Group Representation Theory, eds. M. Geck, D. Testerman, J. Thévenaz,
EPFL Press, Lausanne 2007, pp. 405-450

Bounds for the orders of the finite subgroups of $G(k)$

Jean-Pierre SERRE

Introduction

The present text reproduces - with a number of additions - a series of three two-hour lectures given at the Ecole Polytechnique Fédérale de Lausanne (E.P.F.L.) on May 25-26-27, 2005.

The starting point is a classical result of Minkowski, dating from 1887, which gives a multiplicative upper bound for the orders of the finite subgroups of $\mathbf{GL}_n(\mathbf{Q})$. The method can easily be extended to other algebraic groups than $\mathbf{GL}_n$, and the field $\mathbf{Q}$ can be replaced by any number field. What is less obvious is that:

a) one can work over an arbitrary ground field;

b) in most cases one may construct examples showing that the bound thus obtained is optimal.

This is what I explain in the lectures.

Lecture I is historical: Minkowski (§1), Schur (§2), Blichfeldt and others (§3). The results it describes are mostly well-known, so that I did not feel compelled to give complete proofs.

Lecture II gives upper bounds for the order of a finite ℓ-subgroup of $G(k)$, where G is a reductive group over a field k, and ℓ is a prime number. These bounds depend on G via its root system, and on k via the size of the Galois group of its ℓ-cyclotomic tower (§4). One of these bounds (called here the S-bound, cf. §5) is a bit crude but is easy to prove and to apply. The second one (called the M-bound) is the most interesting one (§6). Its proof follows Minkowski's method, combined with Chebotarev's density theorem (for schemes of any dimension, not merely dimension 1); it has a curious cohomological generalization cf. §6.8. The last subsection (§6.9) mentions some related problems, not on semisimple groups, but on Cremona groups; for instance: does the field $\mathbf{Q}(X, Y, Z)$ have an automorphism of order 11 ?

Lecture III gives the construction of "optimal" large subgroups. The case of the classical groups (§9) is not difficult. Exceptional groups such as E_8 are a different matter; to handle them, we shall use Galois twists, braid groups and Tits groups, cf. §§10-12.

Acknowledgements. A first draft of these notes, made by D. Testerman and R. Corran, has been very useful; and so has been the generous help of D. Testerman with the successive versions of the text. My thanks go to both of them, and to the E.P.F.L. staff for its hospitality. I also thank M. Broué and J. Michel for several discussions on braid groups.

J-P. Serre April 2006

Table of Contents

I. History: Minkowski, Schur, ...

§1. Minkowski

Reference: [Mi 87].

1.1. **Statements.** We shall use the following notation:

ℓ is a fixed prime number; when we need other primes we usually denote them by p;

the ℓ-adic valuation of a rational number x is denoted by $v_\ell(x)$; one has $v_\ell(\ell) = 1$, and $v_\ell(x) = 0$ if x is an integer with $(x, \ell) = 1$;

the number of elements of a finite set A is denoted by $|A|$; we write $v_\ell(A)$ instead of $v_\ell(|A|)$; if A is a group, $\ell^{v_\ell(A)}$ is the order of an ℓ-Sylow of A;

if x is a real number, its integral part ("floor") is denoted by $[x]$.

We may now state Minkowski's theorem ([Mi 87]):

Theorem 1. *Let n be an integer ≥ 1, and let ℓ be a prime number. Define*:

$$M(n,\ell) = \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \left[\frac{n}{\ell^2(\ell-1)}\right] + \cdots$$

Then:
(i) *If A is a finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$, we have $v_\ell(A) \leq M(n,\ell)$.*
(ii) *There exists a finite ℓ-subgroup A of $\mathbf{GL}_n(\mathbf{Q})$ with $v_\ell(A) = M(n,\ell)$.*

The proof will be given in §1.3 and §1.4.

Remarks.

1) Let us define an integer $M(n)$ by:

$$M(n) = \prod_\ell \ell^{M(n,\ell)}.$$

Part (i) of th.1 says that the order of any finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$ *divides* $M(n)$, and part (ii) says that $M(n)$ is the smallest integer having this property. Hence $M(n)$ is a sharp multiplicative bound for $|A|$.

Here are the values of $M(n)$ for $n \leq 8$:

$M(1) = 2$
$M(2) = 2^3 \cdot 3 = 24$
$M(3) = 2^4 \cdot 3 = 48$
$M(4) = 2^7 \cdot 3^2 \cdot 5 = 5760$
$M(5) = 2^8 \cdot 3^2 \cdot 5 = 11520$
$M(6) = 2^{10} \cdot 3^4 \cdot 5 \cdot 7 = 2903040$
$M(7) = 2^{11} \cdot 3^4 \cdot 5 \cdot 7 = 5806080$
$M(8) = 2^{15} \cdot 3^5 \cdot 5^2 \cdot 7 = 1393459200.$

Note that

$$M(n)/M(n-1) = \begin{cases} 2 & \text{if } n \text{ is odd} \\ \text{denominator of } b_n/n & \text{if } n \text{ is even,} \end{cases}$$

where b_n is the n-th Bernoulli number. (The occurence of the Bernoulli numbers is natural in view of the mass formulae which Minkowski had proved a few years before.)

2) One may ask whether there is a finite subgroup A of $\mathbf{GL}_n(\mathbf{Q})$ of order $M(n)$. It is so for $n = 1$ and $n = 3$ and probably for no other value of n (as Burnside already remarked on p.484 of [Bu 11]). Indeed, some incomplete arguments of Weisfeiler and Feit would imply that the upper bound of $|A|$ is $2^n \cdot n!$ if $n > 10$, which is much smaller than $M(n)$. See the comments of Guralnick-Lorenz in [GL 06], §6.1.

Exercise. Let $\left[\frac{n}{\ell-1}\right] = \sum a_i \ell^i, 0 \leq a_i \leq \ell - 1$, be the ℓ-adic expansion of $\left[\frac{n}{\ell-1}\right]$. Show that $M(n,\ell) = \sum a_i \frac{\ell^{i+1}-1}{\ell-1} = \sum M(a_i \ell^i(\ell-1), \ell)$.

1.2. **Minkowski's lemma.** Minkowski's paper starts with the following often quoted lemma:

Lemma 1. *If $m \geq 3$, the kernel of $\mathbf{GL}_n(\mathbf{Z}) \to \mathbf{GL}_n(\mathbf{Z}/m\mathbf{Z})$ is torsion free.*

Proof. Easy exercise ! One may deduce it from general results on formal groups over local rings, cf. Bourbaki [LIE III], §7. Many variants exist. For instance:

Lemma 1′. *Let R be a local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. If ℓ is a prime number distinct from $\mathrm{char}(k)$, the kernel of the map $\mathbf{GL}_n(R) \to \mathbf{GL}_n(k)$ does not contain any element of order ℓ.*

Proof. Suppose $x \in \mathbf{GL}_n(R)$ has order ℓ and gives 1 in $\mathbf{GL}_n(k)$. Write $x = 1 + y$; all the coefficients of the matrix y belong to $\mathfrak{m}$. Since $x^\ell = 1$, we have

$$\ell \cdot y + \binom{\ell}{2} \cdot y^2 + \cdots + \ell \cdot y^{\ell-1} + y^\ell = 0,$$

which we may write as $y \cdot u = 0$, with $u = \ell + \binom{\ell}{2} y + \cdots + y^{\ell-1}$. The image of u in $\mathbf{GL}_n(k)$ is ℓ, which is invertible. Hence u is invertible, and since $y \cdot u$ is 0, this shows that $y = 0$. □

Several other variants can be found in [SZ 96].

Remark. A nice consequence of lemma 1′ is the following result of Malcev and Selberg ([Bo 69], §17):

(*) *Let Γ be a finitely generated subgroup of $\mathbf{GL}_n(K)$, where K is a field of characteristic 0. Then Γ has a torsion free subgroup of finite index.*

Sketch of proof (for more details, see Borel, *loc.cit.*). Let S be a finite generating subset of Γ, and let L be the ring generated by the coefficients of the elements of $S \cup S^{-1}$. We have $\Gamma \subset \mathbf{GL}_n(L)$. Let $\mathfrak{m}$ be a maximal ideal of L; the residue field $k = A/\mathfrak{m}$ is finite ([AC V], p.68, cor.1 to th.3); let p be its characteristic. The kernel Γ_1 of $\Gamma \to \mathbf{GL}_n(k)$ has finite index in Γ; by lemma 1′ (applied to the local ring $R = L_\mathfrak{m}$), Γ_1 does not have any torsion except possibly p-torsion. By choosing another maximal ideal of L, with a different residue characteristic, one gets a torsion free subgroup of finite index of Γ_1, and hence of Γ. □

Remark. When K has characteristic $p > 0$ the same proof shows that Γ has a subgroup of finite index which is "p'-torsion free", i.e. such that its elements of finite order have order a power of p.

