

Zeitschrift: Commentarii Mathematici Helvetici
Herausgeber: Schweizerische Mathematische Gesellschaft
Band: 80 (2005)

Artikel: Optimal $SL(2)$ -homomorphisms
Autor: McNinch, George J.
DOI: <https://doi.org/10.5169/seals-60448>

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 08.05.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Optimal SL_2 -homomorphisms

George J. McNinch*

Abstract. Let G be a semisimple group over an algebraically closed field of *very good* characteristic for G . In the context of geometric invariant theory, G. Kempf and – independently – G. Rousseau have associated optimal cocharacters of G to an unstable vector in a linear G -representation. If the nilpotent element $X \in \text{Lie}(G)$ lies in the image of the differential of a homomorphism $SL_2 \rightarrow G$, we say that homomorphism is optimal for X , or simply optimal, provided that its restriction to a suitable torus of SL_2 is optimal for X in the sense of geometric invariant theory.

We show here that any two SL_2 -homomorphisms which are optimal for X are conjugate under the connected centralizer of X . This implies, for example, that there is a unique conjugacy class of *principal homomorphisms* for G . We show that the image of an optimal SL_2 -homomorphism is a *completely reducible* subgroup of G ; this is a notion defined recently by J.-P. Serre. Finally, if G is defined over the (arbitrary) subfield K of k , and if $X \in \text{Lie}(G)(K)$ is a K -rational nilpotent element with $X^{[p]} = 0$, we show that there is an optimal homomorphism for X which is defined over K .

Mathematics Subject Classification (2000). 20G15.

Keywords. Reductive group, nilpotent orbit, instability flag, completely reducible subgroup, principal homomorphism.

Contents

1	Introduction	392
2	Reductive groups	394
3	Parabolic subgroups	396
4	Springer’s isomorphisms	397
5	Frobenius twists and untwists	399
6	Nilpotent and unipotent elements	402
7	The order formula and a generalization	405
8	Optimal SL_2 -homomorphisms	408
8.1	Conjugacy of optimal SL_2 -homomorphisms	413

*Research of the author supported in part by the US National Science Foundation through DMS-0437482

8.2	Uniqueness of a principal homomorphism	415
8.3	Optimal homomorphisms over ground fields	415
8.4	Complete reducibility of optimal homomorphisms	417
8.5	Comparison with good homomorphisms	420
9	Rational elements of a nilpotent orbit defined over a ground field	421
10	Appendix: Springer isomorphisms (Jean-Pierre Serre, June 1999)	422

1. Introduction

Let G be a semisimple group over the algebraically closed field k , and assume that the characteristic of k is *very good* for G . (Actually, we consider in this paper a slightly more general class of reductive groups; see §2, where we also define *very good* primes).

Premet has recently given a conceptual proof of the Bala–Carter theorem using ideas of Kempf and of Rousseau from geometric invariant theory. An element $X \in \mathfrak{g} = \text{Lie}(G)$ is nilpotent just in case the closure of its adjoint orbit contains 0; such vectors are said to be unstable. The Hilbert–Mumford criteria says that an unstable vector for G is also unstable for certain one-dimensional sub-tori of G . This result has a more precise form due to Kempf and to Rousseau: there is a class of *optimal* cocharacters of G whose images exhibit such one dimensional sub-tori. One of the nice features of these cocharacters is that they each define the same parabolic subgroup of G ; for a nilpotent element $X \in \mathfrak{g}$, this instability parabolic is sometimes called the Jacobson–Morozov parabolic attached to X .

In his proof of the Bala–Carter Theorem in good characteristic, Pommerening constructed cocharacters associated with the nilpotent element $X \in \mathfrak{g}$; see [Ja04] for more on this notion, and see §6 below. Using the results of Kempf, Rousseau, and Premet, one finds (cf. [Mc04]) that the cocharacters associated with a nilpotent $X \in \mathfrak{g}$ are optimal, and that any optimal cocharacter Ψ for X such that $X \in \mathfrak{g}(\Psi; 2)$ is associated with X in Pommerening’s sense.

In this paper, we show that the notion of optimal cocharacters is important in the study of subgroups of G . We say that a homomorphism $\phi: \text{SL}_2 \rightarrow G$ is optimal provided that the restriction of ϕ to the standard maximal torus of SL_2 is a cocharacter associated to the nilpotent element

$$X = d\phi\left(\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\right) \in \mathfrak{g}.$$

More precisely, we say that ϕ is optimal for X .

We prove in this paper that any two optimal homomorphisms for X are conjugate by $C_G^\circ(X)$; cf. Theorem 44. This has an immediate corollary. A principal

homomorphism $\phi: \mathrm{SL}_2 \rightarrow G$ is one for which the image of $d\phi$ contains a regular nilpotent element; the conjugacy result just mentioned implies that there is a unique G -conjugacy class of principal homomorphisms.

Generalizing the notion of completely reducible representations, J.-P. Serre has defined the notion of a G -cr subgroup H of G : H is G -cr if whenever H lies in a parabolic subgroup of G , it lies in a Levi subgroup of that parabolic. We show in Theorem 52 that the image of any optimal homomorphism is G -cr. In a previous paper [Mc03], the author showed the existence of a homomorphism optimal for any p -nilpotent $X \in \mathfrak{g}$; such a homomorphism was essentially obtained (up to G -conjugacy) by base change from a morphism of group schemes defined over a valuation ring in a number field. Suppose that G is defined over the arbitrary subfield K of k . If X is a K -rational p -nilpotent element, we show in this paper that there is an optimal homomorphism ϕ for X which is defined over K ; for this we use the fact, proved in [Mc04], that some cocharacter associated with X is defined over K .

G. Seitz [Sei00] has studied homomorphisms $\phi: \mathrm{SL}_2 \rightarrow G$ with the property that all weights of a maximal torus of SL_2 on $\mathrm{Lie}(G)$ are $\leq 2p - 2$; he calls the image of such a homomorphism a good (or restricted) A_1 -subgroup. We give here a direct proof that an optimal SL_2 -homomorphism is good: we show that the weights of a cocharacter associated with a p -nilpotent element $X \in \mathfrak{g}$ are all $\leq 2p - 2$; see Proposition 30. It follows from results of Seitz that all good homomorphisms are optimal – we do not use this fact in our proofs.

We do use here a result of Seitz (see Proposition 34) to show that $(\mathrm{Ad} \circ \phi, \mathfrak{g})$ is a tilting module for SL_2 when ϕ is the optimal homomorphism obtained previous by the author [Mc03]; this fact is used to prove a unicity result Proposition 38 for certain homomorphisms $G_a \rightarrow G$ which is crucial to the proof of Theorem 44; of course, in the end one knows that $(\mathrm{Ad} \circ \phi, \mathfrak{g})$ is a tilting module for any optimal ϕ .

Seitz *loc. cit.* proved a conjugacy result for good homomorphisms analogous to the result proved here for optimal ones; he also proved that good homomorphisms are G -cr, so in some sense our results are not new. On the other hand, our proofs of conjugacy and of the G -cr property for optimal homomorphisms are free of any case analysis; we do not appeal to the classification of quasisimple groups at all. Moreover, we believe that our results on optimal homomorphisms over ground fields are new and that the ease with which they are obtained is evidence of the value of our techniques.

As further application of the methods of this paper, we include in §9 an extension of a result of Kottwitz; we prove that any nilpotent orbit which is defined over a ground field K contains a K -rational point.

Finally, the appendix contains a note of Jean-Pierre Serre concerning Springer isomorphisms.

I would like to thank Serre for allowing me to include his note on Springer isomorphisms as an appendix; I also thank him for some useful remarks on a preliminary

version of this manuscript. Moreover, I would like to extend thanks to Jens Carsten Jantzen, and to a referee, for several useful comments on the manuscript.

2. Reductive groups

We fix once and for all an algebraically closed field k ; K will be an arbitrary subfield of k , and G will be a connected, reductive algebraic group (over k) which is defined over the ground field K .

If G is quasisimple with root system R , the characteristic p of k is said to be a bad prime for R in the following circumstances: $p = 2$ is bad whenever $R \neq A_r$, $p = 3$ is bad if $R = G_2, F_4, E_6$, and $p = 5$ is bad if $R = E_8$. Otherwise, p is good. [Here is a more intrinsic definition of good prime: p is good just in case it divides no coefficient of the highest root in R].

If p is good, then p is said to be very good provided that either R is not of type A_r , or that $R = A_r$ and $r \not\equiv -1 \pmod{p}$.

If G is reductive, the isogeny theorem [Spr98, Theorem 9.6.5] yields a – not necessarily separable – central isogeny $\prod_i G_i \times T \rightarrow G$ where the G_i are quasisimple and T is a torus. The G_i are uniquely determined by G up to central isogeny, and p is good (respectively very good) for G if it is good (respectively very good) for each G_i .

The notions of good and very good primes are geometric in the sense that they depend only on G over k . Moreover, they depend only on the central isogeny class of the derived group (G, G) .

We record some facts:

Lemma 1. (1) *Let G be a quasisimple group in very good characteristic. Then the adjoint representation of G on $\mathrm{Lie}(G)$ is irreducible and self-dual.*

(2) *Let $M \leq G$ be a reductive subgroup containing a maximal torus of G . If p is good for G , then it is good for M .*

Proof. For the first assertions of (1), see [Hu95, 0.13]. (2) may be found for instance in [MS03, Proposition 16]. \square

Consider K -groups H which are direct products

$$(*) \quad H = H_1 \times S,$$

where S is a K -torus and H_1 is a connected, semisimple K -group for which the characteristic is very good. We say that the reductive K -group G is *strongly standard* if there exists a group H of the form $(*)$ and a separable K -isogeny between G and a K -Levi subgroup of H . Thus, G is separably isogenous to $M = C_H(S)$ for some

K -subtorus $S < H$; note that we do not require M to be the Levi subgroup of a K -rational parabolic subgroup.

We first observe that a strongly standard group G is *standard* in the sense of [Mc04]; this is contained in the following:

Proposition 2. *If G is a strongly standard K -group, then there is a separable K -isogeny between G and \tilde{G} where \tilde{G} is a reductive K -group satisfying the “standard hypotheses” of [Ja04, §2.9], namely:*

- (1) *the derived group of \tilde{G} is simply connected,*
- (2) *p is good for \tilde{G} , and*
- (3) *there is a \tilde{G} invariant nondegenerate bilinear form on $\text{Lie}(\tilde{G})$.*

Proof. Let $\tilde{H} = \tilde{H}_1 \times S$ where $\pi_1: \tilde{H}_1 \rightarrow H_1$ is the simply connected cover, and let $\pi = \pi_1 \times \text{id}: \tilde{H} \rightarrow H$ be the corresponding isogeny; of course, \tilde{H} and π are defined over H [KMRT, Theorem 26.7]. By assumption, $G = C_H(S)$ for some K -subtorus $S < H$. Since $\tilde{S} = \pi^{-1}(S)^\circ < \tilde{H}$ is again a K -torus, its centralizer $\tilde{G} = C_{\tilde{H}}(\tilde{S})$ is a K -Levi subgroup of \tilde{H} and $\pi|_{\tilde{G}}: \tilde{G} \rightarrow G$ is an isogeny. Now, $\text{Lie}(\tilde{G})$ is the 0-weight space of \tilde{S} on $\text{Lie}(\tilde{H})$ and $\text{Lie}(G)$ is the 0-weight space of S (and \tilde{S}) on $\text{Lie}(H)$. Since $d\pi$ is an \tilde{S} -isomorphism, it restricts to an isomorphism $d\pi|_{\text{Lie}(\tilde{G})}: \text{Lie}(\tilde{G}) \rightarrow \text{Lie}(G)$; in other words, π is a separable isogeny.

Since \tilde{G} is a Levi subgroup of \tilde{H} , its derived group \tilde{G} is simply connected, so that (1) holds. Since p is good for H , it is also good for H and for the Levi subgroups G and \tilde{G} ; see for instance [MS03, Proposition 16]. Thus (2) holds for \tilde{G} .

Finally, notice that $\text{Lie}(\tilde{H})$ is semisimple as a \tilde{H} -module and that $\text{Lie}(H')$ is a self-dual, simple H' -module whenever H' is quasi-simple in very good characteristic. It follows that there is a non-degenerate \tilde{H} -invariant bilinear form on $\text{Lie}(\tilde{H})$. This restriction of this form to the 0-weight space for \tilde{S} is again nondegenerate, and so (3) holds. [Note that the same argument gives non-degenerate invariant forms on $\text{Lie}(H)$ and $\text{Lie}(G)$.] □

Remark 3. Suppose that V is a finite dimensional vector space. Then the group $G = GL(V)$ is strongly standard. Indeed, if $\dim V \not\equiv 0 \pmod{p}$, then G is separably isogenous to $SL(V) \times G_m$, and p is very good for $SL(V)$. If $\dim V \equiv 0 \pmod{p}$, then G is isomorphic to a Levi subgroup of $H = SL(V \oplus k)$ and p is very good for H .

On the other hand, $SL(V)$ is only strongly standard when $\dim V \not\equiv 0 \pmod{p}$.

Remark 4. If G is strongly standard, there is always a *symmetric* invariant non-degenerate bilinear form on $\text{Lie}(G)$. Indeed, up to separable isogeny, G is a Levi subgroup of $T \times H$ where H is semisimple in very good characteristic. If the result

holds for H , then it holds for G ; note that any nondegenerate form on $\text{Lie}(T)$ is invariant. Thus we assume that G is semisimple in very good characteristic. For such a group, the simply connected cover is a separable isogeny so we may also assume G to be simply connected. But then G is a direct product of quasisimple groups, hence we may as well suppose that G is quasisimple in very good characteristic. In this case, the adjoint representation is a self-dual simple G -module. If $p = 2$, we are done. Otherwise, one can argue as follows: If G/\mathbb{Q} denotes the split group over \mathbb{Q} with the same root datum as G , then the adjoint representation of G/\mathbb{Q} is also simple; identifying the weight lattice of a maximal torus of G and of G/\mathbb{Q} , the adjoint representations have the “same” highest weight λ . Steinberg [St67, Lemma 79] gives a condition on λ for the invariant form to be symmetric; since this condition is independent of characteristic, and since the Killing form is symmetric on $\text{Lie}(G/\mathbb{Q})$, our claim is verified.

