The Lang Map

January 13, 2004

The Lang-Steinberg theorem on the surjectivity of the Lang map is fundamental. See:

Serge Lang, Algebraic groups over finite fields, *Amer. J. Math.* **78** (1956), 555–563.

T. Springer and R. Steinberg, Conjugacy classes, in *1970 Seminar on Algebraic Groups and Related Finite Groups* (The Institute for Advanced Study, Princeton, N.J., 1968/69) pages 167–266, Lecture Notes in Mathematics **131** (Springer Verlag).

The theorem asserts that if G is a connected algebraic group over an algebraically closed field K and $\varphi : G \longrightarrow G$ is an automorphism whose group of fixed points is finite, then $L(x) = x^{-1}\varphi(x)$ is a surjective map. In the original form proved by Lang, K is the algebraic closure of a finite field, G is defined over $\mathbb{F}_q$ and φ is the Frobenius map, which raises the coordinates of a point in G to the q-th power. We will prove a weakened form applicable when G is $\mathrm{GL}(n)$, a parabolic subgroup of $\mathrm{GL}(n)$, or the centralizer of an element in one of these groups. These are precisely the groups we will need for Shintani's theorem. By proving a weaker theorem, we reduce the prerequisites from algebraic geometry.

We will need a couple of standard facts from commutative algebra.

Proposition 1 *Let $A \subseteq B$ be commutative rings. Assume that B is integral over A. Let $\mathfrak{p}$ be a prime ideal of A. Then there exists a prime ideal $\mathfrak{P}$ of B such that $\mathfrak{P} \cap A = \mathfrak{p}$. If $\mathfrak{p}$ is maximal, so is $\mathfrak{P}$.*

Proof See Lang's *Algebra*, third edition, Propositions 1.10 and 1.11 on page 339, Rotman, *Advanced Modern Algebra* Theorem 11.55 and Corollary 11.52 on p. 928, or Atiyah and Macdonald, Theorem 5.10 on p. 62. $\square$

If E/F are fields, a subset S of E of is called *algebraically independent* if $P(x_1, \cdots, x_n) = 0$ with P a polynomial in F and $x_1, \cdots, x_n$ distinct elements of S implies that $P = 0$ as a polynomial. It is well known that maximal algebraic independent sets all have the same cardinality, which is the *transcendence degree* of E/F. See Lang's *Algebra*, Section VIII.1, or Rotman, *Advanced Modern Algebra*, p. 365. A maximal algebraically independent set is called a *transcendence basis* of E over F.

Let K be an algebraically closed field, which will be fixed. If I is any ideal in the polynomial ring $K[X_1, \cdots, X_n]$, let

$$V(I) = \{x = (x_1, \cdots, x_n) \in K^n \mid f(x) = 0 \text{ for all } f \in I\}.$$

The *Zariski topology* is the topology in which the closed sets are the sets $V(I)$. This defines a topology since $V(I + J) = V(I) \cap V(J)$, and

$$V(I_1 \cdots I_n) = V(I_1) \cup \cdots \cup V(I_n), \qquad V\left(\sum I_j\right) = \bigcap V(I_j).$$

Let $\theta \in K[X_1, \cdots, X_n]$. The set $E_\theta = \{x \in K^n \mid \theta(x) \neq 0\}$ is called a *principal open set*. If $E = E_\theta \subseteq K^n$ is a principal open set, let $\mathcal{O}(E) = K[X_1, \cdots, X_n, 1/\theta]$. Elements of $\mathcal{O}(E)$ may be regarded as functions on E.

Theorem 1 (Nullstellensatz) *Let E be a principal open set, and let $\mathfrak{P}$ be a maximal ideal of $\mathcal{O}(E)$. Then there exists a unique point $x \in E$ such that*

$$\mathfrak{P} = \{f \in \mathcal{O}(E) \mid f(x) = 0\}. \tag{1}$$

Proof The Nullstellensatz as it is often stated asserts that if $\mathfrak{M}$ is a maximal ideal of $K[X_1, \cdots, X_n]$ then there is a unique point $x \in K^n$ such that $\mathfrak{M}$ is the set of all polynomials vanishing at x. We will assume this statement. See Lang, Theorem 1.4 on p. 380 or Rotman, *Advanced Modern Algebra*, Theorem 6.101 on p. 385.