1.3. **Proof of theorem 1 (i).** Let A be a finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$; we have to show that $v_\ell(A) \leq M(n, \ell)$. Note first:

1.3.1. *The group A is conjugate to a subgroup of $\mathbf{GL}_n(\mathbf{Z})$.*

This amounts to saying that there exists an A-stable lattice in $\mathbf{Q}^n$, which is clear: just take the lattice generated by the A-transforms of the standard lattice $\mathbf{Z}^n$.

1.3.2. *There is a positive definite quadratic form on $\mathbf{Q}^n$, with integral coefficients, which is invariant by A.*

Same argument: take the sum of the A-transforms of $x_1^2 + \cdots + x_n^2$, and multiply it by a suitable non-zero integer, in order to cancel any denominator.

Let us now proceed with the proof of $v_\ell(A) \leq M(n,\ell)$. We do it in two steps:

1.3.3. *The case $\ell > 2$.*

By 1.3.1, we may assume that A is contained in $\mathbf{GL}_n(\mathbf{Z})$. Let p be a prime number $\neq 2$. By lemma 1, the map $A \to \mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is injective. Hence

$$v_\ell(A) \leq a(p) = v_\ell\big(\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})\big).$$

The order of $\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is $p^{n(n-1)/2}(p-1)(p^2-1)\dots(p^n-1)$. Let us assume that $p \neq \ell$. Then we have

$$a(p) = \sum_{i=1}^{n} v_\ell(p^i - 1).$$

We now choose p in such a way that $a(p)$ is as small as possible. More precisely, we choose p such that:

(*) *The image of p in $(\mathbf{Z}/\ell^2\mathbf{Z})^*$ is a generator of that group.*

This is possible by Dirichlet's theorem on the existence of primes in arithmetic progressions (of course, one should also observe that $(\mathbf{Z}/\ell^2\mathbf{Z})^*$ is cyclic.)

Once p is chosen in that way, then $p^i - 1$ is divisible by ℓ only if i is divisible by $\ell - 1$; moreover, one has $v_\ell(p^{\ell-1} - 1) = 1$ because of (*), and this implies that $v_\ell(p^i - 1) = 1 + v_\ell(i)$ if i is divisible by $\ell - 1$. (This is where the hypothesis $\ell > 2$ is used.) One can then compute $a(p)$ by the formula above. The number of indices $i \leq n$ which are divisible by $\ell - 1$ is $\left[\frac{n}{\ell-1}\right]$. We thus get:

$$\begin{aligned} a(p) &= \left[\frac{n}{\ell-1}\right] + \sum_{1 \leq j \leq \left[\frac{n}{\ell-1}\right]} v_\ell(j) = \left[\frac{n}{\ell-1}\right] + v_\ell\Big(\left[\frac{n}{\ell-1}\right]!\Big) \\ &= \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \cdots = M(n,\ell). \end{aligned}$$

This proves th.1 (i) in the case $\ell \neq 2$.

1.3.4. *The case $\ell = 2$.*

When $\ell = 2$, the method above does not give the right bound as soon as $n > 1$. One needs to replace $\mathbf{GL}_n$ by an orthogonal group. Indeed, by 1.3.1 and 1.3.2, we may assume, not only that A is contained in $\mathbf{GL}_n(\mathbf{Z})$, but also that it is contained in the orthogonal group $\mathbf{O}_n(q)$, where q is a non-degenerate quadratic form with integral coefficients. Let D be the discriminant of q, and let us choose a prime number $p > 2$ which does not divide D. The image of A in $\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is contained in the orthogonal

group $\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})$ relative to the reduction of q mod p. If we put $r = [n/2]$, the order of $\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})$ is known to be:

$$2 \cdot p^{r^2}(p^2-1)(p^4-1)\dots(p^{2r}-1) \quad \text{if } n \text{ is odd.}$$

and

$$2 \cdot p^{r(r-1)}(p^2-1)(p^4-1)\dots(p^{2r}-1)/(p^r+\varepsilon) \quad \text{if } n \text{ is even,}$$

with $\varepsilon = \pm 1$ equal to the Legendre symbol at p of $(-1)^r D$.

If we choose $p \equiv \pm 3 \pmod 8$, we have $v_2(p^{2i}-1) = 3 + v_2(i)$, and $v_2(p^r+\varepsilon) \geq 1$. If n is odd, this gives

$$v_2(\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})) = 1 + 3r + v_2(r!) = n + r + \left[\frac{r}{2}\right] + \left[\frac{r}{4}\right] + \dots = M(n,2),$$

and, if n is even:

$$v_2(\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})) \leq 3r + v_2(r!) = M(n,2).$$

Hence $v_2(A)$ is at most equal to $M(n,2)$. □

Remark. There are several ways of writing down this proof. For instance:

- There is no need to embed A in $\mathbf{GL}_n(\mathbf{Z})$. It sits in $\mathbf{GL}_n(\mathbf{Z}[1/N])$ for a suitable $N \geq 1$, and this allows us to reduce mod p for all p's not dividing N.

- Minkowski's lemma is not needed either: we could replace it by the trivial fact that a matrix which is different from 1 is not congruent to 1 mod p for all large enough p's.

- Even when $\ell > 2$, we could have worked in $\mathbf{O}_n$ instead of $\mathbf{GL}_n$; that is what Minkowski does.

- When $\ell = 2$ the case n even can be reduced to the case n odd by observing that, if $A \subset \mathbf{GL}_n(\mathbf{Q})$, then $A \times \{\pm 1\}$ embeds into $\mathbf{GL}_{n+1}(\mathbf{Q})$, and $M(n+1,2)$ is equal to $1 + M(n,2)$.

1.4. **Proof of theorem 1 (ii).** The symmetric group S_ℓ has a faithful representation $S_\ell \to \mathbf{GL}(V_1)$ where V_1 is a $\mathbf{Q}$-vector space of dimension $\ell - 1$. Put $r = \left[\frac{n}{\ell-1}\right]$, and let $V = V_1 \oplus \dots \oplus V_r$ be the direct sum of r copies of V_1. Let S be the semi-direct product of S_r with the product $(S_\ell)^r$ of r copies of S_ℓ ("wreath product"). The group S has a natural, and faithful, action on V. We may thus view S as a subgroup of $\mathbf{GL}_{r(\ell-1)}(\mathbf{Q})$, hence also of $\mathbf{GL}_n(\mathbf{Q})$, since $n \geq r(\ell-1)$. We have

$$v_\ell(S) = r + v_\ell(r!) = \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \dots = M(n,\ell).$$

An ℓ-Sylow A of S satisfies the conditions of th.1 (ii). □

Example. When $\ell = 2$ the group S defined above is the "hyper-octahedral group", i.e. the group of automorphisms of an n-cube (= the Weyl group of a root system of type B_n); in ATLAS notation, it may be written as $2^n \cdot S_n$.

1.5. **A conjugacy theorem.** The finite ℓ-subgroups of $\mathbf{GL}_n(\mathbf{Q})$ have the following Sylow-like property:

Theorem 1′. *Let A and A' be two finite ℓ-subgroups of $\mathbf{GL}_n(\mathbf{Q})$. Assume that A has the maximal order allowed by th.*1*. Then A' is conjugate to a subgroup of A.*

Corollary. *If $|A| = |A'| = \ell^{M(n,\ell)}$, then A and A' are conjugate in $\mathbf{GL}_n(\mathbf{Q})$.*

Proof of theorem 1′. See Bourbaki, [LIE III], §7, exerc.6 f) where only the case $\ell > 2$ is given, and Feit [Fe 97] who does the case $\ell = 2$. Let us sketch Bourbaki's method (which we shall use in §6.6 in a more general setting):

We may assume that A and A' are contained in $\mathbf{GL}_n(\mathbf{Z})$. Choose a prime p as in 1.3.3, and reduce mod p. The groups A and A' then become ℓ-subgroups of $G_p = \mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$, and A is an ℓ-Sylow of G_p. By Sylow's theorem applied to G_p, one finds an injection $i : A' \to A$ which is induced by an inner automorphism of G_p. The two linear representations of A':

$$A' \to \mathbf{GL}_n(\mathbf{Q}) \quad \text{and} \quad A' \xrightarrow{i} A \to \mathbf{GL}_n(\mathbf{Q})$$

become isomorphic after reduction mod p. Since $p \neq \ell$, a standard argument shows that they are isomorphic over $\mathbf{Q}$, which proves th.1′ in that case. The case $\ell = 2$ can be handled by a similar, but more complicated, argument: if n is odd, one uses orthogonal groups as in 1.3.4, and one reduces the case n even to the case n odd by the trick mentioned at the end of §1.3. □

Exercise. Let $A(n)$ be a maximal 2-subgroup of $\mathbf{GL}_n(\mathbf{Q})$. Show that the $A(n)$'s can be characterized by the following three properties:

$$\begin{aligned} A(1) &= \{\pm 1\}. \\ A(2n) &= \big(A(n) \times A(n)\big) \cdot \{\pm 1\} \text{ (wreath product) if } n \text{ is a power of 2.} \\ A(n) &= A(2^{m_1}) \times \cdots \times A(2^{m_k}) \text{ if } n = 2^{m_1} + \cdots + 2^{m_k} \text{ with } m_1 < \cdots < m_k. \end{aligned}$$

§2. Schur

Ten years after [Mi 87], Frobenius founded the theory of characters of finite groups. It was then (and still is now) very tempting to use that theory to give a different proof of Minkowski's results. The first people to do so were Schur ([Sch 05]) and Burnside ([Bu 11], Note G). Schur's paper is especially interesting. He works first over $\mathbf{Q}$, as Minkowski did, and uses a very original argument in character theory, see §2.1 below. He then attacks the case of an arbitrary number field, where he gets a complete answer, see §2.2.