Proposition 5. *If G is strongly standard, then each conjugacy class and each adjoint orbit is separable. In particular, if G is defined over K , and if $g \in G(K)$ and $X \in \mathfrak{g}(K)$, then $C_G(g)$ and $C_G(X)$ are defined over K .*

Proof. Separability is [SS70, I.5.2 and I.5.6]. The fact that the centralizers are defined over K then follows from [Spr98, Proposition 12.1.2]. □

3. Parabolic subgroups

In this section, G is an arbitrary reductive group over k . The material we recall here is foundational; the lemmas from this section will be used mainly for our consideration of G -completely reducible subgroups of a reductive group G ; cf. 8.4 below.

If V is an affine variety and $f : G_m \rightarrow V$ is a morphism, we write $v = \lim_{t \rightarrow 0} f(t)$, and we say that the limit exists, if f extends to a morphism $\tilde{f} : k \rightarrow V$ with $\tilde{f}(0) = v$. If γ is any cocharacter of G , then

$$P_G(\gamma) = P(\gamma) = \{x \in G \mid \lim_{t \rightarrow 0} \gamma(t)x\gamma(t^{-1}) \text{ exists}\}$$

is a parabolic subgroup of G whose Lie algebra is $\mathfrak{p}(\gamma) = \sum_{i \geq 0} \mathfrak{g}(\gamma; i)$. Moreover, each parabolic subgroup of G has the form $P(\gamma)$ for some cocharacter γ ; for all this cf. [Spr98, 3.2.15 and 8.4.5].

We note that γ “exhibits” a Levi decomposition of $P = P(\gamma)$. Indeed, $P(\gamma)$ is the semi-direct product $Z(\gamma) \cdot U(\gamma)$, where $U(\gamma) = \{x \in P \mid \lim_{t \rightarrow 0} \gamma(t)x\gamma(t^{-1}) = 1\}$ is the unipotent radical of $P(\gamma)$, and the reductive subgroup $Z(\gamma) = C_G(\gamma(G_m))$ is a Levi factor in $P(\gamma)$; cf. [Spr98, 13.4.2].

Lemma 6. *Let P be a parabolic subgroup of G , and let T be a maximal torus of P . Then there is a cocharacter $\gamma \in X_*(T)$ with $P = P(\gamma)$.*

Proof. Since $P = P(\gamma')$ for some cocharacter γ' , this follows from the conjugacy of maximal tori in P . \square

For later use, we record:

Lemma 7. *Let $P = P(\gamma)$ be the parabolic subgroup determined by the cocharacter $\gamma \in X_*(G)$. Write $L = Z(\gamma)$ for the Levi factor of P determined by the choice of γ . If $\phi: H \rightarrow P$ is any homomorphism of algebraic groups, the rule*

$$\widehat{\phi}(x) = \lim_{s \rightarrow 0} \gamma(s)\phi(x)\gamma(s^{-1})$$

determines a homomorphism $\widehat{\phi}: H \rightarrow L$ of algebraic groups. Moreover, the tangent map $d\widehat{\phi}$ is the composite

$$\text{Lie}(H) \xrightarrow{d\phi} \text{Lie}(P) \xrightarrow{\text{pr}} \text{Lie}(L) = \text{Lie}(P)(\gamma; 0)$$

where pr is projection on the 0 weight space.

Proof. It was already observed that $P = L \cdot U$ is a semidirect product; the map

$$x \mapsto \lim_{s \rightarrow 0} \gamma(s)x\gamma(s^{-1})$$

is the projection of P on L and is thus an algebraic group homomorphism $\psi: P \rightarrow L$. The tangent map to ψ is evidently given by projection onto the 0-weight space for the image of γ , and the lemma follows. \square

Remark 8. If the cocharacter γ is defined over the ground field K , then $P = P(\gamma)$ is a K -parabolic subgroup, and the Levi factor $L = Z(\gamma)$ is defined over K . The projection $P \rightarrow L$ given by $x \mapsto \lim_{s \rightarrow 0} \gamma(s)x\gamma(s^{-1})$ is of course defined over K as well.

4. Springer’s isomorphisms

If the characteristic of k is zero, or is “sufficiently large” with respect to the group G , (some sort of) exponential map defines an equivariant isomorphism $\exp: \mathcal{N} \rightarrow \mathcal{U}$ between the nilpotent variety and the unipotent variety of G . Simple examples show the exponential to be insufficient in general, however, and in 1969, T. A. Springer [Spr69] found (the beginnings of) a good substitute. See also the outline given in

[SS70, III §3]. The unipotent variety is known always to be normal; to make Springer's work complete, one required also the normality of the nilpotent variety. Veldkamp obtained that normality for "most" p , and Demazure proved it for G satisfying our hypothesis; cf. [Ja04, 8.5]. We summarize these remarks in the following:

Proposition 9 (Springer). *Let G be a strongly standard K -reductive group, where K is any subfield of k . There is a G -equivariant isomorphism of varieties $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$ which is defined over K .*

Sketch. We just comment briefly on our assumptions on G . First, note that if G is the direct product of a torus and a semisimple group in very good characteristic, there is a separable isogeny $\tilde{G} \rightarrow G$ where \tilde{G} is the direct product of a K -torus and a simply connected semisimple K -group in (very) good characteristic. Moreover, the separable isogeny is defined over K and induces equivariant K -isomorphisms $\tilde{\mathcal{U}} \rightarrow \mathcal{U}$ and $\tilde{\mathcal{N}} \rightarrow \mathcal{N}$ (using some hopefully obvious notation); see [Mc03, Lemma 27]. Now, Springer proved the proposition holds for \tilde{G} – see the above references – and thus the result for G is true in this case.

Repeating the above argument, we may replace G by a separably isogenous group, and thus we suppose that $G = C_H(S)$, where S is a K -torus in a K -group H as in (*) of section §2; the above remarks show that there is an H -equivariant isomorphism $\Lambda_H: \mathcal{U}_H \rightarrow \mathcal{N}_H$ between the unipotent and nilpotent varieties for H . Since $\mathcal{U} = (\mathcal{U}_H)^S$ and $\mathcal{N} = (\mathcal{N}_H)^S$, it is clear that $\Lambda_H|_{\mathcal{U}}$ defines the required isomorphism for the varieties associated with G . \square

Remark 10. Suppose that $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$ is an equivariant isomorphism defined over K . If $P \leq G$ is a K -parabolic subgroup, Lemma 6 makes clear that the restriction $\Lambda|_U: U \rightarrow \text{Lie}(U)$ is a P -equivariant isomorphism. Similarly, if $L \leq G$ is a K -Levi subgroup, then $\Lambda|_{\mathcal{U}_L}: \mathcal{U}_L \rightarrow \mathcal{N}_L$ is an L -equivariant isomorphism.

The isomorphism Λ of the proposition is *quite far* from being unique; cf. the appendix of J.-P. Serre below. We summarize the result of that appendix with the following statements, which we make only in the "geometric" setting – i.e. over k rather than K .

Proposition 11 (Serre). *Let G be a strongly standard reductive k -group.*

- (1) *Fix a regular nilpotent $X \in \mathfrak{g}$. For each regular unipotent $v \in C_G(X)$, there is a unique G -equivariant isomorphism of varieties $\Lambda_v: \mathcal{U} \rightarrow \mathcal{N}$ with $\Lambda_v(v) = X$.*
- (2) *Any two G -equivariant isomorphisms $\Lambda, \Lambda': \mathcal{U} \rightarrow \mathcal{N}$ induce the same map on the finite sets of orbits.*

5. Frobenius twists and untwists

Let K' be a perfect field of characteristic $p > 0$, and let $K' \subset K$ be an arbitrary extension of K' . We fix an algebraically closed field k containing K .

In this section, algebras are always assumed to be commutative. Consider a K' -algebra A . For $r \in \mathbb{Z}$, we may consider the K' -algebra $A^{(r)}$ which coincides with A as a ring, but where each $b \in K'$ acts on $A^{(r)}$ as $b^{p^{-r}}$ does on A . For an extension field K of K' , we write $A^{(r)}/K$ and A/K for the K -algebras obtained by base-change; thus e.g. $A/K = A \otimes_{K'} K$.

Let $r \geq 0$ and let $q = p^r$. There is a K' -algebra homomorphism $F^r : A^{(r)} \rightarrow A$ given by $x \mapsto x^q$. We write $A^q = \{f^q \mid f \in A\}$; A^q is a K' -subalgebra of A , and the image of F^r coincides with A^q .

Let A be a K' -algebra and an integral domain. We clearly have:

Lemma 12. *If $r \geq 0$, and $q = p^r$, then $F^r : A^{(r)} \rightarrow A^q$ is an isomorphism of K' -algebras.*

Write $B = A/K$. Let us notice that $K[B^q] = K[A^q]$. For $r \geq 0$, consider the algebra homomorphism $F^r_{/K} : A^{(r)}/K \rightarrow K[A^q] \subset A/K$ given on pure tensors by $f \otimes \alpha \mapsto f^q \cdot \alpha$ for $f \in A^{(r)}$ and $\alpha \in K$. We have more generally

Lemma 13. *For $r \geq 0$, $F^r_{/K} : A^{(r)}/K \rightarrow K[B^q]$ is an isomorphism, where again $q = p^r$.*

Proof. We have observed already that $C = K[B^q] = K[A^q]$ is the K -algebra generated by A^q . According to the previous lemma, the image of the restriction of $F^r_{/K}$ to $A^{(r)} \otimes 1$ is the set of K -algebra generators A^q of C ; this implies that $F^r_{/K}$ is surjective.

Since A is a domain, the homomorphism $F^r : A^{(r)} \rightarrow A$ is injective. This implies the injectivity of $F^r_{/K}$ since K is flat over K' . □

Lemma 14. *Assume that A is geometrically irreducible, i.e. that A/k is a domain. Also assume A to be geometrically normal, i.e. that A/k is integrally closed in its field of fractions E . Let $q = p^r$ for $r \geq 0$, and let $f \in A/K$. Then $f \in K[A^q]$ if and only if $f \in E^q$.*

Proof. We have clearly the implication \implies . Now suppose that $f \in E^q$, say $f = g^q$ for $g \in E$. The normality of A/k shows then that $g \in A/k$. We may find $\alpha_1, \dots, \alpha_n \in k$ and elements $f_1, \dots, f_n \in A$ such that $g = \sum_{i=1}^n \alpha_i f_i$; we may assume as well that $\{f_i \mid 1 \leq i \leq n\}$ is a K' -linearly independent set. Since K' is perfect, $\{f_i^q \mid 1 \leq i \leq n\}$ is again K' -linearly independent. Since $f = g^q = \sum_{i=1}^n \alpha_i^q f_i^q \in A/K$, it follows that $\alpha_i^q \in K$ for $1 \leq i \leq n$ and the proof of \impliedby is complete. □

Remark 15. It can happen that $A_{/K}$ is a normal domain, but that $A_{/k}$ is not normal; cf. [Bo98, exerc. V.§1.23(b)].

Lemma 16. *Let X and Y be irreducible affine k -varieties, and let $f: X \rightarrow Y$ be a dominant morphism. Then the following are equivalent:*

- (a) *there is a non-empty open subset $W \subset X$ such that $df_x \neq 0$ for all $x \in W(k)$.*
- (b) *$f^*(k(Y))$ is not contained in $k(X)^p$.*

Proof. For an affine k -variety Z , let $\Omega_Z = \Omega_{k[Z]/k}$ be the module of differentials. The map $f: X \rightarrow Y$ determines a map $\phi: \Omega_Y \rightarrow \Omega_X$ of $k[Y]$ modules and – since f is dominant – a map $\psi: \Omega_{k(Y)/k} \rightarrow \Omega_{k(X)/k}$ of $k(Y)$ -vector spaces.

It follows from [Spr98, Theorem 4.3.3] that there are non-empty affine open subsets U of X and V of Y such that f restricts to a morphism $U \rightarrow V$, Ω_U is a free $k[U]$ -module of rank $\dim X$, and Ω_V is a free $k[V]$ -module of rank $\dim Y$. Now, ϕ restricts to a map $\phi|_{\Omega_V}: \Omega_V \rightarrow \Omega_U$ of $k[V]$ -modules, and it is clear that $\phi|_{\Omega_V} = 0$ if and only if $\psi = 0$ [use that $\Omega_{k(X)/k} = k(X) \otimes_{k[U]} \Omega_U$ together with the corresponding statement for Y].

Choosing bases of the free modules Ω_U and Ω_V , $\phi|_{\Omega_V}$ is given on Ω_V by a matrix M with entries in $k[U]$. For $x \in U(k)$, the map $df_x: T_x U \rightarrow T_{f(x)} V$ identifies with the map

$$\mathrm{Hom}_{k[U]}(\Omega_U, k_x) \rightarrow \mathrm{Hom}_{k[V]}(\Omega_V, k_{f(x)})$$

deduced from $\phi|_{\Omega_V}$. The open subset of U defined by the condition $M_x \neq 0$ is non-empty if and only if $\phi|_{\Omega_V} \neq 0$; thus (a) is equivalent to the statement $\psi \neq 0$.

Applying [Spr98, Theorem 4.2.2], one knows that the restriction mapping

$$\mathrm{Der}_k(k(X), k(X)) \rightarrow \mathrm{Der}_k(f^*k(Y), k(X))$$

is dual to the mapping $\psi: \Omega_{k(Y)/k} \rightarrow \Omega_{k(X)/k}$; in particular, this restriction is 0 if and only if $\psi = 0$.

Now, it is proved for instance in [La93, VIII, Proposition 5.4] that $z \in k(X)$ is contained in $k(X)^p$ if and only if $D(z) = 0$ for each $D \in \mathrm{Der}_k(k(X), k(X))$. The assertion (a) \iff (b) follows at once. \square

If X is an affine K' -variety and $A = K'[X]$, then for $r \in \mathbb{Z}$ we write $X^{(r)}$ for the K' -variety $\mathrm{Spec}(A^{(r)})$. For an arbitrary K' -variety X , one defines the K' -variety $X^{(r)}$ by gluing together the K' -varieties $U_i^{(r)}$ from an affine open covering $\{U_i \mid 1 \leq i \leq n\}$ of X ; this construction is independent of the choice of the covering.

Let $r \geq 0$. When X is affine, the r -th Frobenius morphism $F_X^r: X \rightarrow X^{(r)}$ is defined to have comorphism $F^r: A^{(r)} \rightarrow A$. For an arbitrary K' variety X , there is a unique morphism $F_X^r: X \rightarrow X^{(r)}$ whose restriction to each affine open subset U of X is given by F_U^r .