We have a surjective ring homomorphism

$$\lambda : K[X_1, \cdots, X_{n+1}] \longmapsto \mathcal{O}(E)$$

in which

$$\psi(X_1, \cdots, X_n, X_{n+1}) \longmapsto \psi(X_1, \cdots, X_n, 1/\theta).$$

Let $\mathfrak{M}$ be the ideal of $K[X_1, \cdots, X_{n+1}]$ which is the preimage of $\mathfrak{P}$ under this map. It is of course a maximal ideal, so it consists of all polynomials vanishing at some $(x_1, \cdots, x_{n+1}) \in K^{n+1}$. Now λ maps

$$\theta(X_1, \cdots, X_n)X_{n+1} - 1$$

to zero, so this ideal is in $\mathfrak{M}$; therefore $\theta(x_1, \cdots, x_n)x_{n+1} = 1$ and so

$$\theta(x_1, \cdots, x_n) \neq 0,$$

that is, $x = (x_1, \cdots, x_n) \in E$, and moreover $x_{n+1} = 1/\theta(x_1, \cdots, x_n)$.

To prove (1) it is sufficient to show that if $f \in \mathfrak{P}$ then $f(x) = 0$, since this will prove one inclusion, and $\mathfrak{P}$ is maximal. By definition of $\mathcal{O}(E)$, every element equals λF where F is some polynomial in $n+1$ variables. Writing $f = \lambda F$, we have $F \in \mathfrak{M}$, so

$$0 = F(x_1, \cdots, x_n, x_{n+1}) = F\left(x_1, \cdots, x_n, \frac{1}{\theta(x_1, \cdots, x_n)}\right) = f(x_1, \cdots, x_n).$$

$\square$

Theorem 2 *Let E be a principal open set in K^n, and let $f_1, \cdots, f_n \in \mathcal{O}(E)$. Let $f : E \longrightarrow K^n$ be the map*

$$f(x) = (f_1(x_1, \cdots, x_n), \cdots f_n(x_1, \cdots, x_n)), \qquad x = (x_1, \cdots, x_n).$$

Assume that the Jacobian

$$J(f) = \det\left(\frac{\partial f_i}{\partial x_j}\right)_{i,j}$$

is not identically zero. Then the functions $f_1, \cdots, f_n$ are algebraically independent over K, and the image of f contains a nonzero Zariski open subset.

Note: Any two nonzero Zariski open sets have nonzero intersection, so this means that the image of f is dense in K^n.

Proof First let us show that $f_1, \cdots, f_n$ are algebraically independent. If not, there exists a nonzero P in the polynomial ring $K[X_1, \cdots, X_n]$ such that $P(f_1, \cdots, f_n) = 0$ identically. We assume that P is of minimal degree. Let $P_i = \partial P/\partial X_i$ be the partial derivatives of P. By the chain rule we have

$$\sum_j P_j(f_1, \cdots, f_n)\,\frac{\partial f_j}{dx_i} = 0.$$

Since $J(f)$ is not identically zero, this means that on the principal open set determined by the nonvanishing of $J(f)$, partial derivative $P_j(f_1, \cdots, f_n) = 0$. This means that each $P_j = 0$. Since P is nonzero and therefore nonconstant, the vanishing of all its partial derivatives implies that the characteristic of K is a positive prime p and that the polynomial $P(X_1, \cdots, X_n) = Q(X_1^p, \cdots, X_n^p)$ for some polynomial Q. Noting that the Frobenius map is a bijection $K \longrightarrow K$, applying its inverse to each coefficient of Q gives another polynomial $R(X_1, \cdots, X_n)$ such that $P(X_1, \cdots X_n) = R(X_1, \cdots, X_n)^p$. Now $P(f_1, \cdots, f_n) = 0$ implies that $R(f_1, \cdots, f_n) = 0$. But the degree of R is less than the degree of P, contradicting the minimality of P.

Since the f_i are algebraically independent, they are a transcendence basis of the field $K(X_1, \cdots, X_n)$ of fractions of the polynomial ring $K[X_1, \cdots, X_n]$. Therefore $K(X_1, \cdots, X_n)$ is algebraic over $K(f_1, \cdots, f_n)$.