2.1. **Finite linear groups with rational trace.** What Schur proves in §1 of [Sch 05] is:

Theorem 2. *Let A be a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$. Assume that the traces of the elements of A lie in $\mathbf{Q}$. Then $v_\ell(A) \leq M(n,\ell)$, where $M(n,\ell)$ is as in th.*1*.*

The condition on the traces is obviously satisfied if A is contained in $\mathbf{GL}_n(\mathbf{Q})$. Hence th.2 is a generalization of th.1. (As a matter of fact, it is a genuine generalization only when $\ell = 2$; indeed, when $\ell > 2$, it is known, cf. [Ro 58], that a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$ with rational trace is conjugate to a subgroup of $\mathbf{GL}_n(\mathbf{Q})$.)

Proof. We start from the following general fact, which is implicit in [Sch 05] (and is sometimes called "Blichfeldt's lemma"):

Proposition 1. *Let G be a finite subgroup of $\mathbf{GL}_n(\mathbf{C})$ and let X be the subset of $\mathbf{C}$ made up of the elements $\mathrm{Tr}(g)$ for $g \in G, g \neq 1$. Let $N = \prod(n - x)$ be the product of the $n - x$, for $x \in X$. Then N is a non-zero integer which is divisible by $|G|$.*

(Hence the knowledge of the set X gives a multiplicative bound for the order of G.)

Proof. Let $m = |G|$, and let z be a primitive m-th root of unity. The elements of X are sums of powers of z; hence they belong to the ring of integers of the cyclotomic field $K = \mathbf{Q}(z)$. This already shows that N is an algebraic integer. If s is an element of $\mathrm{Gal}(K/\mathbf{Q})$, one has $s(z) = z^a$ for some $a \in (\mathbf{Z}/m\mathbf{Z})^*$. If $x = \mathrm{Tr}(g)$, with $g \in G$, then $s(x) = \mathrm{Tr}(g^a)$, hence $s(x)$ belongs to X. This shows that X is stable under the action of $\mathrm{Gal}(K/\mathbf{Q})$; hence N is fixed by $\mathrm{Gal}(K/\mathbf{Q})$; this proves that N belongs to $\mathbf{Z}$.

The factors of N are $\neq 0$. Indeed, $\mathrm{Tr}(g)$ is equal to the sum of n complex numbers z_i with $|z_i| = 1$, hence can be equal to n only if all the z_i are equal to 1, which is impossible since $g \neq 1$. This shows that $N \neq 0$ (one could also prove that N is positive, but we shall not need it).

It remains to see that N is divisible by $|G|$. It is well-known that, if χ is a generalized character of G, the sum $\sum_{g\in G} \chi(g)$ is divisible by $|G|$. Let us apply this to the function $g \mapsto \chi(g) = \prod_{x\in X} \big(\mathrm{Tr}(g) - x\big)$, which is a $\mathbf{Z}$-linear combination of the characters $g \mapsto \mathrm{Tr}(g)^m, m \geq 0$. Since $\chi(g) = 0$ for $g \neq 1$ and $\chi(1) = N$, the sum of the $\chi(g)$ is equal to N. Hence N is divisible by $|G|$. □

The next lemma gives an information on the $\mathrm{Tr}(g)$'s:

Lemma 2. *Let A be as in th.2. If $g \in A$, then $\mathrm{Tr}(g)$ may be written as $n - \ell y$ with $y \in \mathbf{Z}$ and $0 \leq y \leq n/(\ell - 1)$.*

Proof. Each eigenvalue of g is of order ℓ^α for some $\alpha \geq 0$, and all the eigenvalues with the same α have the same multiplicity. By splitting $\mathbf{C}^n$ according to the α's, one is reduced to the following three cases:
(1) $g = 1$ and $n = 1$. Here $\mathrm{Tr}(g) = 1$ and we take $y = 0$.
(2) g has order ℓ and $n = \ell - 1$. Here $\mathrm{Tr}(g) = -1$, and $y = 1$.
(3) g has order ℓ^α with $\alpha > 1$ and $n = \ell^{\alpha-1}(\ell - 1)$. Here $\mathrm{Tr}(g) = 0$ and $y = \ell^{\alpha-2}(\ell - 1)$.
In each case we have $0 \leq y \leq n/(\ell - 1)$. □

End of the proof of theorem 2. We apply prop.1 to $G = A$. By lemma 2, each factor $n - x$ of N can be written as ℓy with $1 \leq y \leq d = [n/\ell - 1)]$.

This shows that N divides the product $\ell^d \cdot d!$ and we have

$$v_\ell(N) < d + v_\ell(d!) = [n/(\ell-1)] + [n/\ell(\ell-1)] + \cdots = M(n,\ell).$$

Since $|G|$ divides N, this proves th.2. □

Remark. One may ask whether th.2 can be complemented by a conjugacy theorem analogous to th.1′ of §1.5. The answer is of course "yes" if $\ell > 2$ (because of th.1′), but it is "no" for $\ell = 2$: the dihedral group D_4 and the quaternion group Q_8 are non-conjugate 2-subgroups of $\mathbf{GL}_2(\mathbf{C})$, with rational trace, which have the maximal order allowed by th.2, namely 8.

2.2. **Replacing Q by an arbitrary number field.** This is what Schur does in §§2-6 of [Sch 05]. Before stating his result, some notation is necessary:

- k is a number field, viewed as a subfield of $\mathbf{C}$.
- For each $a \geq 1$, z_a denotes a primitive a-th root of unity.
- (assuming $\ell > 2$). We put $t = [k(z_\ell) : k]$ and we denote by m the maximal a such that $k(z_\ell)$ contains $z_{\ell a}$ (this notation coincides with Schur's, and it will be extended to arbitrary fields in §4 of Lect.II). We put

$$M_k(n,\ell) = m \cdot \left[\frac{n}{t}\right] + \left[\frac{n}{\ell t}\right] + \left[\frac{n}{\ell^2 t}\right] + \dots$$

- (assuming $\ell = 2$). We put $t = [k(i) : k]$ and we define m as explained in §4.2 (warning: t and m do not always coincide with Schur's t_2 and m_2). We put:

$$M_k(n,2) = n + (m'-1)\left[\frac{n}{t}\right] + \left[\frac{n}{2t}\right] + \left[\frac{n}{4t}\right] + \dots,$$

where m' is equal to $m+1$ in case (b) of §4.2 and is equal to m in the other cases.

The main result of [Sch 05] is:

Theorem 2′. *Let A be a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$ such that $\mathrm{Tr}(g)$ belongs to k for every $g \in A$. Then $v_\ell(A) \leq M_k(n,\ell)$.*

Note that, when $k = \mathbf{Q}$, the integer $M_k(n,\ell)$ is equal to Minkowski's $M(n,\ell)$; hence th.2′ is a generalization of th.2.

Proof. I shall not give all the details of Schur's proof, but just explain its main steps. For more information, see [Sch 05] (and also [GL 06] for the case $\ell > 2$).