We write $X^{(r)}/K$ for the base change of the K' -variety $X^{(r)}$ to K .

Theorem 17. *Let X and Y be geometrically irreducible K -varieties. Assume that X is defined over K' and is geometrically normal – i.e. $X_{/k}$ is normal. Suppose that $f: X \rightarrow Y$ is a K -morphism whose image contains a positive dimensional subvariety of Y . There is a unique $r \geq 0$ and a unique K -morphism $g: X^{(r)}_{/K} \rightarrow Y$ such that*

(1) $f = g \circ F_X^r$, and

(2) *there is a non-empty open subset U of $X^{(r)}$ such that $dg_x \neq 0$ for $x \in U(k)$.*

Remark 18. (a) Of course, the image of f contains a non-empty open subset U of its closure $\overline{f(X)}$ [Spr98, Theorem 1.9.5], so the dimension assumption made in the theorem is equivalent to: U has positive dimension.

(b) The theorem has been known for a long time, but it seems to be difficult to give a reference. It was used for instance by J.-P. Serre in his classification of the inseparable isogenies of height 1 of a group variety (and especially of an abelian variety), cf. Amer. J. Math. 80 (1958), pp. 715–739, Section 2.

Proof of Theorem 17. Notice that if the theorem is proved when X and Y are affine, the unicity of r and g shows that it holds as stated; we assume now that X and Y are affine. The affine variety X is defined over K' , and the domain $K'[X]$ is geometrically normal in the sense discussed previously.

Write Y' for the closure of the image of f . Then Y' is defined over K . Moreover, if $i: Y' \rightarrow Y$ denotes the inclusion, di_y is injective for all $y \in Y'(k)$; see e.g. [Spr98, Exercise 4.1.9(4)]. Since Y' is again geometrically irreducible, we may and shall replace Y by Y' ; thus we assume that f is a dominant morphism. Since the tangent maps of F_X^r are all 0, it is clear that if a suitable $r \geq 0$ exists, it is unique.

Assume that $df_x = 0$ for all smooth k -points of X ; Lemma 16 then shows that $f^*k(Y) \subset k(X)^p$. The assumption on the image of f means that the transcendence degree over K of $K(Y)$ is ≥ 1 ; since $k(X)$ is a finitely generated field extension of k , it follows that we may choose $r \geq 1$ such that $f^*k(Y) \subset k(X)^q$ for $q = p^r$ but not for $q = p^{r+1}$.

Put $q = p^r$. We now apply Lemma 14 to see that $f^*(K[Y]) \subset K[A^q]$. Lemma 13 gives then a K -algebra isomorphism $\phi: K[A^q] \rightarrow K[X^{(r)}]$ inverse to F^r , and we define $g: X^{(r)} \rightarrow Y$ to have comorphism $\phi \circ f^*$. It is clear that $f = g \circ F_X^r$ and that g is the unique morphism with this property.

The Frobenius map gives an isomorphism $F^r: k(X^{(r)}) \rightarrow k(X)^q$. If $h \in K[Y]$, and if g^*h is a p -th power in $k(X^{(r)})$ then f^*h is a q' -th power in $k(X)$, where $q' = p^{r+1}$. Since $f^*k(Y)$ is not contained in $k(X)^{q'}$, $g^*(k(Y))$ is not contained in $k(X^{(r)})^p$. It then follows from Lemma 16 that dg_x is non-0 for all x in some non-empty open subset of X , and the result is proved. \square

Remark 19. Let $X \subset \mathbb{A}^2$ denote the irreducible variety with k -points $\{(s, t) \mid s^p = t^p(t - 1)\}$, and let $Y = \mathbb{A}^1$. Consider the morphism $f: X \rightarrow Y$ given on k -points by $f(s, t) = t - 1$. Since $t - 1 = (s/t)^p$ on the open subset U of X defined by $t \neq 0$, we have $df_x = 0$ for each $x \in U(k)$. Since X is over \mathbb{F}_p in an obvious way, we identify X and $X^{(1)}$; the Frobenius map $F: X \rightarrow X$ is then just $F(s, t) = (s^p, t^p)$. There is a unique $\tilde{g}: U \rightarrow \mathbb{A}^1$ with $f|_U = \tilde{g} \circ F$; it is given on k -points by $((s, t) \mapsto s/t)$. Moreover, $d\tilde{g}_x \neq 0$ for each $x \in U(k)$. However, there is no regular function g on X such that $g|_U = \tilde{g}$; thus X is not normal, and the conclusion of the theorem does not hold for f .

Corollary 20. *Let G and H be linear algebraic K -groups. Assume that G is connected, and that G is defined over the perfect subfield K' . Let $\phi: G \rightarrow H$ be a homomorphism of K -groups such that the image of ϕ is a positive dimensional subgroup of H . There is a unique integer $r \geq 1$ and a unique homomorphism of K -groups $\psi: G^{(r)}/K \rightarrow H$ such that*

- (1) $\phi = \psi \circ F_G^r$, and
- (2) the differential $d\psi = d\psi_1$ is non-zero.

Proof. The K' -variety G is geometrically irreducible; since G/k is smooth, G is geometrically normal. Hence we may apply Theorem 17; we find a unique $r \geq 0$ and a morphism of K -varieties $\psi: G^{(r)}/K \rightarrow H/K$ such that $\psi \circ F_G^r$ coincides with the restriction of ϕ and such that $d\psi_x$ is non-zero for x in some non-empty open subset of $G^{(r)}$.

Since the Frobenius homomorphism $F_G^r: G \rightarrow G^{(r)}$ is bijective on k -points, it is clear that ψ is a homomorphism of algebraic groups. Since $d\psi_x \neq 0$ for some $x \in G^{(r)}(k)$, the map induced by ψ on left-invariant differentials in $\Omega_{G^{(r)}/K}$ is non-0; this implies that $d\psi_1 \neq 0$ and the proof is complete. \square

6. Nilpotent and unipotent elements

We return to consideration of a strongly standard reductive K -group G . Let $X \in \mathfrak{g}$ be nilpotent. A cocharacter $\Psi: G_m \rightarrow G$ is said to be associated with X if the following conditions hold:

- (A1) $X \in \mathfrak{g}(\Psi; 2)$, where for any $i \in \mathbb{Z}$ the subspace $\mathfrak{g}(i) = \mathfrak{g}(\Psi; i)$ is the i weight space of the torus $\Psi(G_m)$ under its adjoint action on \mathfrak{g} .
- (A2) There is a maximal torus $S \subset C_G(X)$ such that $\Psi(G_m) \subset (L, L)$ where $L = C_G(S)$.

With the preceding notation, X is a *distinguished* nilpotent element in the Lie algebra of the Levi subgroup L (see the discussion just before Proposition 22 for the definition).

If Ψ is associated to X , the parabolic subgroup $P = P(\Psi)$ is known variously as the canonical parabolic, the Jacobson–Morozov parabolic, or the instability parabolic (“instability flag”) associated with X . Among other things, the following result shows this parabolic subgroup to be independent of the choice of cocharacter associated to X .

Proposition/Definition 21. *Let $X \in \mathfrak{g}(K)$ be nilpotent.*

- (1) *There is a cocharacter Ψ associated with X which is defined over K .*
- (2) *If Ψ is associated to X and $P = P(\Psi)$ is the parabolic determined by Ψ , then $C_G(X) \subset P$. In particular, $\mathfrak{c}_{\mathfrak{g}}(X) \subset \text{Lie}(P)$.*
- (3) *Let U be the unipotent radical of $C = C_G^{\circ}(X)$. Then U is defined over K , and is a K -split unipotent group. If the cocharacter Ψ is associated with X , then $L = C \cap C_G(\Psi(G_m))$ is a Levi factor of C ; i.e. L is connected and reductive, and C is the semidirect product $U \cdot L$.*
- (4) *Any two cocharacters Ψ and Φ which are associated with X are conjugate by a unique element $x \in U$. If Ψ and Φ are each defined over K , then $x \in U(K)$.*
- (5) *The parabolic subgroups $P(\Psi)$ for cocharacters Ψ associated with X all coincide; the subgroup $P(X) = P(\Psi)$ is called the instability parabolic of X .*

See e.g. [Spr98, Chapter 14] for the notion of a K -split unipotent group. We will not need to explicitly refer to this notion here.

Proof. The assertion (1) in the “geometric case” (when $K = k$) is a consequence of Pommerening’s proof of the Bala–Carter theorem in good characteristic; a proof of that theorem which avoids case-checking has been given recently by Premet [Pr02] using results in geometric invariant [Ke78]. One can deduce the assertion from Premet’s work – see [Mc04, Proposition 18]. Working over the ground field K , (1) was proved in [Mc04, Theorem 26].

(2) is [Ja04, Proposition 5.9].

The first assertion of (3) is [Mc04, Theorem 28]; notice that assumption (4.1) of *loc. cit.* holds for strongly standard G , by Proposition 5. The semidirect product decomposition of C may be found in [Ja04, Proposition 5.10 and 5.11]; see also [Mc04, Corollary 29].

We now prove (4). By (3), $C = C_G^{\circ}(X)$ is the semidirect product $C = U \cdot L$ of its unipotent radical U and the Levi factor $L = C \cap C_G(\Psi(G_m))$. One knows by [Ja04, Lemma 5.3] that $\Phi = \text{Int}(g) \circ \Psi$ for an element $g \in C$. Write $g = x \cdot y$ with $x \in U$ and $y \in L$. Since y centralizes Ψ , one sees that $\Phi = \text{Int}(x) \circ \Psi$ as well.

Since $U \cap L = \{1\}$, we see that Φ and Ψ are indeed conjugate by the *unique* element $x \in U$.

Assume that Ψ and Φ are defined over K , and write $S = \Psi(G_m)$ and $S' = \Phi(G_m)$; thus $S, S' \leq C$ are tori defined over K . We have just seen that the transporter

$$N_C(S, S') = \{g \in C \mid gSg^{-1} = S'\}$$

is non-empty (it has geometric points); it follows from [Spr98, 13.3.1] that $N_C(S, S')$ is defined over K .

Choose a separable closure $K_{\text{sep}} \subset k$ of the ground field K ; [Spr98, Theorem 11.2.7] shows that $N_C(S, S')(K_{\text{sep}})$ is dense in $N_C(S, S')$; we may thus find $g \in N_C(S, S')(K_{\text{sep}})$. Since S and S' are one dimensional, and since $\text{Int}(g)$ induces an isomorphism between the respective groups of cocharacters of these tori, we must have $\text{Int}(g) \circ \Psi = \pm\Phi$. Since $g \in C$, the cocharacter $\text{Int}(g) \circ \Psi$ is associated with X ; it follows that $\text{Int}(g) \circ \Psi = \Phi$ e.g. since $X \in \mathfrak{g}(\text{Int}(g) \circ \Psi, 2)$.

Writing $g = y \cdot x$ with $x \in U$ and $y \in L$, we have $y = \lim_{t \rightarrow 0} \Psi(t)g\Psi(t^{-1})$. By Remark 8, $y \in C(K_{\text{sep}})$, so that $x = y^{-1}g \in U(K_{\text{sep}})$. Thus $x \in U(K_{\text{sep}})$ is the unique element of U for which $\text{Int}(x) \circ \Psi = \Phi$. Let $\Gamma = \text{Gal}(K_{\text{sep}}/K)$ be the Galois group. Since Ψ and Φ are Γ -stable, if $\gamma \in \Gamma$, we see that

$$\text{Int}(\gamma(x)) \circ \Psi = \Phi;$$

the unicity of x shows that $x = \gamma(x)$ and we deduce that $x \in U(K)$ as required.

To see (5), let Ψ and Φ be cocharacters associated with X . Since we have $U \leq C \leq P(\Psi)$ by (2), it follows from (4) that the parabolic subgroups $P(\Psi)$ and $P(\Phi)$ are equal. □

Recall that a nilpotent element $X \in \mathfrak{g}$ is said to be *distinguished* if the connected center of G is a maximal torus of $C_G(X)$. A parabolic subgroup $P \leq G$ is said to be distinguished if

$$\dim P/U = \dim U/(U, U) + \dim Z$$

where U is the unipotent radical of P , and Z is the center of G .

Proposition 22. *Assume that $X \in \mathfrak{g}$ is a distinguished nilpotent element. Then the instability parabolic $P = P(X)$ is a distinguished parabolic subgroup, and X lies in the dense (Richardson) orbit of P on $\text{Lie}(R_u P)$.*

Proof. [Mc04, Proposition 16]. □

Remark 23. Fixing an equivariant isomorphism $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$ defined over K , we may say that a cocharacter Ψ is associated with the unipotent element $u \in G$ if it is associated with $\Lambda(u)$. The analogous assertions of the proposition then hold for

unipotent elements of G . Note that, with this definition, the notion of cocharacter associated with a unipotent element *depends on the choice of* Λ . If Ψ is a cocharacter associated with $X = \Lambda(u)$ and if Λ' is a second Springer isomorphism, easy examples show that $\Lambda'(u)$ need not be a weight vector for Ψ . On the other hand, if Ψ' is associated with $X' = \Lambda'(u)$, then $P(\Psi) = P(\Psi')$. To see this, note that X and X' have the same centralizer. Fix a maximal torus S of this centralizer and write $L = C_G(S)$; since both Λ and Λ' restrict to isomorphisms $\mathcal{U}_L \rightarrow \mathcal{N}_L$ (see Remark 10), we may as well suppose that X and X' are distinguished. Since e.g. Λ' restricts to an isomorphism $U \rightarrow \text{Lie}(U)$ where $U = R_u(P(\Psi))$, it follows that X and X' are both Richardson elements for $P(\Psi)$. Thus Ψ and Ψ' are conjugate by an element of $P(\Psi)$ and it is then clear that $P(\Psi) = P(\Psi')$. In fact, it is even clear that Ψ and Ψ' are conjugate by an element of the unipotent radical of $P(\Psi)$; this shows that Ψ is an *optimal cocharacter* for X' (in the sense of [Ke78]) even though it need not be associated to X' .

7. The order formula and a generalization

Throughout this section, G is a strongly standard reductive k -group defined over K . Let P be a parabolic subgroup of G ; we may fix representatives $u \in U = R_u(P)$ and $X \in \text{Lie}(U)$ for the dense (Richardson) P -orbits on U and $\text{Lie}(U)$.