Since E is a principal open set, $E = E_\theta$ for some $\theta = \theta(X_1, \cdots, X_n) \in K[X_1, \cdots, X_n]$. Then $\mathcal{O}(E) = K[X_1, \cdots, X_n, \theta^{-1}]$. Since the f_i are a transcendence basis, each X_i is algebraic over the field $K(f_1, \cdots, f_n)$ of fractions of $K[f_1, \cdots, f_n]$, so we have polynomial equations

$$X_i^{r_i} + \varphi_{i,r_i-1}(f_1, \cdots, f_n)X^{r_i-1} + \cdots + \varphi_{i,0}(f_1, \cdots, f_n) = 0,$$

where $\varphi_{i,j}$ is a rational function in the f_i; and similarly θ^{-1} is a root of a monic polynomial with coefficients in $K(f_1, \cdots, f_n)$. Let $g = g(f_1, \cdots, f_n)$ be the least common multiple of the denominators of all of these coefficients $\varphi_{i,j}$, and also of the coefficients in the monic polynomial satisfied by θ^{-1}. Then $X_1, \cdots, X_n$ and θ^{-1} are all integral over the ring $K[f_1, \cdots, f_n, g^{-1}]$, since this ring contains the coefficients of all these polynomials. We now claim that the image of f contains the principal open set E_g. Since the f_i are algebraically independent, the ring $K[f_1, \cdots, f_n]$ is a polynomial ring. Therefore, given $y = (y_1, \cdots, y_n) \in E_g$, there is a homomorphism $K[f_1, \cdots, f_n] \longrightarrow K$ in which $f_i \longmapsto y_i$. In this homomorphism, $g = g(f_1, \cdots, f_n) \longmapsto g(y_1, \cdots, y_n) \neq 0$ because $y \in E_g$. Thus the homomorphism extends to a homomorphism $K[f_1, \cdots, f_n, g^{-1}] \longrightarrow K$. Let $\mathfrak{p}$ be the kernel of this homorphism. It is a maximal ideal, since K is a field. By Proposition 1, since $K[X_1, \cdots, X_n, \theta^{-1}, g^{-1}] = K[X_1, \cdots, X_n, (\theta g)^{-1}]$ is integral over $K[f_1, \cdots, f_n, g^{-1}]$, there exists a maximal ideal $\mathfrak{P}$ of $K[X_1, \cdots, X_n, (\theta g)^{-1}]$ such that $\mathfrak{P} \cap K[f_1, \cdots, f_n, g^{-1}] = \mathfrak{p}$. By the Nullstellensatz, there exists a point $x \in E_{\theta g} \subseteq E_\theta$ such that $\mathfrak{P}$ is the set of elements of $K[X_1, \cdots, X_n, (\theta g)^{-1}]$ which vanish at x. Each $f_i - y_i \in \mathfrak{p} \subset \mathfrak{P}$, so this polynomial vanishes at x, i.e. $f(y) = x$. $\square$

If $E \subset K^n$ is a Zariski open set, and $f : E \longrightarrow K^m$ is a polynomial map, and $x \in K^n$ let $df(x) : K^n \longrightarrow K^m$ be the linear map with matrix $(\partial f_i/dX_j)(x)$. This map is called the *differential* of f.

Let $F = \mathbb{F}_q$, and let K be the algebraic closure of F. If $\mathrm{Mat}_n(K)$, let $x^{(q)}$ denote the image of x under the *Frobenius endomorphism*, which is the result of raising every element of x to the q-th power. The *Lang map* $L : \mathrm{GL}(n, K) \longrightarrow \mathrm{GL}(n, K)$ is defined by $L(x) = x^{-1}x^{(q)}$.

Theorem 3 *Let A be an $\mathbb{F}_q$-subalgebra of $\mathrm{Mat}_n(\mathbb{F}_q)$, and let $A(K)$ be the K-subalgebra of $\mathrm{Mat}_n(\mathbb{F}_q)$ spanned by A.*
(i) If $g \in A(K)$, and if $\det(g) \neq 0$ then $g^{-1} \in A(K)$.
(ii) Let $G(K)$ be the group of invertible elements of $A(K)$. Then the Lang map $L : G(K) \longrightarrow G(K)$ is surjective.