One of the inputs of the proof is the following result, which had just been proved by Blichfeldt ([Bl 04] - see also §3 below):

2.2.1. *Every linear representation of A is monomial.*

Hence one can decompose the vector space $\mathbf{C}^n$ as a direct sum of n lines $D_1, \dots, D_n$ which are permuted by A. This gives a homomorphism $A \to S_n$; its kernel A' is a normal abelian subgroup of A. Hence:

2.2.2. *The group A has a normal abelian subgroup A' such that $(A : A')$ divides $n!$.*

This led Schur to investigate the case where A is abelian. He proved:

2.2.3. *If A is as in th.2′, and is abelian, then* :

$$v_\ell(A) \leq \begin{cases} m \cdot \left[\frac{n}{t}\right] & \text{if } \ell > 2 \\ (m'-t) \cdot \left[\frac{n}{t}\right] + n & \text{if } \ell = 2. \end{cases}$$

Sketch of proof. Since A is abelian, and the traces of its elements belong to k, it is conjugate to a subgroup of $\mathbf{GL}_n(k)$. Let R be the k-subalgebra of $\mathbf{M}_n(k)$ generated by A. We may write R as a product $\prod K_i$, where the K_i are cyclotomic extensions of k, of the form $k(z_{\ell^{a_i}})$, with $a_i \geq 0$. Let $n_i = [K_i : k]$; then $\sum n_i \leq n$. The image of A in K_i^* is a cyclic group of order ℓ^{a_i}. If $\ell > 2$, it is not difficult to see that $a_i \leq m \cdot \left[\frac{n_i}{t}\right]$ for every i. Adding up, we find $\sum a_i \leq m \cdot \left[\frac{n}{t}\right]$, and since $v_\ell(A) \leq \sum a_i$, we get the inequality (2.2.3). The case $\ell = 2$ is similar. □

Once this is done, the case $\ell = 2$ follows. Indeed (2.2.2) and (2.2.3) give $v_2(A) \leq v_2(A') + v_2(n!) \leq n + (m'-t) \cdot \left[\frac{n}{t}\right] + v_2(n!)$, and this is equivalent to $v_2(A) \leq M_k(n, 2)$. The case $\ell > 2$ requires more work, cf. [Sch 05], §5. □

Remarks

1) The bound $v_\ell(A) \leq M_k(n, \ell)$ is *optimal*; this is proved by the same explicit constructions as in §1.4, cf. [Sch 05], §6.

2) As we already pointed out in §2.1, the hypothesis $\mathrm{Tr}(A) \subset k$ implies, when $\ell > 2$, that A is conjugate to a subgroup of $\mathbf{GL}_n(k)$. One may then use Minkowski's method, as will be explained in §6 for semisimple algebraic groups (of course $\mathbf{GL}_n$ is not semisimple, but the method applies with almost no change – the invariant degrees d_i of §6 have to be replaced by $1, 2, \ldots, n$). The bound found in that way coincides with Schur's.

For $\ell = 2$, if one does not assume that A can be embedded in $\mathbf{GL}_n(k)$, I do not see how to apply either Minkowski's method or the cohomological method of §6.8. This raises interesting questions. For instance, consider a finite subgroup A of $E_8(\mathbf{C})$, and suppose that the conjugacy classes of the elements of A are $\mathbf{Q}$-rational. Is it true that $v_2(A) \leq 30$, $v_3(A) \leq 13, \ldots$, as would be the case if A were contained in the rational points of a $\mathbf{Q}$-form of E_8, cf. §6.3.2 ?

§3. Blichfeldt and others

Blichfeldt's theorem (§3.1 below) has already been used in §2.2. The results of §3.3 will be applied in §5.4, in order to prove what I call the "S-bound".

3.1. **Blichfeldt's theorem.** Recall that a finite group A is called *supersolvable* if it has a composition series

$$1 = A_0 \subset A_1 \subset \cdots \subset A_m = A$$

where the A_i are normal in A (and not merely in A_{i+1}) and the quotients A_i/A_{i-1} are cyclic. One has

$$\text{nilpotent} \Rightarrow \text{supersolvable} \Rightarrow \text{solvable}.$$

In particular, an ℓ-group is supersolvable.

One proves easily:

(*) *If A is supersolvable and non abelian, there exists an abelian normal subgroup A' of A which is not contained in the center of A.*

Recall also that a linear representation V of a group A is called *monomial* if one can split V as a direct sum of lines which are permuted by A. When V is irreducible, this amounts to saying that V is induced by a 1-dimensional representation of a subgroup of A.

We may now state Blichfeldt's theorem ([Bl 04], see also [Bu 11], §258):

Theorem 3. *Every complex linear representation of a supersolvable finite group is monomial.*

(As a matter of fact, Blichfeldt was only interested in the case where A is nilpotent.)

Proof. The argument is now standard. We may assume that the given representation V is irreducible and faithful. If A is abelian, we have $\dim V = 1$ and there is nothing to prove. If not, we choose A' as in (*) above, and we split V as $V = \oplus V_\chi$, where χ runs through the 1-dimensional characters of A', and V_χ is the corresponding eigenspace; let V_ψ be a non-zero V_χ; it is distinct from V (otherwise, A' would be central), and every non-zero V_χ is an A-transform of V_ψ (because V is irreducible). Call B the subgroup of A stabilizing V_ψ. We have $A' \subset B \subset A$, and $|B| < |A|$. Using induction on $|A|$, we may assume that th.3 is true for B; this gives a splitting of V_ψ as a direct sum of lines which are stable under B. By transforming them by A, we get the desired splitting of V. □

3.2. **Borel-Serre.** In [BS 53], Borel and I proved:

Theorem 3′. *Let G be a compact real Lie group, and let A be a finite supersolvable subgroup of G. There exists a maximal torus T of G which is normalized by A.*

Remark. When one applies th.3′ to $G = \mathbf{U}_n(\mathbf{C})$, one recovers th.3. Hence th.3′ may be viewed as a generalization of Blichfeldt's theorem.

Proof of theorem 3′ (sketch).

Lemma 3. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra over a field of characteristic 0, and let s be an automorphism of prime order of $\mathfrak{g}$. If s has no fixed point $\neq 0$, then $\mathfrak{g}$ is nilpotent.*

(Note the analogy with a - much deeper - theorem of Thompson [Th 60-64]: if a finite group G has an automorphism of prime order with no non-trivial fixed point, then G is nilpotent.)

Proof of lemma 3. By extending scalars, we may assume that the ground field is algebraically closed. Let p be the order of s, and let z be a primitive p-th root of unity. Let $\mathfrak{g}_i$ be the kernel of $s - z^i$ in $\mathfrak{g}$. We have

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_{p-1}\,,$$

and the hypothesis made on s means that $\mathfrak{g}_0 = 0$. One then shows that $\mathrm{ad}(x)^{p-1} = 0$ for every x belonging to one of the $\mathfrak{g}_i$'s. This implies that the

Killing form of $\mathfrak{g}$ is 0, hence that $\mathfrak{g}$ is solvable (Cartan's criterion). The fact that $\mathfrak{g}$ is nilpotent follows easily. (For more details, see §4 of [BS 54].) □

Once this is done, th.3′ is proved by an induction argument similar to the one used in the proof of Blichfeldt's theorem, cf. [BS 53], §3. □

3.3. **Steinberg and Springer-Steinberg.** We now come to the setting of linear algebraic groups. Let k be a field, and let G be an algebraic group over k. We shall assume in what follows that G is linear and smooth over k; the connected component of the identity of G is denoted by G°. Recall that G is said to be *reductive* if it is connected and if its unipotent radical (over an algebraic closure of k) is trivial, cf. [Bo 91], §11.21. If $k = \mathbf{C}$, such groups correspond (by a standard dictionary, cf. [Se 93], §5) to the connected compact Lie groups. [In the literature, a group G such that G° is reductive is sometimes called "reductive"; this is reasonable in characteristic 0, but not otherwise. Here we prefer that "reductive" implies "connected".]

Theorem 3′ has the following analogue:

Theorem 3″. *Let A be a finite supersolvable group of order prime to* $\mathrm{char}(k)$ *and let G be a reductive group over k on which A acts by k-automorphisms. Then there exists a maximal torus T of G, defined over k, which is stable under the action of A.*

(When $k = \mathbf{C}$, this is equivalent to th.3′, thanks to the dictionary mentioned above.)

Corollary. *If A is a finite supersolvable subgroup of $G(k)$, of order prime to* $\mathrm{char}(k)$*, there is a maximal k-torus T of G whose normalizer N is such that A is contained in $N(k)$.*

(Recall that, if X is a k-variety, $X(k)$ is the set of k-points of X.)