Recall that if the nilpotence class of U is $< p$, then $\text{Lie}(U)$ may be regarded as an algebraic K -group using the Hausdorff formula; cf. [Sci00, §5].

Proposition 24. *Assume that P is a distinguished parabolic subgroup. The following conditions are equivalent:*

- (1) u has order p ,
- (2) $X^{[p]} = 0$,
- (3) $\mathfrak{g}(\Psi; i) = 0$ for all $i \geq 2p$ and some (any) cocharacter Ψ associated to u or to X ,
- (4) the nilpotence class of U is $< p$.

Proof. The equivalence of (1) and (2) follows e.g. from [Mc03, Theorem 35]. The equivalence of (2), (3) and (4) is [Mc02, Theorem 5.4] – note that there is a misstatement (“off by 1 glitch”) concerning the nilpotence class in [Mc02] which is explained and corrected in the footnote to [Mc03, Lemma 11]. \square

Remark 25. Let X be a distinguished nilpotent element with $X^{[p]} = 0$, and let U be the unipotent radical of the instability parabolic of X . The proposition shows that the nilpotence class of $U < p$. This is not true in general for nilpotent elements which are

not distinguished. For example, let $G = \mathrm{GL}_5$, and let $X \in \mathfrak{g}$ be a nilpotent element with partition $(3, 2)$. Then X is distinguished in $\mathrm{Lie}(L)$, where L is a Levi subgroup whose derived group is $\mathrm{SL}_3 \times \mathrm{SL}_2$. If $\Psi \in X_*(L)$ is associated to X , then $P_G(\Psi)$ is a Borel subgroup of G . In particular, if $p = 3$, $X^{[p]} = 0$ but a Richardson element Y for $P_G(\Psi)$ has $Y^{[p]} \neq 0$.

Proposition 26. *Let P be a distinguished parabolic subgroup. If the equivalent conditions of Proposition 24 hold, and if P is defined over K , then:*

- (1) *there is a unique P -equivariant isomorphism of algebraic groups*

$$\varepsilon: \mathrm{Lie}(U) \rightarrow U$$

such that $d\varepsilon_0: \mathrm{Lie}(U) \rightarrow \mathrm{Lie}(U)$ is the identity.

- (2) *ε is defined over K .*

- (3) *Any homomorphism $G_a \rightarrow U$ over K has the form*

$$s \mapsto \varepsilon(sX_0) \cdot \varepsilon(s^p X_1) \cdot \varepsilon(s^{p^2} X_2) \cdots \varepsilon(s^{p^n} X_n)$$

for some elements $X_0, X_1, \dots, X_n \in \mathrm{Lie}(U)(K)$ with $[X_i, X_j] = 0$ for all $0 \leq i, j \leq n$.

Proof. Since the conditions of Proposition 24 hold, the unipotent radical $U = R_u P$ has nilpotence class $< p$. In §5 of [Sei00] – a section contributed by J.-P. Serre – one now finds the necessary results. (1) and (2) follow from Proposition 5.3 of *loc. cit.*, while (3) is Proposition 5.4 of *loc. cit.* □

Remark 27. Recall from Remark 10 that the restriction of *any* Springer isomorphism $\mathcal{N} \rightarrow \mathcal{U}$ gives a P -equivariant isomorphism $\mathrm{Lie}(U) \rightarrow U$. If $p \geq h$, there is always a Springer isomorphism whose restriction is ε . It does not seem to be clear (to the author, at least) whether a suitable analogue of this statement is true if one weakens the assumption on p .

Recall that we may regard G/k as arising by base change from a split reductive group scheme G/\mathbb{Z} over \mathbb{Z} . Write T/\mathbb{Z} for a split maximal torus of G/\mathbb{Z} .

Lemma 28. *Let $X \in \mathfrak{g}$, let L be a Levi subgroup of G with $X \in \mathrm{Lie}(L)$ distinguished, and let $\Psi \in X_*(L)$ be associated with X . We may find a number field $F \supset \mathbb{Q}$, a valuation ring $\Lambda \subset F$ whose residue field embeds in k , a standard Levi subgroup M/\mathbb{Z} of G/\mathbb{Z} , a cocharacter $\Psi' \in X_*(T/\mathbb{Z})$, and an element $Y_\Lambda \in \mathrm{Lie}(M/\Lambda)(\Psi'; 2)$ such that $(Y, M, \Psi') = g.(X, L, \Psi)$ for some $g \in G$, where $Y = Y_\Lambda \otimes 1_k$. Moreover, we may arrange that $Y_F = Y_\Lambda \otimes 1_F$ is also a Richardson element for the parabolic subgroup $P_{M/F}(\Psi') \leq M/F$.*

Proof. L is evidently conjugate to some standard Levi subgroup M , which we may regard as arising from the Levi subgroup scheme M/\mathbb{Z} . Replacing X , L , and Ψ by a G -conjugate we may thus suppose that L is standard. Replacing (X, L, Ψ) by an L -conjugate, we may then assume that X is a Richardson element for a standard distinguished parabolic of L . The remainder of the lemma is now essentially the content of [Mc02, Lemma 5.2]. \square

Proposition 29 (Spaltenstein). *Let $\Lambda \subset F$ be a valuation ring in a number field, as in the previous lemma. Let $\Psi \in X_*(T/\Lambda)$, let $X_\Lambda \in \mathfrak{g}/\Lambda(\Psi; 2)$, and assume that Ψ is associated to X_k and to X_F . Then*

$$\dim \mathfrak{c}_{\mathfrak{g}}(X_k) = \dim \mathfrak{c}_{\mathfrak{g}/F}(X_F).$$

Proof. This is essentially [Mc02, Proposition 5.2] when G is semisimple in very good characteristic. As observed in *loc. cit.*, it was proved by Spaltenstein for such G . A look at the proof of Spaltenstein in [Spa84] shows that the result remains valid for strongly standard reductive groups [the only conditions on G used in the proof in [Spa84] are: the validity of the Bala–Carter theorem and the separability of nilpotent orbits]. \square

Proposition 30. *Let $X \in \mathfrak{g}$ satisfy $X^{[p]} = 0$. If Ψ is a cocharacter associated with X and if $\mathfrak{g}(\Psi; n) \neq 0$, then $-2p + 2 \leq n \leq 2p - 2$.*

Remark 31. The analogue of the proposition for unipotent elements of order p was essentially observed by G. Seitz [Sei00] and is crucial to the proof of the existence of good A_1 -subgroups in *loc. cit.* It is proved for the classical groups in [Sei00, Proposition 4.1], and for the exceptional groups it is observed in the proof of [Sei00, Proposition 4.2] that it follows either from an explicit calculation with the associated cocharacter (“labeled diagram”) of each nilpotent orbit, or from some computer calculations of R. Lawther.

Proof of Proposition 30. It is enough to verify the proposition for a G -conjugate of Ψ and X . Lemma 28 shows that, after replacing the data X, L, Ψ by a G -conjugate, we may assume, as in that lemma, that X, L , and Ψ are “defined over Λ ” for a suitable valuation ring Λ . We write X_Λ for the element of \mathfrak{g}/Λ giving rise to $X_k = X$ by base change, and we write $X_F = X_\Lambda \otimes 1_F \in \mathfrak{g}/F$; note that Ψ is a cocharacter both of G/F and of G/k , and Ψ is associated to both X and X_F .

We now contend that if $\mathfrak{g}(\Psi; n) \neq 0$ for some $n \geq 2p - 1$, then $\text{ad}(X_k)^p \neq 0$; this implies the proposition. The proof is essentially like that of [Mc02, Theorem 5.4] except that we must also deal with the fact that the (in general, not distinguished) orbit of X may not be “even”.

Let $\mathcal{L} = \bigoplus_{i \geq -1} \mathfrak{g}/\Lambda(\Psi; i)$, and $\mathcal{L}^+ = \bigoplus_{i \geq 1} \mathfrak{g}/\Lambda(\Psi; i)$. Since we may embed X_F in an $\mathfrak{sl}_2(F)$ -triple normalized by the image of Ψ , the representation theory of

$\mathfrak{sl}_2(F)$ implies that $\mathrm{ad}(X_F): \mathcal{L}_F \rightarrow \mathcal{L}_F^+$ is surjective, where the subscript indicates “base change” – e.g. $\mathcal{L}_F = \mathcal{L} \otimes_{\Lambda} F$. In view of Proposition 29 and Proposition 21, one knows that the kernels of the maps $\mathrm{ad}(X_k): \mathcal{L}_k \rightarrow \mathcal{L}_k^+$ and $\mathrm{ad}(X_F): \mathcal{L}_F \rightarrow \mathcal{L}_F^+$ have the same dimension. We may therefore argue as in [Mc02, Proposition 5.1] and see that $\mathrm{ad}(X_k): \mathcal{L}_k \rightarrow \mathcal{L}_k^+$ is also surjective, hence that $\mathrm{ad}(X_k)^{n/2} \neq 0$ if n is even, and that $\mathrm{ad}(X_k)^{(n+1)/2} \neq 0$ if n is odd, whence our claim and the proposition. \square

8. Optimal SL_2 -homomorphisms

Throughout this section, G will denote a strongly standard reductive K -group. We first ask the reader’s patience while we fix some convenient notation for SL_2 . We choose the standard basis for \mathfrak{sl}_2 :

$$X_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \text{and} \quad Y_1 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Now put

$$x_1(t) = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad y_1(t) = \begin{pmatrix} 1 & 0 \\ t & 1 \end{pmatrix} \quad \text{for } t \in k,$$

and write $\mathcal{X} = \{x_1(t) \mid t \in k\}$ and $\mathcal{X}^- = \{y_1(t) \mid t \in k\}$. Finally, write

$$\mathcal{T} = \left\{ \begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \mid t \in k^\times \right\}$$

for the standard maximal torus of SL_2 .

We fix once and for all one of the two isomorphisms $G_m \simeq \mathcal{T}$, so that if $\phi: \mathrm{SL}_2 \rightarrow G$ is a homomorphism, it determines a cocharacter $\Psi = \phi|_{\mathcal{T}} \in X_*(G)$ by restriction to \mathcal{T} ; explicitly, Ψ is given by the rule

$$\Psi(t) = \phi \left(\begin{pmatrix} t & 0 \\ 0 & t^{-1} \end{pmatrix} \right) \quad \text{for } t \in k^\times.$$

Definition 32. The homomorphism $\phi: \mathrm{SL}_2 \rightarrow G$ is an *optimal SL_2 -homomorphism* if the cocharacter $\Psi = \phi|_{\mathcal{T}}$ is associated to the nilpotent element $X = d\phi(X_1) \in \mathfrak{g}$. Briefly, we say that ϕ is optimal for X .

We first recall that the main result of [Mc03] shows that optimal homomorphisms always exist. More precisely, let $X \in \mathfrak{g}$ with $X^{[p]} = 0$, and let Ψ be a cocharacter associated with X . If S is a maximal torus of C_Ψ , then X is distinguished in $\mathrm{Lie}(L)$ where $L = C_G(S)$. We may apply Proposition 26 to $P_L(\Psi)$; let $\varepsilon: \mathrm{Lie}(U) \rightarrow U$ be the isomorphism of that proposition, where we have written U for the unipotent radical of $P_L(\Psi)$. Now the main result of [Mc03] says the following:

Proposition 33. *There is an optimal SL_2 -homomorphism ϕ for X with the following properties:*

- (1) $\phi|_{\mathcal{F}} = \Psi$, and
- (2) $\phi(x(t)) = \varepsilon(tX)$ for each $t \in k$.

We wish to see that $\varepsilon(tX)$ is independent of the choice of the maximal torus S of C_Ψ . For this, we will use the following result due to Seitz; the result is essentially [Sei00, Proposition 4.2].

Proposition 34 (Seitz). *Let $\Lambda \subset F$ be a valuation ring in a number field whose residue field is embedded in k , let \mathcal{L} be a Λ lattice, and let $\rho_{/\Lambda}: SL_{2/\Lambda} \rightarrow GL(\mathcal{L})$ be a representation over Λ . Assume that*

- (1) *all weights of the standard maximal Λ -torus $\mathcal{T}_{/\Lambda}$ on \mathcal{L} are $\leq 2p - 2$,*
- (2) *the representation $\rho_{/k}$ of $SL_{2/k}$ is self-dual,*
- (3) *the dimension of the fixed point space of $u_F = \rho_{/F}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)$ on \mathcal{L}_F is the same as the dimension of the fixed point space of $u_k = \rho_{/k}\left(\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\right)$ on \mathcal{L}_k .*

Then the representation $(\rho_{/k}, \mathcal{L}_k)$ is a tilting module for $SL_{2/k}$.

Proof. One decomposes the $SL_{2/k}$ -module \mathcal{L}_k according to the blocks of $SL_{2/k}$. In view of the assumption on the weights of $\mathcal{T}_{/k}$ on \mathcal{L}_k , the blocks that can conceivably occur are those of the simple modules $L(d)$ with $0 \leq d < p$. The summand corresponding to the block for $d = p - 1$ is isomorphic to $L(d)^{v(d)}$ for some integer $v(d) \geq 0$. Otherwise, the summand corresponding to a block with $d < p - 1$ is isomorphic to a module of the form

$$T(c_d)^{r(d)} \oplus W(c_d)^{s(d)} \oplus (W(c_d)^\vee)^{t(d)} \oplus L(c_d)^{u(d)} \oplus L(d)^{v(d)}$$

where $c_d = 2p - 2 - d$ and where the exponents $r(d), s(d), t(d), u(d), v(d)$ are non-negative integers. [We are using Seitz's notation for $SL_{2/k}$ -representations: $W(d)$ is the Weyl module with high weight d , and $T(d)$ is the indecomposable tilting module with high weight d ; cf. [Sei00, §2].]

The assumption (2) implies that $s(d) = t(d)$ for all $0 \leq d < p - 1$. As in [Sei00, Proposition 4.2], one now expresses the dimensions of the fixed point spaces of u_k and u_F in terms of the exponents and finds that $u(d) = s(d) = t(d) = 0$ for all d . Thus \mathcal{L}_k is the direct sum of various simple tilting modules $L(d)$ for $0 \leq d < p$, and various indecomposable tilting modules $T(c_d) = T(2p - 2 - d)$ for $0 \leq d < p - 1$, so indeed \mathcal{L}_k is a tilting module. □

Proposition 35. *With notation as above, we have*

- (1) $C_G^o(X) = C_G^o(\varepsilon(X))$; in particular, $\Psi(\mathbf{G}_m)$ normalizes $C_G^o(\varepsilon(X))$.
- (2) $C_G^o(\varepsilon(X)) = C_G^o(\varepsilon(tX))$ for each $t \in k^\times$.