Proof If $\det(g) \neq 0$ we show that $g^{-1} \in A(K)$. By the Cayley-Hamilton theorem, g satisfies its characteristic polynomial, which we write

$$g^n + a_{n-1}g^{n-1} + \ldots + a_0 = 0,$$

where $a_0 = (-1)^n \det(g) \neq 0$. Then

$$g^{-1} = -a_0^{-1}(g^{n-1}a_{n-1}g^{n-2} + a_{n-2}g + \ldots + a_1)$$

is expressible as a polynomial in g, hence is in $A(K)$. This proves (i).

Since $A(K)$ is just a vector space over K we may identify it with K^n where n is its dimension. Thus it has a Zariski topology. Fix $t \in G(K)$. We will show that the image of $f_t : G(K) \longrightarrow G(K)$ given by $f_t(x) = xtx^{-(q)}$ contains a Zariski open set in $A(K)$. This will suffice, because the images of f_1 and f_t will then both be nonempty open sets, and hence intersect; if $x\,tx^{-(q)} = y\,y^{-(q)}$ then $t = L(y^{-1}x)$.

By (i), $G(K)$ is a Zariski open set in $A(K)$. Thus we can apply Theorem 2. We will show that the Jacobian of f_t is nonzero. Note that multiplying on the right by $x^{-(q)}$ multiplies each coefficient of a matrix by q-th powers of coordinate functions. Any partial derivative of such a q-th power is zero since q is a power of the characteristic, so as far as differentiation is concerned, these terms just act as constants. We see that the differential of f_t is just the linear transformation of $A(K)$ consisting of right multiplication by t, which is invertible. So its Jacobian is nonzero. $\square$

Lang proved in general that if G is a connected algebraic group defined over $\mathbb{F}_q$, then the Lang map $G(K) \longrightarrow G(K)$ is surjective. The proof is the same, but we did not prove the general case because it requires a bit more foundations— for example, if one wants to argue as in the proof of Theorem **??**, one needs to define the Zariski tangent space and discuss smoothness. Steinberg proved a further generalization, in which the Frobenius map $x \longmapsto x^{(q)}$ is replaced by an arbitrary automorphism whose group of fixed points is finite.

Let $n_1, \cdots, n_r$ be given such that $\sum_{n_i} = n$. If k is any field, then the *standard parabolic subgroup* $P(k)$ consists of all invertible matrices in $\mathrm{Mat}_n(k)$ of the form

$$\begin{pmatrix} g_1 & * & \cdots & * \\ 0 & g_2 & \cdots & * \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & g_r \end{pmatrix}, \qquad g_i \in \mathrm{Mat}_{n_i}(k). \tag{2}$$

Theorem 4 *The Lang map $G(K) \longrightarrow G(K)$ is surjective if $G(K)$ is one of the following groups:*
(i) $\mathrm{GL}(n, K)$*;*
(ii) the centralizer of an element of $\mathrm{GL}(n, K)$*;*
(iii) $P(K)$ for any standard parabolic subgroup P of $\mathrm{GL}(n)$.
(iv) The centralizer in $P(K)$ of an element of this group.

Proof In each case, the group consists of the invertible elements in some algebra. Indeed, $\mathrm{GL}(n, K)$ is the group of invertible elements in $\mathrm{Mat}_n(K)$, the centralizer of g in $\mathrm{GL}(n, K)$ is the group of invertible elements in the centralizer of g in $\mathrm{Mat}_n(K)$, the group $P(K)$ consists of the invertible elements in the algebra of matrices of the form (2), and the centralizer of $g \in P(K)$ is the group of invertible elements in the centralizer of g in the algebra of matrices of the form (2). $\square$

We temporarily drop the assumption that $F = \mathbb{F}_q$ and prove some facts about $\mathrm{GL}(n)$ using the rational canonical form. Later we will give different proofs of the same facts extending their validity to parabolic subgroups (for finite fields only) using the Lang map.

We recall the basics of the rational canonical form. If

$$q(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_0$$

is a monic polynomial, its *companion matrix* is

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

If F is a field, every matrix A in $\mathrm{Mat}_n(F)$ is conjugate to a unique matrix of the form

$$\begin{pmatrix} R_1 & & \\ & \ddots & \\ & & R_d \end{pmatrix}, \tag{3}$$

where R_i are companion matrices of polynomials p_i, and each p_i divides p_{i+1}. The polynomials p_i are called the *invariants* of A. The matrix (3) is called the *rational canonical form* of A.

