**Proof of Theorem 5.** Consider any matrix $\varphi \in \mathrm{GL}_d(\mathbb{F}_q)$. Then by the choice of $r$ and Lagrange's theorem the $r$-th power $\varphi^r$ is the identity matrix. Thus $W := \mathbb{F}_q^d$ carries a unique representation of the cyclic group $\Gamma_r$ such that $\gamma_r$ acts as $\varphi$. The equivalence (a)$\Leftrightarrow$(b) in Theorem 5 thus follows from the equivalence (a)$\Leftrightarrow$(b) in Lemma 10. By Lemma 11 these conditions are also equivalent to the existence of an injective homomorphism of $\mathbb{F}_q[\Gamma_r]$-modules $W \hookrightarrow k_r$. Giving such a homomorphism amounts to giving a $\gamma_r$-invariant $\mathbb{F}_q$-subspace $U \subset k_r$ and an isomorphism of $\mathbb{F}_q$-vector spaces $i\colon W \xrightarrow{\sim} U$ satisfying $i\circ\gamma_r = \gamma_r\circ i$.

By the definition of the actions of $\gamma_r$ the last relation is equivalent to $i \circ \varphi = \sigma_q^n \circ i$. By Lemma 12 such data is therefore the same as giving a $\sigma_q^n$-invariant $\mathbb{F}_q$-subspace $U \subset \bar{k}$ and an isomorphism of $\mathbb{F}_q$-vector spaces $i\colon W \xrightarrow{\sim} U$ satisfying $i \circ \varphi = \sigma_q^n \circ i$.

As explained above, the set of zeros $V_f$ of any separable $q$-linear polynomial $f$ over $k$ is a finite dimensional $\sigma_q^n$-invariant $\mathbb{F}_q$-subspace of $\bar{k}$. Lemma 13 asserts that, conversely, every finite dimensional $\sigma_q^n$-invariant $\mathbb{F}_q$-subspace of $\bar{k}$ arises in this way. Giving the above data is therefore equivalent to giving a separable $q$-linear polynomial $f$ over $k$ and an isomorphism of $\mathbb{F}_q$-vector spaces $i\colon W \xrightarrow{\sim} V_f$ satisfying $i \circ \varphi = \sigma_q^n \circ i$. But the existence of such an isomorphism $i$ means that $\dim_{\mathbb{F}_q} V_f = d$ and that $\varphi$ represents the conjugacy class of Frobenius associated to $f$, in other words, that $\varphi_f$ is conjugate to $\varphi$. Thus altogether we find that the conditions (a) and (b) of Theorem 5 are also equivalent to condition (c), and we are done. $\square$

# References


[1] Goss, D.: *Basic structures in function field arithmetic.* Springer-Verlag, 1996.

[2] Moore, E. H.: A two-fold generalization of Fermat's theorem. *Amer. Math. Soc. Bull.* **2** no.7 (1896) 189-199.

[3] Ziegler, P.: *Das Umkehrproblem der Galoistheorie.* Bachelor thesis ETH Zürich, May 2008, 27p. See:
`http://www.math.ethz.ch/~pink/Theses/2008-Bachelor-Paul-Ziegler.pdf`