Proof. If X is distinguished, (1) holds since ε is $P = P(\Psi)$ equivariant, since $\varepsilon(X) \in R_u(P)$ is again a Richardson element, and since $C_G(X), C_G(\varepsilon(X)) \leq P$ by Proposition 21. [In fact, $C_G(X) = C_G(\varepsilon(X))$ always holds in this case.] It remains to prove (1) when X is no longer distinguished; we essentially follow the proof in [Sei00, Lemma 6.3].

By the unicity of ε , it is enough to prove the result with L, Ψ , and X replaced by a G -conjugate. We will regard $G = G/k$ as arising by base change from the split reductive group scheme G/\mathbb{Z} over \mathbb{Z} ; let T/\mathbb{Z} be a \mathbb{Z} -split maximal torus of G/\mathbb{Z} .

According to Lemma 28, we may find a suitable valuation ring in a number field $\Lambda \subset F$ and assume that the Levi subgroup L contains T/k and arises by base change from a standard split reductive Levi subgroup scheme $L/\mathbb{Z} \leq G/\mathbb{Z}$ containing T/\mathbb{Z} , that $\Psi \in X_*(T/\mathbb{Z})$, and that the nilpotent element $X_\Lambda \in \text{Lie}(L/\Lambda)(\Psi; 2)$ gives X on base change.

After possibly enlarging Λ and F , [Mc03, Theorem 13] gives a homomorphism

$$f: \text{SL}_{2/\Lambda} \rightarrow G/\Lambda$$

such that the restriction of f to the subgroup scheme $\begin{pmatrix} 1 & * \\ 0 & 1 \end{pmatrix}$ of $\text{SL}_{2/\Lambda}$ is given by $t \mapsto \varepsilon(tX_\Lambda)$, where $X_\Lambda \in \mathfrak{g}/\Lambda$ gives X upon extension of scalars to k (recall from [Sei00, Prop. 5.1] that ε is indeed defined over $\mathbb{Z}_{(p)}$ hence over Λ). Moreover, the restriction of f to the standard maximal torus of $\text{SL}_{2/\Lambda}$ gives the cocharacter Ψ of T/Λ .

Since G is strongly standard, its adjoint representation is self-dual. Together with Proposition 29, this shows that we may apply Proposition 34 to the representation $\text{Ad} \circ f: \text{SL}_{2/\Lambda} \rightarrow \text{GL}(\text{Lie}(G/\Lambda))$. Thus the SL_2 -representation $(\text{Ad} \circ f/k, \mathfrak{g})$ is a tilting module, and it follows from [Sei00, Lemma 2.3(d)] that

$$c_{\mathfrak{g}}(\varepsilon(tX)) = c_{\mathfrak{g}}(X)$$

for each $t \in k^\times$. The orbits of $\varepsilon(tX)$ and X are separable by Proposition 5; thus we know that $\text{Lie } C_G(\varepsilon(tX)) = \text{Lie } C_G(X)$. In particular, $C_G(X)$ and $C_G(\varepsilon(X))$ have the same dimension; assertion (1) will follow if we show that $C_G^o(X) \leq C_G^o(\varepsilon(X))$.

For any connected linear group H , we write H_t for the subgroup generated by the maximal tori in H . Applying [Spr98, 13.3.12], to the group $H = C_G^o(X)$, we find that H is generated by H_t and $C_H(S)$, where S is our fixed maximal torus of H ; i.e.

$$(*) \quad H = \langle H_t, C_H(S) \rangle.$$

Working for the moment inside the Levi subgroup $L = C_G(S)$ of G , the “distinguished” case of part (1) of the proposition means that $C_H(S) = C_L(X) = C_L(\varepsilon(X))$; in particular $C_H(S)$ centralizes $\varepsilon(X)$. So according to (*), the containment $H \leq C_G^\circ(\varepsilon(X))$, and hence (1), will follow if we just show that $\varepsilon(X)$ is centralized by each maximal torus T of $C_G(X)$. Since $\mathfrak{c}_{\mathfrak{g}}(\varepsilon(X)) = \mathfrak{c}_{\mathfrak{g}}(X) = \text{Lie } C_G(X)$, one knows that $\varepsilon(X)$ centralizes $\text{Lie}(T)$. We claim that (*) $C_G(T) = C_G(\text{Lie}(T))$; this shows that T centralizes $\varepsilon(X)$ as desired.

Write $M = C_G(T)$. Since T is a maximal torus of $C_G^\circ(X)$, it follows that T is a maximal torus of the center of M . Thus (*) is a consequence of the next lemma (Lemma 36), and (1) is proved. For (2), notice that if $s^2 = t$, we have by (1) that

$$C_G^\circ(\varepsilon(X)) = \Psi(s)C_G^\circ(\varepsilon(X))\Psi(s^{-1}) = C_G^\circ(\varepsilon(\text{Ad}(\Psi(s))X)) = C_G^\circ(\varepsilon(tX)). \quad \square$$

Lemma 36. *Let G be a strongly standard reductive group, let $T \leq G$ be a torus, and write $M = C_G(T)$. If T is a maximal torus of the center of M , then $C_G(T) = C_G(\text{Lie}(T))$.*

Proof. We essentially just reproduce the proof of [Sei00, Lemma 6.2]. Let T_0 be a maximal torus of G containing T . Denote by $R \subset X^*(T_0)$ the roots of G and by $R_L \subset R$ the roots of L . Choose a system $\alpha_1, \dots, \alpha_r \in X_*(T_0)$ of simple roots for G such that $\alpha_1, \dots, \alpha_t$ is a system of simple roots for $M = C_G(T)$ (so $t \leq r$). If we write $U_\alpha \leq G$ for the root subgroup corresponding to $\alpha \in R$, then $U_\alpha \leq L$ for $\alpha \in R_L$; moreover,

$$C_G(T) = \langle T_0; U_\alpha \mid \alpha|_T = 1 \rangle, \quad \text{and} \quad C_G(\text{Lie}(T)) = \langle T_0; U_\alpha \mid d\alpha|_{\text{Lie}(T)} = 0 \rangle.$$

We have always $C_G(T) \leq C_G(\text{Lie}(T))$. If the lemma were not true, there would be some root β of G such that $\beta|_T \neq 1$ but $d\beta|_{\text{Lie}(T)} = 0$. We may write $\beta = \alpha + \sum_{i=t+1}^r c_i \alpha_i$ with $\alpha \in R_L$. Since p is good, the c_i are integers with $0 \leq c_i < p$ [SS70, I.4.3]. Since $\beta|_T \neq 1$, it follows that c_j is non-zero in k for some $t+1 \leq j \leq r$.

Since G and M are strongly standard, [SS70, Corollary I.5.2] implies that $\mathfrak{z}(\mathfrak{g}) = \text{Lie } Z(G)$ and $\mathfrak{z}(\mathfrak{m}) = \text{Lie } Z(M)$ (where $\mathfrak{z}(\cdot)$ denotes the center of a Lie algebra, and $Z(\cdot)$ that of a group). We thus have $\dim T = \dim \mathfrak{z}(\mathfrak{g}) + (r - t)$. It follows that $\{d\alpha_{t+1}, \dots, d\alpha_r\}$ is a linearly independent subset of $\text{Lie}(T)^\vee$ (the dual space of $\text{Lie}(T)$). In particular, there is $A \in \text{Lie}(T)$ such that

$$d\alpha_i(A) = \delta_{i,j}.$$

But then $d\beta(A) = c_j \neq 0$, contradicting the choice of β . This completes the proof. \square

Remark 37. If $S, S' \leq C_\Psi$ are maximal tori, let us write U and U' for the unipotent radicals of the distinguished parabolic subgroups $P_L(\Psi) \leq L$ and $P_{L'}(\Psi) \leq L'$

where $L = C_G(S)$ and $L' = C_G(S')$. If $\varepsilon: \text{Lie}(U) \rightarrow U$ and $\varepsilon': \text{Lie}(U') \rightarrow U'$ are the isomorphisms of Proposition 26, then $\varepsilon(tX) = \varepsilon'(tX)$ for each $t \in k$. Indeed, we may choose $g \in C_\Psi^o(X)$ with $gSg^{-1} = S'$. It is then clear that $U' = gUg^{-1}$ and the uniqueness statement of Proposition 26 shows that $\varepsilon' = \text{Int}(g) \circ \varepsilon \circ \text{Ad}(g^{-1}): \text{Lie}(U') \rightarrow U'$. Let $t \in k^\times$. Proposition 35 shows that g centralizes $\varepsilon(tX)$ in addition to X . So indeed

$$\varepsilon'(tX) = \text{Int}(g) \circ \varepsilon \circ \text{Ad}(g^{-1})(tX) = \text{Int}(g) \circ \varepsilon(tX) = \varepsilon(tX)$$

as asserted.

Now let $\phi: G_a \rightarrow G$ be an injective homomorphism of algebraic groups with $X = d\phi(1)$, and assume that the cocharacter Ψ associated to X has the property that

$$\Psi(t)\phi(s)\Psi(t^{-1}) = \phi(t^2s) \quad \text{for each } t \in k^\times \text{ and } s \in k.$$

Since ϕ is injective, the cocharacter Ψ is non-trivial; this means in particular that $X \neq 0$ and so $d\phi$ is non-zero.

We remark that the homomorphism $h: G_a \rightarrow G$ given by $t \mapsto \varepsilon(tX)$ is injective. Indeed, as in the proof of Proposition 35, there is an optimal homomorphism $f: \text{SL}_2 \rightarrow G$ such that $h(s) = f(x_1(s))$ for $s \in G_a$. The group SL_2 is almost simple; its unique normal subgroup is contained in each maximal torus. In particular, $\ker h$ is trivial as asserted.

Fix now a maximal torus S of $C_G(X)$ centralized by the image of Ψ , and hence a Levi subgroup $L = C_G(S)$ such that $\Psi(G_m) \leq L$ and $X \in \text{Lie}(L)$.

Proposition 38. *With ϕ and Ψ as above, we have $\phi(t) = \varepsilon(tX)$ for each $t \in k$, where $\varepsilon: \text{Lie}(U) \rightarrow U$ is the isomorphism of Proposition 26 for the unipotent radical U of the distinguished parabolic subgroup $P_L(\Psi) \leq L$. In particular, $\phi(G_a) \leq L$.*

Proof. Notice that $\phi(s) \in C_G^o(X)$ for all $s \in G_a$. According to Proposition 35 this shows that $\phi(s) \in C_G^o(\varepsilon(tX))$ for all $t \in k^\times$, hence that

$$s \mapsto \varepsilon(-sX) \cdot \phi(s)$$

is a homomorphism $\phi_1: G_a \rightarrow G$. Moreover, $\Psi(t)\phi_1(s)\Psi(t^{-1}) = \phi_1(t^2s)$ for $t \in k^\times$ and $s \in k$, and a quick calculation shows $d\phi_1$ to be trivial.

Assume that the proposition is not true, hence that $\phi_1 \neq 1$; it has positive dimensional image and so by Corollary 20 there is a homomorphism $\phi_2: G_a \rightarrow G$ and an integer $r \geq 1$ such that $\phi_1 = \phi_2 \circ F^r$, where F denotes the Frobenius morphism for SL_2 , and such that $d\phi_2 \neq 0$. On the additive group, F is given by $s \mapsto s^p$, so we know that $\phi_1(s) = \phi_2(s^{p^r})$ for $s \in k$. [Notice we have used the fact that G_a is defined over \mathbb{F}_p , so that G_a identifies with $G_a^{(r)}$ for $r \geq 0$.]

Observe that if $\phi_1(s_0) = 1$ for some $s_0 \neq 0$, then $1 = \phi_1(s_0) = \varepsilon(-s_0X)\phi(s_0)$ so that $\varepsilon(s_0X) = \phi(s_0)$; applying $\text{Int}(\Psi(t))$ for $t \in k^\times$, we see that $\varepsilon(sX) = \phi(s)$ for all $s \in k$, so that $\phi_1 = 1$. Thus if $\phi_1 \neq 1$, then ϕ_1 is an injective map on the points of G_a . It is then clear that ϕ_2 is injective as well [since $d\phi_2$ is non-zero, ϕ_2 is even an injective homomorphism of algebraic groups].

Since $\Psi(G_m)$ normalizes the image of ϕ_2 , we have $\Psi(t)\phi_2(s)\Psi(t^{-1}) = \phi_2(t^n s)$ for some $n \in \mathbb{Z}$. Let now $t \in k^\times$ and $s \in k$. Then

$$\phi_1(t^2s) = \Psi(t)\phi_1(s)\Psi(t^{-1}) = \Psi(t)\phi_2(s^{p^r})\Psi(t^{-1}) = \phi_2(t^n s^{p^r});$$

since ϕ_1 and ϕ_2 are injective, we have $(t^2s)^{p^r} = t^n s^{p^r}$ for all $t \in k^\times$ and $s \in k$. It follows that $n = 2p^r$.

Denoting by $0 \neq Y$ an element in the image of $d\phi_2$, it is clear that $\text{Ad}(\Psi(t))Y = t^{2p^r}Y$ so that $Y \in \mathfrak{g}(\Psi; 2p^r)$. Since $r \geq 1$, since Ψ is associated with X , and since $X^{[p]} = 0$, this contradicts Proposition 30; hence $\phi_1 = 1$ and $\phi(s) = \varepsilon(sX)$ for all $s \in k$ as asserted. \square

Remark 39. Assume that $p \geq h$, where h is the Coxeter number of G . Then the nilpotence class of the unipotent radical U of a Borel subgroup B of G is $< p$. Thus there is a B -equivariant isomorphism $\varepsilon: \text{Lie}(U) \rightarrow U$ as in Proposition 26. Fix a regular nilpotent element $X \in \text{Lie}(U)$ and write $u = \varepsilon(X)$. According to Proposition 11, there is a unique Springer isomorphism $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$ with $\Lambda(u) = X$. It is then clear by the unicity of ε that $\Lambda^{-1}_{|\text{Lie}(U)} = \varepsilon$ for the unipotent radical U of any Borel subgroup of G . Since the unipotent radical V of any parabolic subgroup P of G is contained in that of some Borel subgroup, it is then clear that $\Lambda^{-1}_{|\text{Lie}(V)}$ is the isomorphism of Proposition 26 (of course, the nilpotence class of V is $< p$). This permits for these p a simple proof of Proposition 35 and hence of Proposition 38 (i.e. a proof independent of the tilting module considerations of Proposition 34)

8.1. Conjugacy of optimal SL_2 homomorphisms. The goal of this paragraph is to show that any two optimal SL_2 -homomorphisms for X are conjugate by an element of $C_G^o(X)$.