The existence and uniqueness of the rational canonical form are deduced from the fact that the polynomial ring $F[X]$ in one indeterminate is a principal ideal domain. One makes the vector space F^n into an $F[X]$ module by $f \cdot v = f(A)v$, then applies the structure theorem for finitely generated torsion modules for a principal ideal domain. The module F^n is a direct sum of cyclic modules of the form $F[X]/(p_i)$, and this decomposition is uniquely determined if each p_i divides p_{i+1}.

For more proofs see Lang's *Algebra*, Section XIV.2, or Rotman, *Advanced Modern Algebra*, Section 9.2.

Theorem 5 *Let F be a field, E/F an extension field. If g_1 and g_2 are elements of* $\mathrm{GL}(n, F)$ *which are conjugate over* $\mathrm{GL}(n, E)$*, then they are already conjugate in* $\mathrm{GL}(n, F)$.

This property fails for $\mathrm{SL}(n)$. For example,

$$\begin{pmatrix} 1 & 1 \\ & 1 \end{pmatrix}, \qquad \begin{pmatrix} 1 & -1 \\ & 1 \end{pmatrix}$$

are conjugate in $\mathrm{SL}(2, \mathbb{C})$ but not in $\mathrm{SL}(2, \mathbb{R})$. We say that the two distinct $\mathrm{SL}(2, \mathbb{R})$ conjugacy classes are *amalgamated* in the larger group $\mathrm{SL}(2, \mathbb{C})$.

Proof The uniqueness of the rational canonical form means that the rational canonical form of A is unchanged if the ground field is extended, so the theorem is a clear consequence. □

Theorem 6 *Let E/F be a finite Galois extension, and let $g \in \mathrm{GL}(n,E)$ be such that for every $\sigma \in \mathrm{Gal}(E/F)$, g is conjugate to $\sigma(g)$. Then g is conjugate to an element of* $\mathrm{GL}(n,F)$.

Proof Applying σ to the invariants of g gives the invariants of $\sigma(g)$, so if these are conjugate, the invariants of g are Galois invariant, hence are polynomials with coefficients in F. Thus the rational canonical form is a matrix with coefficients in F. $\square$

Now let us give alternate proofs of these facts when $F = \mathbb{F}_q$ using the Lang map. As before, K is the algebraic closure of F.

Theorem 7 *Let $G = P$ be a parabolic subgroup of* $\mathrm{GL}(n)$. *If g_1 and g_2 are elements of $P(\mathbb{F}_q)$ which are conjugate over $P(\mathbb{F}_{q^m})$, then they are already conjugate in $P(\mathbb{F}_q)$.*

Proof Let $g_2 = hg_1h^{-1}$ with $h \in P(\mathbb{F}_{q^m})$. If $\gamma \in \mathrm{GL}(n,K)$ is in the centralizer of g_1, then $g_2 = (h\gamma^{-1})g_1(h\gamma^{-1})^{-1}$, so what we need to show is that we can find γ such that $h\gamma^{-1}$ is in $P(\mathbb{F}_q)$. What we need is for $(h\gamma^{-1}) = (h\gamma^{-1})^{(q)}$, or $L(\gamma) = L(h^{-1})$. Since the Lang map is surjective, we can find such γ. $\square$

Theorem 8 *Let $G = P$ be a parabolic subgroup of* $\mathrm{GL}(n)$. *If $g \in P(\mathbb{F}_{q^r})$ is conjugate to $g^{(q)}$, then g is conjugate in $P(\mathbb{F}_{q^r})$ to an element of $P(\mathbb{F}_q)$.*

Proof Suppose that $g^{(q)} = hgh^{-1}$. First we exhibit $w \in P(K)$ such that $wgw^{-1} \in P(\mathbb{F}_q)$. We need

$$(wgw^{-1})^{(q)} = wgw^{-1},$$

and this will be satisfied if $h^{-1} = L(w)$. Thus g and wgw^{-1}, which are both elements of $P(\mathbb{F}_{q^m})$ are conjugate in $P(K)$, and it follows from Theorem 7 that they are conjugate in $P(\mathbb{F}_{q^m})$. $\square$