Proof of theorem 3″. When k is algebraically closed, this is proved in [SS 68], I.5.16, with the help of several results from [St 68]. For an arbitrary field k, the same proof works with very little change. One starts with the following basic result of Steinberg ([St 68], th.7.2):

Proposition 2. *Assume k is algebraically closed. Let $s : G \to G$ be a surjective homomorphism. Then there exists a Borel subgroup B of G such that $s(B) = B$.*

When s has finite order prime to $\mathrm{char}(k)$, one can say much more:

Proposition 3. *Let s be an automorphism of G of finite order prime to* $\mathrm{char}(k)$*, and let G^s be the subgroup of G fixed by s. Then*:

a) *The connected component of G^s is reductive.*

b) *One has* $\dim G^s > 0$ *if G is not a torus.*

c) *If k is algebraically closed, there exists a Borel subgroup B of G and a maximal torus T of B such that $s(B) = B$ and $s(T) = T$.*

Proof (*sketch*). We may assume k is algebraically closed, since assertions a) and b) are "geometric". A proof of a) is given in [St 68], cor.9.4. A proof of c) is given in [SS 68], I.2.9, as an application of prop.2. Assertion b) follows from c) by the following method of Steinberg: one observes that a pair (B, T)

with $B \supset T$, determines *canonically* a homomorphism $h : \mathbf{G}_m \to T$ (indeed B gives a basis of the root system of (G, T), and one takes for h twice the sum of the corresponding coroots). Moreover, h is non-trivial if G is not a torus. The canonicity of h implies that it is fixed by s. Hence G^s contains $\mathrm{Im}(h)$. □

End of the proof of th.3″. By induction on $|A| + \dim G$. When $A = 1$, one takes for T any maximal k-torus of G; by a theorem of Grothendieck, there is such a torus (cf. [Bo 91], th.18.2). We may thus assume $A \neq 1$. In that case A contains a cyclic subgroup $< s >$, non-trivial, which is normal. We may also assume that G is semisimple and that A acts faithfully. Let G_1 be the connected component of G^s; we have $\dim G_1 > 0$, cf. prop.3 b). The group A/A' acts on G_1. By the induction assumption, there is a maximal torus T_1 of G_1, defined over k, which is stable under the action of A/A', hence of A. Let G_2 be the centralizer of T_1 in G. It is a reductive group of the same rank as G. We have $\dim G_2 < \dim G$, since T_1 is not contained in the center of G. Moreover, G_2 is stable under the action by A. By applying the induction assumption to the pair (G_2, A) we get a maximal k-torus T of G_2 which is A-stable. Since G_2 and G have the same rank, T is a maximal torus of G. □

II. Upper bounds

Let G be a reductive group over a field k, and let ℓ be a prime number, different from $\mathrm{char}(k)$. Let A be a finite subgroup of $G(k)$. We want to give an upper bound for $v_\ell(A)$, in terms of invariants of G, k and ℓ. We give two such bounds. The first one (§5) is less precise, but very easy to apply; we call it the S-bound (S for Schur). The other bound (§6) is the M-bound (M for Minkowski). Both bounds involve some cyclotomic invariants of k, which are defined in §4 below.

§4. The invariants t and m

4.0. **Cyclotomic characters.** Let $\bar{k}$ be an algebraic closure of k, and let k_s be the separable closure of k in $\bar{k}$. For each $n \geq 1$ prime to $\mathrm{char}(k)$, let $\mu_n \subset k_s^*$ be the group of n-th roots of unity and let z_n be a generator of μ_n.

The Galois group $\Gamma_k = \mathrm{Gal}(k_s/k)$ acts on $\langle z_n \rangle = \mu_n$. This action defines a continuous homomorphism

$$\chi_n : \Gamma_k \to \mathrm{Aut}(\mu_n) = (\mathbf{Z}/n\mathbf{Z})^*,$$

which is called the *n-th cyclotomic character of k*.

This applies in particular to $n = \ell^d$ $(d = 0, 1, \dots)$; by taking inverse limits we get the ℓ^∞-cyclotomic character

$$\chi_{\ell^\infty} : \Gamma_k \to \mathbf{Z}_\ell^* = \varprojlim (\mathbf{Z}/\ell^d\mathbf{Z})^*,$$

where $\mathbf{Z}_\ell$ is the ring of ℓ-adic integers. What matters for us is the image $\mathrm{Im}\ \chi_{\ell^\infty}$, which is a closed subgroup of $\mathbf{Z}_\ell^*$. To discuss its structure, it is convenient to separate the cases $\ell \neq 2$ and $\ell = 2$.

4.1. **The case $\ell \neq 2$.** We have

$$\mathbf{Z}_\ell^* = C_{\ell-1} \times \{1 + \ell \cdot \mathbf{Z}_\ell\}$$

where $C_{\ell-1}$ is cyclic of order $\ell - 1$ (i.e. $C_{\ell-1}$ is the group $\mu_{\ell-1}$ of the ℓ-adic field $\mathbf{Q}_\ell$; it is canonically isomorphic to $\mathbf{F}_\ell^*$). As for $1 + \ell \cdot \mathbf{Z}_\ell$, it is procyclic, generated by $1 + \ell$, and isomorphic to the additive group $\mathbf{Z}_\ell$; its closed subgroups are the groups $1 + \ell^d \cdot \mathbf{Z}_\ell = \langle 1 + \ell^d \rangle$, $d = 1, 2, \ldots, \infty$, with the convention $\ell^\infty = 0$.

Since $\ell - 1$ and ℓ are relatively prime, the subgroup $\operatorname{Im} \chi_{\ell^\infty}$ of $\mathbf{Z}_\ell^*$ decomposes as a direct product:

$$\operatorname{Im} \chi_{\ell^\infty} = C_t \times \{1 + \ell^m \cdot \mathbf{Z}_\ell\} \ ,$$

where t is a divisor of $\ell - 1$, C_t is cyclic of order t and $m = 1, 2, \ldots$ or ∞.

Remark. An alternative definition of the invariants t and m is:

$$\begin{aligned} t &= [k(z_\ell) : k] = k\text{-degree of } z_\ell \\ m &= \text{upper bound of the } d \geq 1 \text{ such that } z_{\ell^d} \text{ is contained in } k(z_\ell). \end{aligned}$$

Examples. If $k = \mathbf{Q}$ or $\mathbf{Q}_\ell$, χ_{ℓ^∞} is surjective and we have $t = \ell - 1$, $m = 1$. If $k = k_s$, then χ_{ℓ^∞} is trivial and $t = 1$, $m = \infty$. If k is finite with q elements, $\operatorname{Im} \chi_{\ell^\infty}$ is the closed subgroup of $\mathbf{Z}_\ell^*$ generated by q and we have:

$$\begin{aligned} t &= \text{order of } q \text{ in } \mathbf{F}_\ell^* \\ m &= v_\ell(q^t - 1) = v_\ell\left(q^{\ell-1} - 1\right). \end{aligned}$$

4.2. **The case $\ell = 2$.** Here $\mathbf{Z}_2^* = C_2 \times \{1 + 4 \cdot \mathbf{Z}_2\}$, where $C_2 = \{1, -1\}$ and the multiplicative group $1 + 4 \cdot \mathbf{Z}_2$ is isomorphic to the additive group $\mathbf{Z}_2$. There are three possibilities for $\operatorname{Im} \chi_{2^\infty}$:

(a) $\operatorname{Im} \chi_{2^\infty} = 1 + 2^m \cdot \mathbf{Z}_2 = \langle 1 + 2^m \rangle$, with $m = 2, \ldots, \infty$. We put $t = 1$.
(b) $\operatorname{Im} \chi_{2^\infty} = \langle -1 + 2^m \rangle$, with $m = 2, \ldots, \infty$. We put $t = 2$.
(c) $\operatorname{Im} \chi_{2^\infty} = C_2 \times \{1 + 2^m \cdot \mathbf{Z}_2\} = \langle -1, 1 + 2^m \rangle$, $m = 2, \ldots, \infty$. We put $t = 2$.

If $m < \infty$, these types are distinct. If $m = \infty$, types (b) and (c) coincide; in that case $\operatorname{Im} \chi_{2^\infty}$ is equal to C_2.

Remark. We have $t = [k(i) : k]$ with the usual notation $i = z_4$. Hence case (a) means that -1 is a square in k, and in that case m is the largest $d \geq 2$ such that $z_{2^d} \in k$.

If $t = 2$, case (c) is characterized by the fact that -1 belongs to $\operatorname{Im} \chi_{2^\infty}$. As for m, it is given by:

$$\begin{aligned} m &= -1 + \text{ upper bound of the } d \geq 2 \text{ such that } z_{2^d} \in k(i) \text{ in case (b)} \\ m &= \text{upper bound of the } d \geq 2 \text{ such that } z_{2^d} \in k(i) \text{ in case (c)}. \end{aligned}$$

Examples. If $k = \mathbf{Q}$ or $\mathbf{Q}_2$, we have type (c) with $t = 2, m = 2$. If $k = \mathbf{R}$, we have types (b) and (c) with $m = \infty$. If k is separably closed, we have type (a) with $t = 1$ and $m = \infty$.

When $\operatorname{char}(k) \neq 0$, type (c) is impossible unless $m = \infty$. If k is finite with q elements, we have type (a) with $m = v_2(q - 1)$ if $q \equiv 1 \pmod 4$ and type (b) with $m = v_2(q + 1)$ if $q \equiv -1 \pmod 4$.