Let ϕ be an optimal SL_2 -homomorphism for $X \in \mathfrak{g}$ with cocharacter $\Psi = \phi|_{\mathcal{T}}$. Choose a maximal torus $S \leq C_\Psi$, so that X is distinguished in $\text{Lie}(L)$, where $L = C_G(S)$ is a Levi subgroup of G . If ϕ is defined over K , then the maximal torus S – and so also L – may be chosen over K .

We will write $P_L = P_L(\Psi)$ for the parabolic subgroup of L determined by the cocharacter Ψ , and U for the unipotent radical of P_L . Denote by $\varepsilon: \text{Lie}(U) \rightarrow U$ the unique P_L -equivariant isomorphism of Proposition 26.

Proposition 40. (1) *The torus S centralizes $\phi(\mathcal{X})$; in particular, $\phi(\mathcal{X}) \subset U$.*

(2) *$\phi(x_1(t)) = \varepsilon(tX)$ for each $t \in k$.*

(3) *For each $t \in k^\times$, $C_G^o(X) = C_G^o(u_t)$ where $u_t = \phi(x_1(t))$.*

Proof. We apply the result of Proposition 38; that proposition shows that $\phi(t) = \varepsilon(tX)$. (1) and (2) are then immediate, and (3) follows from Proposition 35. \square

Proposition 41. *The image of ϕ lies in the derived group of the Levi subgroup $L = C_G(S)$.*

Proof. Since SL_2 is equal to its own derived group, we only must see that the image of ϕ lies in L .

Now write

$$Y = d\phi(Y_1) \in \mathfrak{g} \quad \text{and} \quad u_t^- = \phi(y_1(t)) \in G \quad \text{for } t \in k.$$

Since SL_2 is generated by the subgroups \mathcal{X} and \mathcal{X}^- , it suffices to show that $u_t, u_t^- \in L = C_G(S)$ for all $t \in k^\times$. Fix $t \in k^\times$. It was proved in Proposition 40(1) that $u_t \in L$.

Now, there is $g \in \phi(SL_2)$ with $gu_tg^{-1} = u_t^-$ and $\text{Ad}(g)X = Y$. Together with Proposition 40, this implies that $C_G^o(u_t^-) = C_G^o(Y)$ for $t \in k^\times$. So the proof is complete once we show that $S \leq C_G(Y)$.

Since S and the image of Ψ commute, $\mathfrak{g}(\Psi; -2)$ is S -stable and is thus a direct sum of S -weight spaces

$$\mathfrak{g}(\Psi; -2) = \sum_{\gamma \in X^+(S)} \mathfrak{g}(\Psi; -2)_\gamma.$$

Hence, we may write $Y \in \mathfrak{g}(\Psi; -2)$ as a sum of S -weight vectors:

$$Y = \sum_{\gamma} Y_{\gamma} \quad \text{with } Y_{\gamma} \in \mathfrak{g}(\Psi; -2)_{\gamma}.$$

We need to show that $Y = Y_0$, or equivalently that $Y_{\gamma} = 0$ for $\gamma \neq 0$.

As Ψ is associated to X , it follows from Proposition 21 that $\mathfrak{c}_{\mathfrak{g}}(X) \subseteq \sum_{i \geq 0} \mathfrak{g}(\Psi; i)$. Since S centralizes X , it follows that $\text{ad}(X): \mathfrak{g}(\Psi; 2) \rightarrow \mathfrak{g}(\Psi; 0)$ is an injective map of S -representations. Writing $H = d\Psi(1) \in \mathfrak{g}$, we have $\text{ad}(X)Y = [X, Y] = H \in \mathfrak{g}(\Psi; 0)_0$. Since $\text{ad}(X)Y_{\gamma} \in \mathfrak{g}(\Psi; 0)_{\gamma}$, the injectivity of $\text{ad}(X)$ implies that $Y_{\gamma} = 0$ unless $\gamma = 0$, as desired. Thus $Y = Y_0$ and the proof is complete. \square

Proposition 42. *Let $X \in \mathfrak{g}$ satisfy $X^{[p]} = 0$. If ϕ_1 and ϕ_2 are optimal SL_2 -homomorphisms for X and if $\phi_1|_{\mathcal{T}} = \phi_2|_{\mathcal{T}}$, then $\phi_1 = \phi_2$.*

Proof. Combined with Proposition 41, the hypotheses yield a maximal torus $S \leq C_G(X)$ such that the image of ϕ_i lies in $L = C_G(S)$ for $i = 1, 2$. Thus we may replace G by the strongly standard reductive group L and so suppose that X is distinguished.

Proposition 40 shows that $\phi_1(x_1(t)) = \varepsilon(tX) = \phi_2(x_1(t))$ for all $t \in k$. It follows that ϕ_1 and ϕ_2 coincide on the Borel subgroup $B = \mathcal{T}\mathcal{X}$ of SL_2 . Using this, we argue

that ϕ_1 and ϕ_2 coincide on all of SL_2 . Indeed, consider the morphism of varieties $SL_2 \rightarrow G$ given by

$$g \mapsto \phi_1(g)\phi_2(g^{-1}).$$

Since the ϕ_i are homomorphisms, this morphism factors through the flag variety $SL_2/B = \mathbb{P}^1$ (the projective line); since \mathbb{P}^1 is an irreducible complete variety, and since G is affine, this morphism must be constant. The proof is complete. \square

Corollary 43. *If ϕ is an optimal homomorphism, let as usual $X = d\phi(X_1)$ and $\Psi = \phi|_{\mathcal{G}}$. Then the centralizer of $\phi(SL_2)$ is $C_\Psi = C_G(X) \cap C_G(\Psi(\mathbf{G}_m))$.*

Proof. This is just a restatement of the previous proposition. \square

Theorem 44. *Suppose that G is strongly standard, and that $X \in \mathfrak{g}$ satisfies $X^{[p]} = 0$. Then any two optimal SL_2 -homomorphisms for X are conjugate by a unique element of the unipotent radical of $C_G^\circ(X)$.*

Proof. Let ϕ_1, ϕ_2 be optimal SL_2 -homomorphisms for X , and write $\Psi_i = \phi_i|_{\mathcal{G}}$ for the corresponding cocharacters. According to Proposition 21, the cocharacters Ψ_1 and Ψ_2 associated with X are conjugate by a unique element of the unipotent radical U of $C_G^\circ(X)$. Replacing ϕ_2 by a U -conjugate, we may thus suppose that $\Psi_1 = \Psi_2$. It then follows from Proposition 42 that $\phi_1 = \phi_2$. \square

8.2. Uniqueness of a principal homomorphism. Suppose that $X \in \mathfrak{g}$ is a distinguished nilpotent element. Then any cocharacter $\Psi \in X_*(G)$ with $X \in \mathfrak{g}(\Psi; 2)$ is associated to X . In particular, if $\phi: SL_2 \rightarrow G$ is any homomorphism with $d\phi(X_1) = X$, then $\Psi = \phi|_{\mathcal{G}}$ is a cocharacter associated with X ; thus ϕ is optimal.

An application of Theorem 44 now gives:

Proposition 45. *If $\phi_1, \phi_2: SL_2 \rightarrow G$ are homomorphisms such that $d\phi_1(X_1) = d\phi_2(X_1) = X$ is a distinguished nilpotent element, then ϕ_1 and ϕ_2 are conjugate by an element of $C_G^\circ(X)$.*

A principal homomorphism $\phi: SL_2 \rightarrow G$ is one for which $d\phi(X_1)$ is a regular nilpotent element. Since a regular nilpotent element is distinguished, we have:

Proposition 46. *A principal homomorphism is optimal. Any two principal homomorphisms are conjugate in G .*

8.3. Optimal homomorphisms over ground fields. Recall that K is an arbitrary ground field. The following theorem gives both an existence result and a conjugacy result for optimal homomorphisms over the ground field K . If $X \in \mathfrak{g}(K)$, write

$C = C_G^o(X)$ for its connected centralizer; recall by Proposition 21 that the unipotent radical of C is defined over K .

Theorem 47. *Let G be a strongly standard reductive K -group, and let $X \in \mathfrak{g}(K)$ satisfy $X^{[p]} = 0$.*

- (1) *There is an optimal SL_2 -homomorphism ϕ for X which is defined over K .*
- (2) *Let U be the unipotent radical of $C = C_G^o(X)$. Any two optimal SL_2 -homomorphisms for X defined over K are conjugate by a unique element of $U(K)$.*

Proof. To prove (1), we need first to quote a more precise form of Proposition 33. The proof of that Proposition given in [Mc03] shows that there is a nilpotent element X'' in the orbit of X which is rational over the separable closure K_{sep} of K in k and an optimal SL_2 -homomorphism ϕ'' for X'' defined over K_{sep} . Since the orbit of X is separable, one can mimic the proof of [Spr98, 12.1.4] to see that X and X'' are conjugate by an element rational over K_{sep} . Indeed, let \mathcal{O} be the orbit of X and let $\mu: G \rightarrow \mathcal{O}$ be the orbit map $\mu(g) = \mathrm{Ad}(g)X$. The separability of the orbit \mathcal{O} means that $d\mu_1: T_1(G) \rightarrow T_X(\mathcal{O})$ is surjective, and it follows for each $g \in G$ that $d\mu_g: T_g(G) \rightarrow T_{\mathrm{Ad}(g)X}(\mathcal{O})$ is surjective. It follows from [Spr98, 11.2.14] that the fiber $\mu^{-1}(X'')$ is defined over K_{sep} , so that by [Spr98, 11.2.7] there is a K_{sep} -rational point g in this fiber. It follows that $\phi' = \mathrm{Int}(g) \circ \phi''$ is an optimal SL_2 -homomorphism for X which is defined over K_{sep} .

According to Proposition 21, we can find a cocharacter Ψ associated with X which is defined over K . Writing $C = C_G^o(X)$, that same proposition shows that the cocharacters Ψ and $\Psi' = \phi'_{|\mathcal{T}}$ are conjugate by an element $h \in C(K_{\mathrm{sep}})$ [in fact, h can be chosen to be a K_{sep} -rational element of the unipotent radical of C].

It now follows that $\phi = \mathrm{Int}(h^{-1}) \circ \phi'$ is an optimal SL_2 -homomorphism for X which is defined over K_{sep} . We argue that ϕ is actually defined over K . Let $\gamma \in \mathrm{Gal}(K_{\mathrm{sep}}, K)$. Then $\phi_\gamma = \gamma \circ \phi \circ \gamma^{-1}: \mathrm{SL}_2 \rightarrow G$ is another optimal SL_2 -homomorphism for X ; since $\Psi = \phi_{|\mathcal{T}}$ is defined over K , $\phi_{|\mathcal{T}} = \phi_{\gamma|\mathcal{T}}$. Thus Proposition 42 shows that $\phi = \phi_\gamma$. Since ϕ is defined over K_{sep} , Galois descent (e.g. [Spr98, Cor. 11.2.9]) shows that ϕ is defined over K .

We now give the proof of (2), which is the same as the proof of Theorem 44. If ϕ and ψ are optimal SL_2 -homomorphisms for X , each defined over K , then by Proposition 21, the K -cocharacters $\Phi = \phi_{|\mathcal{T}}$ and $\Psi = \psi_{|\mathcal{T}}$ associated with X are conjugate by a unique element of $U(K)$. Thus we may replace ψ by a $U(K)$ -conjugate and suppose that $\phi_{|\mathcal{T}} = \psi_{|\mathcal{T}}$. Proposition 42 then shows that $\phi = \psi$ and the proof is complete. \square

Remark 48. In the case of a *finite* ground field K , Seitz [Sei00, Proposition 9.1] obtained existence and conjugacy over K for *good* A_1 subgroups (see §8.5 below for their definition).

8.4. Complete reducibility of optimal homomorphisms. Let G be any reductive group. Generalizing the notion of a completely reducible representation of a group, J.-P. Serre has introduced the following definition. A subgroup $H \leq G$ is said to be G -completely reducible (for short: G -cr) if for every parabolic subgroup P of G containing H there is a Levi subgroup of P which also contains H . See [Ser04] for more on this notion.

We are going to prove that the image of an optimal homomorphism is G -cr. We establish some technical lemmas needed in the proof. First, we show that a suitable generalization of Proposition 35 is valid.

Lemma 49. *Let $\Psi \in X_*(G)$ and suppose that $P = P(\Psi)$ is a distinguished parabolic subgroup with unipotent radical $U = R_u P$. Suppose that the nilpotence class of U is $< p$, and let*

$$\varepsilon: \text{Lie}(U) \rightarrow U$$

be the isomorphism of Proposition 26. If $X_0 \in \mathfrak{g}(\Psi; n)$ for some $n \geq 1$, then $X_0 \in \text{Lie}(U)$ and $C_G^o(X_0) = C_G^o(\varepsilon(X_0))$.

Proof. Let $N(X_0) = \{g \in G \mid \text{Ad}(g)X_0 \in kX_0\} \leq G$. By assumption, the torus $\Psi(G_m)$ is contained in $N(X_0)$; in particular, this torus normalizes $C_G(X_0)$. We may choose a maximal torus S of $C_G(X_0)$ centralized by $\Psi(G_m)$; thus $S' = S \cdot \Psi(G_m)$ is a maximal torus of $N(X_0)$. According to [Mc04, Lemma 25], there is a cocharacter $\Lambda \in X_*(S')$ which is associated to X_0 . Let T be a maximal torus of G containing S' ; thus T lies in the centralizer of $\Lambda(G_m)$, of S , and of $\Psi(G_m)$.

Since a Richardson orbit representative X for the dense P -orbit on U satisfies $X^{[p]} = 0$, we have also $X_0^{[p]} = 0$. Now consider the Levi subgroup $L = C_G(S)$; the nilpotent element X_0 is distinguished in $\text{Lie}(L)$. Let $Q = P_L(\Lambda)$, and let $V = R_u Q$ be the unipotent radical of Q . Proposition 26 gives a unique isomorphism

$$\varepsilon': \text{Lie}(V) \rightarrow V,$$

and we know from Proposition 35 that $C_G^o(X_0) = C_G^o(\varepsilon'(X_0))$. Thus our lemma will follow if we show that $\varepsilon(X_0) = \varepsilon'(X_0)$.

Notice that T is contained in the Levi factors $Z_G(\Psi)$ of P and $Z_L(\Lambda)$ of Q , so that T normalizes the connected unipotent subgroup $W = (U \cap V)^\circ$ of G . Since the nilpotence class of W is $< p$, [Sei00, Proposition 5.2] gives a unique isomorphism of algebraic groups

$$\varepsilon'': \text{Lie}(W) \rightarrow W$$

whose tangent map is the identity and which is compatible with the action of the connected solvable group $T \cdot W$ by conjugation. On the other hand, the tangent maps of the restrictions $\varepsilon|_{\text{Lie}(W)}$ and $\varepsilon'|_{\text{Lie}(W)}$ are the identity, and these maps are compatible

with the action of $T \cdot W$; we thus have

$$\varepsilon|_{\text{Lie}(W)} = \varepsilon'' = \varepsilon'_{|\text{Lie}(W)}.$$

This implies that $\varepsilon(X_0) = \varepsilon'(X_0)$ as desired, and the proof is complete. \square

We now show that a suitable deformation of an optimal homomorphism remains optimal.

Lemma 50. *Let $\phi: \text{SL}_2 \rightarrow G$ be an optimal SL_2 -homomorphism, and suppose that ϕ takes its values in the parabolic subgroup P .*

- (1) *There is a cocharacter $\gamma \in X_*(P)$ such that $\gamma(G_m)$ centralizes $\phi(\mathcal{T})$ and such that $P = P(\gamma)$.*
- (2) *Denoting by $L = Z(\gamma)$ the Levi factor of P determined by γ , write $\widehat{\phi}: \text{SL}_2 \rightarrow L$ for the homomorphism*

$$x \mapsto \lim_{t \rightarrow 0} \gamma(t)\phi(x)\gamma(t^{-1})$$

of Lemma 7. Then $\widehat{\phi}$ is an optimal SL_2 -homomorphism as well.

Proof. Since $\phi(\mathcal{T})$ lies in some maximal torus of P , (1) follows from Lemma 6.

Let us prove (2). Let $X = d\phi(X_1)$ as usual, and write Ψ for the cocharacter $\phi|_{\mathcal{T}}$; it is associated with X . Denoting by C_Ψ the corresponding Levi factor of the centralizer of X , we may choose a maximal torus $S \leq C_\Psi$ and Proposition 41 implies that ϕ takes its values in the Levi subgroup $C_G(S)$. We may evidently replace G by L and so assume that X is distinguished.

Now let $X = X_0 + X'$, $Y = Y_0 + Y'$ with $X_0, Y_0 \in \text{Lie}(L) = \mathfrak{g}(\gamma; 0)$ and with $X', Y' \in \text{Lie}(R_u P)$. Lemma 7 shows that $d\widehat{\phi}(X_1) = X_0$ and $d\widehat{\phi}(Y_1) = Y_0$.

To show that $\widehat{\phi}$ is optimal for X_0 , it is enough to show that $\widehat{\phi}$ takes values in some Levi subgroup M of L such that $X_0 \in \text{Lie}(M)$ is distinguished. Indeed, since SL_2 is its own derived group, this will imply that $\Psi = \phi|_{\mathcal{T}}$ takes its values in (M, M) , so that Ψ is indeed associated with X_0 .

Note that the torus $\Psi(G_m)$ normalizes $C_L(X_0)$. Since $\Psi(G_m)$ lies in a maximal torus of the semidirect product of $C_L(X_0)$ and $\Psi(G_m)$, it is clear that there is a maximal torus S of $C_L(X_0)$ centralized by $\Psi(G_m)$. Taking $M = C_L(S)$, we claim that ϕ takes its values in M .

Notice that

$$\widehat{\phi}(x_1(t)) = \lim_{s \rightarrow 0} \gamma(s)\varepsilon(tX)\gamma(s^{-1}) = \lim_{s \rightarrow 0} \varepsilon(t(X_0 + \text{Ad}(\gamma(s))X')) = \varepsilon(tX_0)$$

for each $t \in k$. Similarly, $\widehat{\phi}(y_1(t)) = \varepsilon(tY_0)$ for each $t \in k$.

Since S is contained in the centralizer of X , it is contained in the instability parabolic P_X for X by Proposition 21. Thus ε is S -equivariant. Since SL_2 is generated by \mathcal{X} and \mathcal{X}^- , this equivariance shows that we are done if S centralizes both X_0 and Y_0 – of course, S centralizes X_0 by assumption.

Write $H = d\Psi(1)$; since Ψ and γ commute, $\widehat{\phi}|_{\mathcal{T}} = \Psi$. Now, $\text{ad}(X_0)Y_0 = [X_0, Y_0] = H$. As in the proof of Proposition 41, we write $Y_0 = \sum_{\lambda \in X^*(S)} Y_{0,\lambda}$ as a sum of weight vectors for the torus S . Since $\Psi(G_m)$ commutes with S , H is centralized by S , and so we have $[X_0, Y_{0,\lambda}] = 0$ when $\lambda \neq 0$; we want to conclude that $Y_{0,\lambda} = 0$. We do not know that Ψ is associated with X_0 , so we can not simply invoke Proposition 21. However, since $Y_{0,\lambda} \in \mathfrak{g}(\Psi; -2)$, the general theory of SL_2 -representations shows: if $Y_{0,\lambda} \neq 0$, then $\widehat{\rho}(x_1(t)) = \varepsilon(tX_0)$ acts non-trivially on $Y_{0,\lambda}$ for some $t \in k^\times$. On the other hand, according to Lemma 49 we have $C_L^o(X_0) = C_L^o(\varepsilon(tX_0))$, so that $Y_{0,\lambda} \in \mathfrak{c}_{\text{Lie}(L)}(X_0) = \mathfrak{c}_{\text{Lie}(L)}(\varepsilon(tX_0))$. Thus indeed $Y_{0,\lambda} = 0$ for each non-zero λ , as required. Thus $Y_0 = Y_{0,0}$ so that S centralizes Y_0 ; the proof is now complete. \square

Lemma 51. *Let $X \in \mathfrak{g}$ be any nilpotent element, let $\psi \in X_*(G)$ a cocharacter associated with X , and let $L = C_G(\psi(G_m))$ be the Levi factor in the instability parabolic determined by ψ .*

- (1) *The L orbit $\mathcal{V} = \text{Ad}(L)X$ is a Zariski open subset of $\mathfrak{g}(\psi; 2)$.*
- (2) *Let $Y \in \mathfrak{g}$ be nilpotent. Then ψ is a cocharacter associated with Y if and only if $Y \in \mathcal{V}$.*

Proof. To prove (1), note that the orbit map

$$y \mapsto \text{Ad}(y)X : L \rightarrow \mathfrak{g}(\psi; 2)$$

has differential $\text{ad}(X) : \text{Lie}(L) = \mathfrak{g}(\psi; 0) \rightarrow \mathfrak{g}(\psi; 2)$; if we know that the differential is surjective, then the orbit map is dominant and separable and (1) follows. To see the surjectivity, we argue as follows. Recall from Proposition 21 that $\mathfrak{c}_{\mathfrak{g}}(X)$ is contained in $\sum_{i \geq 0} \mathfrak{g}(\psi; i)$; in particular, $\mathfrak{g}(\psi; -2) \cap \mathfrak{c}_{\mathfrak{g}}(X) = 0$. According to [Ja04, Lemma 5.7] this last observation implies (in fact: is equivalent to) the statement $[\mathfrak{g}(\psi; 0), X] = \mathfrak{g}(\psi; 2)$; this proves the required surjectivity (note that [Ja04, 5.7] is applicable since the Lie algebra of a strongly standard reductive group has on it a nondegenerate, invariant, symmetric, bilinear form – cf. Proposition 2).

For (2) note first that ψ is evidently associated to any $Y \in \mathcal{V}$. Conversely, if ψ is associated to Y , then $Y \in \mathfrak{g}(\psi; 2)$, and (1) shows that $\text{Ad}(L)Y$ is also open and dense in $\mathfrak{g}(\psi; 2)$. Thus $\text{Ad}(L)X \cap \text{Ad}(L)Y \neq \emptyset$, so that $Y \in \text{Ad}(L)X = \mathcal{V}$. \square

Theorem 52. *Let G be strongly standard, and let $\phi : SL_2 \rightarrow G$ be an optimal SL_2 homomorphism. Then the image of ϕ is G -cr.*

Proof. Let $X = d\phi(X_1)$ as usual, and write Ψ for the cocharacter $\phi|_{\mathcal{G}}$; it is associated with X . Denoting by C_Ψ the corresponding Levi factor of the centralizer of X , we may choose a maximal torus $S \leq C_\Psi$ and Proposition 41 implies that ϕ takes its values in the Levi subgroup $L = C_G(S)$. Applying [Ser04, Proposition 3.2], one knows that $\phi(\mathrm{SL}_2)$ is G -cr if and only if it is L -cr. We replace G by L , and thus suppose that X is *distinguished*.

Let P be a parabolic subgroup of G and suppose that the image of ϕ lies in P . We claim that since X is distinguished, we must have $P = G$; this will prove the theorem.

To prove our claim, first notice that by Lemma 50(1) we may choose $\gamma \in X_*(P)$ with $P = P(\gamma)$ and such that $\gamma(\mathbf{G}_m)$ commutes with $\Psi(\mathbf{G}_m)$.

Let us write $X = \sum_{i \geq 0} X_i$ with $X_i \in \mathfrak{g}(\gamma; i)$. Consider the homomorphism $\widehat{\phi}: \mathrm{SL}_2 \rightarrow Z(\gamma)$ constructed in Lemma 50; according to (2) of that lemma, $\widehat{\phi}$ is optimal for X_0 , so that the cocharacter Ψ is associated to X_0 as well as to X .

We now claim that X and X_0 are conjugate. This will show that X_0 is distinguished in G , hence that $G = Z(\gamma)$ so that also $P = G$ as desired. Let $L = C_G(\Psi(\mathbf{G}_m))$. Then Lemma 51 implies that X_0 is contained in the orbit $\mathcal{V} = \mathrm{Ad}(L)X \subset \mathfrak{g}(\Psi; 2)$, proving our claim. \square

8.5. Comparison with good homomorphisms. According to Seitz [Sei00], an SL_2 homomorphism $\phi: \mathrm{SL}_2 \rightarrow G$ is called *good* (or *restricted*) provided that the weights of a maximal torus of SL_2 on $\mathrm{Lie}(G)$ are all $\leq 2p - 2$.

Proposition 53. *Let $\phi: \mathrm{SL}_2 \rightarrow G$ be a homomorphism, where G is a strongly standard reductive group. Then ϕ is good if and only if it is optimal for $X = d\phi(X_1)$. In particular, all good SL_2 -homomorphisms whose image contains the unipotent element v are conjugate by $C_G^o(v)$.*

Proof. That an optimal homomorphism is good follows from Proposition 30. Choose a Springer isomorphism $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$. If u is a unipotent element of order p , choose a Levi subgroup L in which u is distinguished; this just means that $X = \Lambda(u) \in \mathfrak{g}$ is distinguished. It follows from Proposition 24 that $X^{[p]} = 0$. Choose an optimal homomorphism ϕ' for X ; we know that ϕ' takes values in L (Proposition 41), and if $v = \phi'(x(1))$, it is clear from Proposition 40 that v and u are Richardson elements in the same parabolic subgroup of L ; thus v and u are conjugate. This proves that u is in the image of some optimal homomorphism ϕ .

To prove that good homomorphisms are optimal, we use a result of Seitz. Since ϕ is optimal, we just observed that it is good, and Seitz proved [Sei00, Theorem 1.1] that any good homomorphism with u in its image is conjugate by $C_G(u)$ to ϕ . Thus, any good homomorphism is indeed optimal. \square

9. Rational elements of a nilpotent orbit defined over a ground field

In this section, we extend a result first obtained by R. Kottwitz [Ko82] in the case where K has characteristic 0. We give here a proof which is also valid in positive characteristic (under some assumptions on G). For the most part, we follow the original argument of Kottwitz.

Theorem 54. *Let K be any field, and let G be a strongly standard connected reductive K -group which is K -quasisplit. If the nilpotent orbit $\mathcal{O} \subset \mathcal{N}$ is defined over K , then \mathcal{O} has a K -rational point.*

Proof. If K is a finite field, the theorem is a consequence of the Lang–Steinberg theorem; cf. [St68, §10] and [St65]. Suppose now K to be infinite.

We fix a Borel subgroup B of G which is defined over K , and a maximal torus $T \subset B$ which is also over K . The roots of G in $X^*(T)$ which appear in the Lie algebra of the unipotent radical of B are declared positive, and we will write $\bar{C} \subset X_*(T)$ for the positive Weyl chamber determined by B :

$$\bar{C} = \{\mu \mid \langle \alpha, \mu \rangle \geq 0 \text{ for all positive roots } \alpha \text{ of } G \text{ in } X^*(T)\}.$$

If $W = N_G(T)/T$ denotes the Weyl group of T , then each $\mu \in X_*(T)$ is W -conjugate to a unique point in \bar{C} . We also write $\Gamma = \text{Gal}(K_{\text{sep}}/K)$ for the absolute Galois group of the field K .

The K -variety \mathcal{O} has a point X' rational over the separable closure K_{sep} of K in k (e.g. by [Spr98, 11.2.7]). According to Proposition 21, there is a cocharacter Ψ' associated with X' and defined over K_{sep} . Let T' be a maximal torus of G defined over K_{sep} which contains the image of Ψ' .

For $\gamma \in \Gamma$, the cocharacter Ψ'^{γ} is associated with the nilpotent X'^{γ} . Since \mathcal{O} is defined over K , X'^{γ} and X' are conjugate. Hence Ψ' and Ψ'^{γ} are conjugate by another application of Proposition 21.

According to [Spr98, Proposition 13.3.1 and 11.2.7] we may find $g \in G(K_{\text{sep}})$ such that $gT'g^{-1} = T$; the same reference shows that any element w of the Weyl group of T may be represented by an element $\dot{w} \in N_G(T)$ rational over K_{sep} . We have that $\Psi = \text{Int}(g) \circ \Psi' \in X_*(T)$ is defined over K_{sep} . Replacing Ψ by $\text{Int}(\dot{w}) \circ \Psi$ for a suitable w in the Weyl group of T , we may suppose that $\Psi \in \bar{C} \subset X_*(T)$ and is defined over K_{sep} . Of course, Ψ is associated with the nilpotent element $X = \text{Ad}(\dot{w}g)X'$.