4.3. **The case of finitely generated fields.** Let k_0 be the prime subfield of k, i.e. $\mathbf{Q}$ if $\mathrm{char}(k) = 0$ or $\mathbf{F}_p$ if $\mathrm{char}(k) = p > 0$. Suppose that k is *finitely generated over* k_0. Then *the invariant* m *is finite*, i.e. $\mathrm{Im}\,\chi_{\ell^\infty}$ is infinite.

Indeed, if not, there would be a finite extension k' of k containing the group μ of all the ℓ^d-th roots of unity ($d = 1, 2, \dots$). Let $K = k_0(\mu)$ be the extension of k_0 generated by μ. Then:

(a) K is algebraic over k_0
(b) K is finitely generated over k_0 (because it is contained in k', cf. [A V], §14, cor. 3 to prop. 17).

Hence K is either a finite field or a number field, which is absurd since such a field only contains finitely many roots of unity.

§5. The S-bound

We start with the case of tori:

5.1. **The S-bound for a torus: statements.**

Theorem 4. *Let T be a torus over k, and let A be a finite subgroup of $T(k)$. Then*

$$v_\ell(A) \,\le\, m \left[\frac{\dim T}{\varphi(t)}\right],$$

where m and t are defined as in §4 above and φ is Euler's totient function.

The bound given by th.4 is optimal. More precisely:

Theorem 4′. *Assume $m < \infty$. For every $n \ge 1$ there exist a k-torus T of dimension n and a finite subgroup A of $T(k)$ such that $v_\ell(A) = m \cdot [n/\varphi(t)]$.*

Example. Take $k = \mathbf{Q}$ and $\ell = 2$, so that $t = m = 2$. Then th.4 says that any finite 2-subgroup of $T(\mathbf{Q})$ has order $\le 4^{\dim T}$, and th.4′ says that this bound can be attained.

5.2. **Proof of theorem 4.**

Lemma 4. *Let $u \in \mathbf{M}_n(\mathbf{Z}_\ell)$ be an $n \times n$ matrix with coefficients in $\mathbf{Z}_\ell$, which we view as an endomorphism of $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^n$. Then*

$$v_\ell\big(\ker(u)\big) = v_\ell\big(\det(u)\big).$$

Proof. This is clear if u is a diagonal matrix, and one reduces the general case to the diagonal one by multiplying u on the right and on the left by invertible matrices. □

Now let n be the dimension of the torus T. Let $Y(T) = \mathrm{Hom}_{k_s}(\mathbf{G}_m, T)$ be the group of cocharacters of T. The action of Γ_k on $Y(T)$ gives a homomorphism $\rho : \Gamma_k \to \mathrm{Aut}\,\big(Y(T)\big) \cong \mathbf{GL}_n(\mathbf{Z})$. If we identify T with $\mathbf{G}_m \times \cdots \times \mathbf{G}_m$ (over k_s) by choosing a basis of $Y(T)$, the ℓ^∞-division points of $T(k_s)$ form a group isomorphic to $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^n$ and the action of $g \in \Gamma_k$ on that group is by $\rho(g)\chi(g)$, where $\chi = \chi_{\ell^\infty}$.

Lemma 5. *Let A be a finite subgroup of $T(k)$. For every $g \in \Gamma_k$ we have*

$$v_\ell(A) \ \le \ v_\ell\big(\det(\rho(g)\,\chi(g)-1)\big) = v_\ell\big(\det(\rho(g^{-1})-\chi(g)\big).$$

Proof. By replacing A by its ℓ-Sylow subgroup, we may assume that A is an ℓ-group, hence is contained in the ℓ-division points of $T(k_s)$. Since the points of A are rational over k, they are fixed by g, i.e. they belong to the kernel of $g-1$. The inequality then follows from lemma 4, applied to $u = \rho(g)\,\chi(g) - 1$. □

We now choose $g \in \Gamma_k$ such that the inequality of lemma 5 gives that of th.4. Here is the choice:

$$\chi(g) = z_t u, \qquad \text{where } z_t \in \mathbf{Z}_\ell^* \text{ has order } t, \text{ and } v_\ell(1-u) = m.$$

(This works for $\ell = 2$ as well as for $\ell \neq 2$, thanks to the definition of t in §4.1 and §4.2. Note that in all cases but $\ell = 2$, type (c), $\chi(g)$ is a topological generator of $\operatorname{Im}\chi$.)

We have $\rho(g) \in \mathbf{GL}_n(\mathbf{Z})$, and $\rho(g)$ is of finite order (because the image of $\rho : \Gamma_k \to \mathbf{GL}_n(\mathbf{Z})$ is finite). Hence the characteristic polynomial F of $\rho(g^{-1})$ is a product of cyclotomic polynomials:

$$F = \prod \Phi_{d_j}, \qquad \text{with } \sum \varphi(d_j) = n. \tag{5.2.1}$$

The inequality of lemma 5 gives

$$v_\ell(A) \le \sum v_\ell\big(\Phi_{d_j}(z_t u)\big).$$

We thus need to compute $v_\ell\big(\Phi_d(z_t u)\big)$ for every $d \ge 1$. The result is:

Lemma 6. *We have*

$$v_\ell\big(\Phi_d(z_t u)\big) = \begin{cases} m & \text{if } d = t \\ 1 & \text{if } d = t\cdot\ell^\alpha,\ \alpha \ge 1 \quad \text{or} \quad \alpha = -1 \ (\text{if } t = 2 = \ell) \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (We restrict ourselves to the case $\ell \neq 2$. The case $\ell = 2$ is analogous but slightly different.)

We have $\Phi_d(z_t u) = \prod(z_t u - z)$ where z runs through the primitive d-th roots of unity in $\overline{\mathbf{Q}}_\ell$. Write d as $d = \delta\cdot\ell^\alpha$ with $(\delta,\ell) = 1$ and $\alpha \ge 0$. The images of the z's in the residue field $\overline{\mathbf{F}}_\ell$ of $\overline{\mathbf{Q}}_\ell$ are primitive δ-th roots of unity. If $\delta \neq t$, none of them is equal to the image of $z_t u$, which has order t. In that case, all the $z_t u - z$ are units in $\overline{\mathbf{Q}}_\ell$ hence have valuation 0 and we have $v_\ell\big(\Phi_d(z_t u)\big) = 0$. If $\delta = t$, i.e. $d = t\cdot\ell^\alpha$ with $\alpha \ge 0$, there are two cases:

(a) $\alpha = 0$, i.e. $d = t$. In that case, one of the z's is equal to z_t and we have $v_\ell(z_t u - z) = v_\ell(u-1) = m$; the other z's contribute 0.

(b) $\alpha \ge 1$. Here z can be written as $z'\cdot z''$ where z' runs through the t-th primitive roots of 1, and z'' through the ℓ^α-th primitive roots of 1. The valuation of $z - z_t u$ is 0 unless $z' = z_t$, in which case $v_\ell(z - z_t u) = v_\ell(z'' - u)$. It is well-known that $v_\ell(z''-1) = \frac{1}{(\ell-1)\ell^{\alpha-1}}$. Since $v_\ell(u-1) = m$, which is strictly larger, we have

$$v_\ell(z''-u) = v_\ell\big((z''-1)-(u-1)\big) = \frac{1}{(\ell-1)\ell^{\alpha-1}} = \frac{1}{\varphi(\ell^\alpha)}.$$

Since the number of the z'' is $\varphi(\ell^\alpha)$, we thus get $v_\ell\big(\Phi_d(z_t u)\big) = 1$, as claimed. □

We can now prove theorem 4: With the notation of (5.2.1), denote by r_1 the number of j's with $d_j = t$, and by r_2 the number of j's with $d_j = t\cdot\ell^{\alpha_j}$, $\alpha_j \geq 1$, or $\alpha_j = -1$ in case $\ell = 2, t = 2$. Using lemmas 5 and 6 we get

$$v_\ell(A) \leq r_1 m + r_2$$

and of course

$$r_1\varphi(t) + \sum \varphi(t\cdot\ell^{\alpha_j}) \leq n = \dim T.$$

Since $\varphi(t\cdot\ell^{\alpha_j}) \geq \varphi(t)(\ell-1)$ this shows that $r_1\varphi(t) + r_2\varphi(t)(\ell-1) \leq n$. Hence $r_1 + r_2(\ell-1) \leq [n/\varphi(t)]$, and we have:

$$v_\ell(A) \leq r_1 m + r_2 \;\leq\; r_1 m + r_2(\ell-1)m \;\leq\; m[n/\varphi(t)]\,,$$

which concludes the proof. □

Remark. Since $(\ell-1)m > 0$ in all cases (even if $\ell = 2$), the above proof shows that $v_\ell(A)$ can be equal to $m[n/\varphi(t)]$ only when $r_2 = 0$. In other words:

Complement to theorem 4. *Assume $v_\ell(A) = m[n/\varphi(k)]$, where $n = \dim T$. If $g \in \Gamma_k$ is such that $\chi(g) = z_t u$, with $v_\ell(u-1) = m$ as above, the characteristic polynomial of $\rho(g)$ is divisible by $(\Phi_t)^N$, with $N = [n/\varphi(k)]$.*
(In other words, the primitive t-th roots of unity are eigenvalues of $\rho(g)$ with multiplicity N.)