Since B and T are Γ -stable, γ permutes the positive roots in $X^*(T)$. Thus, γ leaves \bar{C} invariant; in particular, $\Psi^{\gamma} \in \bar{C}$. We know Ψ and Ψ^{γ} to be conjugate in G . Since T is a maximal torus of the centralizer of both $\Psi(G_m)$ and of $\Psi^{\gamma}(G_m)$, we may suppose that $\Psi^{\gamma} = \text{Int}(\dot{w})\Psi$ for some w in the Weyl group of T . But \bar{C} is a

fundamental domain for the W -action on $X_*(T)$, so we see that $\Psi = \Psi^\nu$. Since Ψ is defined over K_{sep} and is Γ -stable, Ψ is defined over K [Spr98, 11.2.9].

This shows in particular that the subspace $\mathfrak{g}(\Psi; 2)$ is defined over K . According to Lemma 51, there is a Zariski open subset of $\mathfrak{g}(\Psi; 2)$ consisting of elements in \mathcal{O} . Since K is infinite, the K -rational points of $\mathfrak{g}(\Psi; 2)$ are Zariski dense in $\mathfrak{g}(\Psi; 2)$. Hence there is a K -rational point in \mathcal{O} and the proof is complete. \square

Corollary 55. *Let G be a strongly standard reductive K -group which is K -quasisplit. There is a regular nilpotent element $X \in \mathfrak{g}(K)$. In particular, there is an optimal homomorphism $\phi: \text{SL}_2 \rightarrow G$ defined over K with $d\phi(X_1) = X$.*

Proof. Since G is split over a separable closure K_{sep} of K , there is a K_{sep} rational regular nilpotent element. Thus the regular nilpotent orbit is defined over K_{sep} . Since this orbit is clearly stable under $\text{Gal}(K_{\text{sep}}/K)$, it is defined over K . So the theorem shows that there is a K -rational regular nilpotent element X . The final assertion follows from Theorem 47. \square

Remark 56. With G as in the theorem, there is a Springer isomorphism $\Lambda: \mathcal{U} \rightarrow \mathcal{N}$ defined over K . Thus a unipotent conjugacy class defined over K has a K -rational point.

10. Appendix: Springer isomorphisms (Jean-Pierre Serre, June 1999)

Let G be a simple algebraic group in char. p , which I assume to be “good” for G . I also assume the ground field k to be algebraically closed. Call G^u the variety of unipotent elements of G and \mathfrak{g}^n the subvariety of $\mathfrak{g} = \text{Lie}(G)$ made up of the nilpotent elements.

Springer has shown that there exist algebraic morphisms

$$f: G^u \rightarrow \mathfrak{g}^n$$

with the following properties:

- a) f is compatible with the action of G by conjugation on both sides.
- b) f is bijective.

In fact, it was later shown that these properties imply (at least when p is “very good”, which is always the case if G is not of type A):

- b') f is an isomorphism of algebraic varieties.

Despite the fact that there are *many such f 's* (they make up an algebraic variety of dimension ℓ , where ℓ is the rank of G), one often finds in the literature the expression

“the Springer isomorphism” used – and abused –, especially to conclude that the G -classes of unipotent elements of G and nilpotent elements of \mathfrak{g} are in a natural correspondence, namely “the” Springer correspondence.

It might be good for the reader to consider the case of $G = SL_n$ (or rather PGL_n , if one wants an adjoint group). In that case a Springer isomorphism is of the form

$$1 + e \mapsto a_1 e + \dots + a_{n-1} e^{n-1},$$

where $e^n = 0$ (so that $u = 1 + e$ is unipotent), and the a_i are elements of k with $a_1 \neq 0$. Every such family $\vec{a} = (a_1, \dots, a_{n-1})$ defines a unique Springer isomorphism $f_{\vec{a}}$, and one gets in this way every Springer isomorphism, once and only once. This example also shows that the Springer isomorphisms can be quite different: e.g., for some one may have $f(u^m) = m \cdot f(u)$ for all u and all $m \in \mathbb{Z}$ (such an f exists if and only if $p \geq n$), and for some one does not even have $f(u^{-1}) = -f(u)$!

In what follows, I want to repair this unfortunate mix-up by showing that all the different Springer isomorphisms give *the same* bijection between the G -classes of G^u and the G -classes of \mathfrak{g}^n , so that one can indeed speak (in that case) of *the* Springer bijection.

I have to recall first how the Springer isomorphisms are defined. Call G^{ur} the set of regular unipotent elements of G ; it is an open dense set in G^u ; same definition for \mathfrak{g}^n in $\mathfrak{g} = \text{Lie}(G)$. Choose an element u in G^{ur} and let $C(u)$ be its centralizer. It is known that $C(u)$ is smooth, connected, unipotent, commutative, of dimension ℓ ($= \text{rank } G$). Let $\mathfrak{c}(u) = \text{Lie } C(u)$ be its Lie algebra. Choose an element X of $\mathfrak{c}(u)$ which is regular. Then its centralizer is $C(u)$, and the Springer construction shows that there is a *unique* Springer isomorphism $f = f_{u,X}$ which has the property that $f(u) = X$. Let us fix X ; then it is clear that every Springer isomorphism is equal to $f_{v,X}$ for some $v \in C(u)^r$, where $C(u)^r = C(u) \cap G^{ur}$; moreover, v is uniquely defined by f . Hence we have a *one-to-one parametrization of the Springer isomorphisms by the elements v of $C(u)^r$* .

The next step consists in showing that this parametrization is “algebraic”. The precise meaning of this is the following:

Proposition. *There exists an algebraic morphism $F : C(u)^r \times G^u \rightarrow \mathfrak{g}^n$ such that $F(v, z) = f_{v,X}(z)$ for every $v \in C(u)^r$ and $z \in G^u$.*

Proof. Call N_u the normalizer of $C(u)$ in G . Since all regular unipotents are conjugate, N_u acts transitively on $C(u)^r$, so that one can identify the algebraic variety $C(u)^r$ with the coset space $N_u/C(u)$. Similarly, one may identify G^{ur} with $G/C(u)$. Let us now define an algebraic map

$$F' : N_u \times G \rightarrow \mathfrak{g}^n$$

by the formula

$$F'(n, z) = \text{Ad}(zn^{-1}).X$$

(i.e. the image of $X \in \mathfrak{g}$ by the inner automorphism defined by zn^{-1}). It is clear that $F'(n, z)$ depends on n only mod. $C(u)$, and that it depends on z also mod $C(u)$. Hence F' factors out and gives a map of $N_u/C(u) \times G/C(u)$ into \mathfrak{g}^n . If we identify $N_u/C(u)$ with $C(u)^r$ and $G/C(u)$ with G^{ur} , we thus get a map

$$F_0: C(u)^r \times G^{\text{ur}} \rightarrow \mathfrak{g}^n.$$

It is well-known that G^u is a normal variety and that $G^u - G^{\text{ur}}$ has codimension > 1 in G^u . Hence the same is true for $C(u)^r \times G^{\text{ur}}$ in $C(u)^r \times G^u$. Since \mathfrak{g}^n is an affine variety, the map F_0 extends uniquely to an algebraic map $F: C(u)^r \times G^u \rightarrow \mathfrak{g}^n$. One checks immediately that for every fixed $v \in C(u)^r$, the map $z \mapsto F(v, z)$ has the following properties: a) it commutes with the action of G ; b) it maps v to X . (Property a) is checked on G^{ur} first; by continuity, it is valid everywhere.) This shows that F is the map we wanted. \square

Corollary. *The bijection*

$$G\text{-classes of } G^u \rightarrow G\text{-classes of } \mathfrak{g}^n$$

given by a Springer isomorphism f is independent of the choice of f .

This is easy. One uses the following elementary lemma:

Lemma. *Let Y, Z be two G -spaces. Assume G has finitely many orbits in each. Let T be a connected space, and $F: T \times Y \rightarrow Z$ a morphism such that, for every $t \in T$, the map $y \mapsto F(t, y)$ is a G -isomorphism of Y on Z .*

Then, for every $y \in Y$, the points $F(t, y)$, $t \in T$, belong to the same G -orbit.

Proof by induction on $\dim Y = \dim Z$. The statement is clear in dimension zero, because of the connexity of T . If $\dim Y > 0$, there are finitely many open orbits in Y (resp. Z); call Y_0 and Z_0 their union. It is clear that, for every t , the isomorphism $F_t: y \mapsto F(t, y)$ maps Y_0 into Z_0 . Moreover, the connexity of T implies that the F_t 's map a given connected component of Y_0 into the same connected component of Z_0 . And the induction hypothesis applies to $Y - Y_0$ and $Z - Z_0$.

The corollary follows from the lemma, applied with $T = C(u)^r$, $Y = G^u$ and $Z = \mathfrak{g}^n$.

Note. The structure of $N_u/C(u)$ seems interesting. If I am not mistaken, it is the semi-direct product of G_m by a unipotent connected group V of dimension $\ell - 1$; moreover, the action of G_m on $\text{Lie } V$ has weights equal to $k_2 - 1, k_3 - 1, \dots, k_\ell - 1$, where the k_i 's are the exponents of the Weyl group.

Another interesting (and related) question is the behaviour of a Springer isomorphism f when one restricts f to $C(u)$. The tangent map to f is an endomorphism of $c(u) = \text{Lie } C(u)$. Is it always a non-zero multiple of the identity?

J.-P. Serre, June 1999

References

- [Bo98] Bourbaki, Nicolas, *Commutative algebra. Chapters 1–7*. Elements of Mathematics, translated from the French; reprint of the 1989 English translation, Springer-Verlag, Berlin 1998. Zbl 0902.13001 MR 1727221
- [Hu95] Humphreys, James E., *Conjugacy classes in semisimple algebraic groups*. Math. Surveys Monogr. 43, Amer. Math. Soc., Providence, RI, 1995. Zbl 0834.20048 MR 1343976
- [Ja04] Jantzen, Jens Carsten, Nilpotent orbits in representation theory. In *Lie Theory: Lie Algebras and Representations* (J.-P. Anker and B. Orsted, eds.), Progr. Math. 228 Birkhäuser, Boston 2004, 1–211. Zbl MR 2042688
- [KMRT] Knus, Max-Albert, Merkurjev, Alexander, Rost, Markus, Tignol, Jean-Pierre, *The book of involutions*. Amer. Math. Soc. Colloq. Publ. 44, Amer. Math. Soc., Providence, RI, 1998. Zbl 0955.16001 MR 1632779
- [Ke78] Kempf, George R., Instability in invariant theory. *Ann. of Math. (2)* **108** (1978), 299–316. Zbl 0406.14031 MR 0506989
- [Ko82] Kottwitz, Robert, Rational conjugacy classes in reductive groups. *Duke Math J.* **49** (1982), Zbl 0506.20017 MR 0683003
- [La93] Lang, Serge, *Algebra*. Third edition, Addison-Wesley, 1993. Zbl 0848.13001
- [Mc02] McNinch, George J., Abelian unipotent subgroups of reductive groups. *J. Pure Appl. Algebra* **167** (2002), 269–300; arXiv:math.RT/0007056. Zbl 0999.20035 MR 1874545
- [Mc03] McNinch, George J., Sub-principal homomorphisms in positive characteristic. *Math. Z.* **244** (2003), 433–455; arXiv:math.RT/0108140. Zbl 1036.20042 MR 1992546
- [Mc04] McNinch, George J., Nilpotent orbits over ground fields of good characteristic. *Math. Annalen* **329** (2004), 49–85; arXiv:math.RT/0209151. Zbl 02123994 MR 2052869
- [MS03] McNinch, George J., Sommers, Eric, Component groups of unipotent centralizers in good characteristic. *J. Algebra* **260** (2003), 323–337; arXiv:math.RT/0204275. Zbl 1026.20026 MR 1976698
- [Pr02] Premet, Alexander, Nilpotent orbits in good characteristic and the Kempf-Rousseau theory. *J. Algebra* **260** (2003), Zbl 1020.20031 MR 1976699
- [Sei00] Seitz, Gary M., Unipotent elements, tilting modules, and saturation. *Invent. Math.* **141** (2000), 467–502. Zbl 01515362 MR 1779618
- [Ser04] Serre, Jean-Pierre, Complète Réductibilité. *Astérisque*, Séminaire Bourbaki 2003/2004 Exposés **924/937**, S.M.F., to appear.
- [Spa84] Spaltenstein, Nicolas, Existence of good transversal slices to nilpotent orbits in good characteristic. *J. Fac. Sci. Univ. Tokyo* **31** (1984), 283–286. Zbl 0581.17003 MR 0763422
- [Spr69] Springer, Tonny A., The unipotent variety of a semi-simple group. In *Algebraic geometry* (Internat. Colloq., Tata Inst. Fund. Res., Bombay, 1968), Oxford University Press, London 1969, 373–391. Zbl 0195.50803 MR 0263830
- [Spr98] Springer, Tonny A., *Linear algebraic groups*. Second edition, Progr. Math. 9, Birkhäuser, Boston 1998. Zbl 0927.20024 MR 1642713

- [SS70] Springer, Tony A., Steinberg, Robert, Conjugacy classes. In *Seminar on algebraic groups and related finite groups* (The Institute for Advanced Study, Princeton, N.J., 1968/69), Lecture Notes in Math. 131, Springer-Verlag, Berlin 1970, 167–266. Zbl 0249.20024 MR 0268192
- [St67] Steinberg, Robert, *Lectures on Chevalley Groups*. Yale University mimeographed notes, 1967. Zbl 0307.22001 MR 0466335
- [St68] —, Endomorphisms of linear algebraic groups. Mem. Amer. Math Soc. **80** (1968), 1–108. Collected Papers no. 23. Zbl 0164.02902 MR 0230728
- [St65] —, Regular elements of semisimple algebraic groups, *Inst. Hautes Études Sci. Publ. Math.* **25** (1965), 49–80. Collected Papers no. 20. Zbl 0136.30002 MR 0180554

Received January 20, 2004

George J. McNinch, Department of Mathematics, Tufts University, 503 Boston Avenue,
Medford, MA 02155, U.S.A.
E-mail: george.mcninch@tufts.edu