When $t = 1$ or 2 (i.e. when $\varphi(t) = 1$), this can be used to determine the structure of an "optimal" T:

Corollary. *Assume $t = 1$ or 2, and $v_\ell(A) = mn$. Then :*
(i) *If $t = 1$, the torus T is split* (i.e. isomorphic to the product of n copies of $\mathbf{G}_m$).
(ii) *If $t = 2$, T is isomorphic to the product of n non-split tori of dimension 1 which are split by the quadratic extension $k(z_\ell)/k$ if $\ell \neq 2$ and by $k(i)/k$ if $\ell = 2$.*

Proof. We give the proof for $t = 2$ and $\ell > 2$: the case $t = 1$ is easier and the case $t = 2 = \ell$ requires similar, but more detailed, arguments.

Let $\gamma \in \Gamma_k$. We may write $\chi(\gamma)$ as $e_\gamma \cdot u_\gamma$, with $e_\gamma \in \{1, -1\}$ and $u_\gamma \in \{1 + \ell^m\mathbf{Z}_\ell\}$. There are three cases:

(a) $e_\gamma = -1$ and $v_\ell(u_\gamma - 1) = m$
(b) $e_\gamma = -1$ and $v_\ell(u_\gamma - 1) > m$
(c) $e_\gamma = 1$.

In case (a), the "complement" above shows that $\rho(\gamma)$ has -1 for eigenvalue with multiplicity n, hence $\rho(\gamma) = -1$ in $\mathrm{Aut}(T) \simeq \mathbf{GL}_n(\mathbf{Z})$.

In case (b), choose $g \in \Gamma_k$ of type (a); this is possible by the very definition of t and m. The element $g^2\gamma$ is of type (a) (this uses the fact that ℓ is odd); hence we have $\rho(g^2\gamma) = -1$ and since $\rho(g) = -1$ this shows that $\rho(\gamma) = -1$.

If γ is of type (c), then $g\gamma$ is of type (a) or (b) and we have $\rho(g\gamma) = -1$ hence $\rho(\gamma) = 1$.

In all cases, we have $\rho(\gamma) \in \{1, -1\}$, and more precisely $\rho(\gamma) = e_\gamma$. The corollary follows. □

It would be interesting to have a similar classification for $t > 2$.

5.3. **Proof of theorem 4′: construction of tori with large A's.** To prove th.4′ it is enough to construct a k-torus T, of dimension $n = \varphi(t)$, such that $T(k)$ contains a cyclic subgroup of order ℓ^m. Here is the construction:

Let K be the field $k(z_\ell)$ if $\ell \neq 2$ and the field $k(i)$ if $\ell = 2$. It is a cyclic extension of k of degree t with Galois group C_t. Let $T_1 = R_{K/k}\mathbf{G}_m$ be the torus: "multiplicative group of K"; we have $T_1(k) = K^*$, and $T_1(k)$ contains the group $\langle z_{\ell^m} \rangle$, cf. §4. If σ is a generator of C_t, σ acts on T_1, and we have $\sigma^t - 1 = 0$ in the ring $\mathrm{End}(T_1)$. Let us write the polynomial $X^t - 1$ as $\Phi_t(X) \cdot \Psi(X)$, where Φ_t is the t-th cyclotomic polynomial. We have $\Phi_t(\sigma)\Psi(\sigma) = 0$ in $\mathrm{End}(T_1)$. Let $T = \mathrm{Im}\,\Psi(\sigma)$ be the image of

$$\Psi(\sigma) : T_1 \to T_1.$$

One checks that

(a) $\dim T = \varphi(t)$

(b) $T(k)$ contains z_{ℓ^m}.

(For $\ell \neq 2$, (b) follows from the fact that the restriction of $\Psi(\sigma)$ to $\langle z_{\ell^m} \rangle$ is an automorphism. For $\ell = 2$, use the fact that T is the kernel of $\Phi_t(\sigma)$.)

Hence T has the required properties. □

Alternate description of T. It is enough to describe its character group $T^* = \mathrm{Hom}_{k_s}(T, \mathbf{G}_m)$, together with the action of Γ_k on T^*:

- $T^* = \mathbf{Z}[X]/\Phi_t(X) =$ algebraic integers of the cyclotomic field $\mathbf{Q}(\mu_t)$
- Γ_k acts on T^* by $\Gamma_k \to \mathrm{Im}\,\chi_{\ell^\infty} \to C_t \xrightarrow{\sim} \mathrm{Aut}\big(\mathbf{Q}(\mu_t)\big)$.

(It does not matter which isomorphism of C_t onto $\mathrm{Aut}\big(\mathbf{Q}(\mu_t)\big)$ one chooses; they all give isomorphic tori.)

5.4. **The S-bound for reductive groups.** Recall, cf. §3.3, that "reductive" ⇒ "connected".

Theorem 5. *Let G be a reductive group over k, of rank r, with Weyl group W. If A is a finite subgroup of $G(k)$, one has*

$$v_\ell(A) \leq m \left[\frac{r}{\varphi(t)} \right] + v_\ell(W).$$

Proof. As usual, we may assume that A is an ℓ-group. In that case it is nilpotent, and by the corollary to th.3″ of §3.3 there exists a maximal k-torus T of G whose normalizer $N = N_G(T)$ contains A. Put $W_T = N/T$; this is a finite k-group such that $W_T(k_s) \simeq W$. If A_T denotes the intersection of A with $T(k)$, we have an exact sequence

$$1 \to A_T \to A \to W_T(k)\,.$$

Hence $v_\ell(A) \leq v_\ell(A_T) + v_\ell\big(W_T(k)\big)$. By th.4, we have $v_\ell(A_T) \leq m \cdot [r/\varphi(t)]$; on the other hand $W_T(k)$ is isomorphic to a subgroup of W, hence $v_\ell\big(W_T(k)\big)$ $\leq v_\ell(W)$. The theorem follows. □

Corollary. *If $r < \varphi(t)$, then $G(k)$ is ℓ-torsion free* (i.e. does not contain any elements of order ℓ).

Proof. We have $\left[\frac{r}{\varphi(t)}\right] = 0$. Hence by th.5 it is enough to show that $v_\ell(W) = 0$, but this follows from th.1 of §1.1 since W is isomorphic to a subgroup of $\mathbf{GL}_r(\mathbf{Z})$ and $r < \varphi(t) \leq t \leq \ell - 1$. □

Remark. The "S-bound" given by th.5 looks *a priori* rather coarse:

(a) The torus T is not an arbitrary torus of dimension r; the fact that it is a subtorus of G puts non-trivial conditions on it; for instance the action of Γ_k on $T^* = \mathrm{Hom}_{k_s}(T, \mathbf{G}_m)$ stabilizes the set of roots.

(b) The group $W_T(k)$ is in general smaller than W itself, and the image of $N(k) \to W_T(k)$ may be even smaller.

It is therefore surprising how often the S-bound is close to being optimal. As an example, take $k = \mathbf{Q}$ and G of type E_8. We have $m = 1$ and $t = \ell - 1$ (except when $\ell = 2$ in which case $m = t = 2$), $r = 8$, $|W| = 2^{14}3^5 5^2 7$. The S-bound tells us that, if A is a finite subgroup of $G(\mathbf{Q})$, its order divides the number

$$M_S = 2^{30}\cdot 3^{13}\cdot 5^6\cdot 7^5\cdot 13^2\cdot 17\cdot 19\cdot 31\,.$$

We shall see later (cf. §6.3.2 and §7) that the best bound is $M = M_S/5\cdot 7\cdot 17$. In particular, the ℓ-factors of M_S are optimal for all ℓ's except $\ell = 5, 7$ and 17.

§6. The M-bound

6.1. **Notation.** From now on, G is a semisimple [1] group over k. We denote by R its root system (over k_s), by W its Weyl group, and by r its rank. The group W has a natural linear representation of degree r. The invariants of W acting on $\mathbf{Q}[x_1, \ldots, x_r]$ make up a graded polynomial algebra of the form $\mathbf{Q}[P_1, \ldots, P_r]$, where the P_i are homogeneous of degrees d_i, with $d_1 \leq d_2 \leq \cdots \leq d_r$, (Shephard-Todd theorem, cf. e.g. [LIE V], §5, th.4 or [Se 00], p.95). The d_i's are called the *invariant degrees* of W (or of G). One has

$$\prod d_i = |W| \quad \text{and} \quad \sum (2d_i - 1) = \dim G\,.$$

When G is quasi-simple (i.e. when R is irreducible) d_r is equal to the Coxeter number $h = (\dim G)/r - 1$, and one has the symmetry formula

$$d_i + d_{r+1-i} = h + 2\,.$$

Moreover, if $j < h$ is prime to h, then $j+1$ is one of the d_i's. These properties make $d_1, \ldots, d_r$ very easy to compute (see e.g. the tables of [LIE VI]).

For instance, for G of type E_8, the d_i's are: 2, 8, 12, 14, 18, 20, 24, 30.

Let $\mathrm{Dyn}(R)$ be the Dynkin diagram of R. There is a natural action of the Galois group Γ_k on $\mathrm{Dyn}(R)$: this follows from the fact that $\mathrm{Dyn}(R)$ can be defined intrinsically from $G_{/k_s}$ (cf. [LIE VIII], §4, no 4, Scholie, or [SGA 3],

[1] We could also accept inner forms of reductive groups, for instance $\mathbf{GL}_n$ or more generally $\mathbf{GL}_D$, where D is a central simple k-algebra with $[D : k] = n^2$. In that case, one has $r = n$, the d_i's are the integers $1, 2, \ldots, n$ and th.6 is valid, with the same proof.

exposé XXIV, §3, p.344). In what follows (with the only exception of §6.7) we make the assumption that *the action of* Γ_k *on* $\mathrm{Dyn}(R)$ *is trivial*: one then says that G is *of inner type* (it can be obtained from a split group G_0 by a Galois twist coming from the adjoint group of G_0).

Examples of groups of inner type :
- $\mathbf{SL}_n$, or more generally, $\mathbf{SL}_D$, where D is a central simple algebra over k.
- Any group G whose root system has no non-trivial automorphism, e.g. any group of type $A_1, B_r, C_r, G_2, F_4, E_7, E_8$.

6.2. **Statement of the theorem.** We fix ℓ, k, and the root system R of G. Recall that $\mathrm{Im}\,\chi_{\ell^\infty}$ is a closed subgroup of $\mathbf{Z}_\ell^*$. Define:

$$M(\ell,k,R) = \inf_{x\,\in\,\mathrm{Im}\,\chi_{\ell^\infty}} \sum v_\ell(x^{d_i}-1) = \inf_{g\,\in\,\Gamma_k} \sum v_\ell(\chi_{\ell^\infty}(g)^{d_i}-1)\,.$$

This is either an integer ≥ 0 or ∞ (it is ∞ if and only if the invariants m, t of k defined in §4 are such that $m = \infty$ and t divides one of the d_i's, see prop.4 below).

Theorem 6. *Let A be a finite subgroup of $G(k)$. Then $v_\ell(A) \leq M(\ell,k,R)$.* (Recall that G is semisimple of inner type, cf. §6.1.)

This is what we call the "M-bound" for $v_\ell(A)$. It will be proved in §6.5 below by a method similar to Minkowski's. We shall see in Lect. III that it is "optimal" except possibly in the case $\ell = 2$, type (c) of §4.2.

For computations, it is useful to write $M(\ell,k,R)$ explicitly in terms of the invariants t and m of §4:

Proposition 4. (1) *If $\ell \neq 2$ or if $\ell = 2, t = 1$* (case (a)), *one has*

$$M(\ell,k,R) = \sum_{\substack{i \\ d_i \equiv\, 0 \pmod t}} \big(m + v_\ell(d_i)\big)$$

(2) *If $\ell = 2$ and $t = 2$* $\big($cases (b) and (c)$\big)$, *one has*

$$M(2,k,R) = r_1 + mr_0 + v_2(W)\,,$$

where r_0 (resp. *r_1*) *is the number of indices i such that d_i is even* (resp. *d_i is odd*).

Proof. Let us begin with the case $\ell \neq 2$. One shows first that, if $t|d$, one has $v_\ell(x^d-1) \geq m + v_\ell(d)$ for every $x \in \mathrm{Im}\,\chi_{\ell^\infty}$. (This is easy, since x can be written as zu with $z^t = 1$ and $v_\ell(u-1) \geq m$, hence $x^d - 1 = u^d - 1$.)

This already shows that $M(\ell,k,R) \geq \sum_{t|d_i} \big(m + v_\ell(d_i)\big)$. To prove the opposite inequality, one chooses $x \in \mathrm{Im}\,\chi_{\ell^\infty}$ of the form zu with z of order t and $v_\ell(u-1) = m$. One gets (1).

The same argument works if $\ell = 2$ and $t = 1$. If $\ell = 2$ and $t = 2$, one has

$$\begin{aligned} v_2(x^d-1) &\geq m + v_2(d) \quad \text{if } d \text{ is even} \\ v_2(x^d-1) &\geq 1 \quad \text{if } d \text{ is odd}\,, \end{aligned}$$

for every $x \in \mathrm{Im}\,\chi_{2^\infty}$. This gives:

$$M(2,k,R) \geq \sum_{d_i \text{ odd}} 1 + \sum_{d_i \text{ even}} \big(m + v_2(d_i)\big) = r_1 + mr_0 + v_2(W)\,.$$

To get the opposite inequality, observe that $x = -1 + 2^m$ belongs to $\operatorname{Im} \chi_{2^\infty}$ and check that $\sum v_2(x^{d_i} - 1)$ is equal to $r_1 + mr_0 + v_2(W)$. □

Corollary. *Let $a(t)$ be the number of indices i such that $d_i \equiv 0 \pmod t$. If $a(t) = 0$, then $G(k)$ is ℓ-torsion free.*

Indeed, if $a(t) = 0$, the sum occurring in prop.4 is an empty sum, hence $M(\ell, k, R) = 0$ and one applies th.6. □

6.3. **Two examples: A_1 and E_8.** We take $k = \mathbf{Q}$, so that $t = \ell - 1$ and $m = 1$ if $\ell > 2$ and $t = m = 2$ if $\ell = 2$.

6.3.1. *Type A_1.* There is only one d_i, namely $d_1 = 2$, and prop.4 gives:

$$M\big(\ell, \mathbf{Q}, A_1\big) = \begin{cases} 3 & \text{if} \quad \ell = 2 \\ 1 & \text{if} \quad \ell = 3 \\ 0 & \text{if} \quad \ell > 3\,. \end{cases}$$

In other words, every finite subgroup of $G(\mathbf{Q})$ has an order which divides $2^3 \cdot 3$. This bound is optimal in the following sense:

(a) The split adjoint group $\mathbf{PGL}_2(\mathbf{Q})$ contains both a subgroup of order 3 and a dihedral subgroup of order 8 (but no subgroup of order 24).

(b) The simply connected group $\mathbf{SL}_{\mathbf{H}}(\mathbf{Q})$, where $\mathbf{H}$ is the standard quaternion division algebra, contains a subgroup of order 24 which is isomorphic to $\mathbf{SL}_2(\mathbf{F}_3)$. However the split group $\mathbf{SL}_2(\mathbf{Q})$ does not contain any subgroup of order 8 (but it does contain cyclic subgroups of order 3 and 4).

6.3.2. *Type E_8.* If we define $M(\mathbf{Q}, E_8)$ as $\prod_\ell \ell^{M(\ell, \mathbf{Q}, E_8)}$, prop.4 gives:

$$M(\mathbf{Q}, E_8) = 2^{30} \cdot 3^{13} \cdot 5^5 \cdot 7^4 \cdot 11^2 \cdot 13^2 \cdot 19 \cdot 31, \text{ see e.g. [Se 79], §3.3.}$$

By th.6, the order of every finite subgroup of $G(\mathbf{Q})$ divides $M(\mathbf{Q}, E_8)$. As we shall see in the next lecture, this multiplicative bound is optimal.

6.4. **A Chebotarev-style result.** We need such a result in order to generalize Minkowski's method of §1.

Let L be a normal domain which is finitely generated over $\mathbf{Z}$ as a ring, and let k be its field of fractions. If $d = \dim(L)$ denotes the Krull dimension of L ([AC VIII], §1), one has (*loc.cit.*, §2):

$$\begin{array}{lcllcl} d & = & 1 + \operatorname{tr.deg}(k/\mathbf{Q}) & \text{if} & \operatorname{char}(k) & = & 0 \\ d & = & \operatorname{tr.deg}(k/\mathbf{F}_p) & \text{if} & \operatorname{char}(k) & = & p > 0\,. \end{array}$$

Let $\operatorname{Specmax}(L)$ be the set of the maximal ideals of L (= set of closed points of $\operatorname{Spec}(L)$). If $x \in \operatorname{Specmax}(L)$, the residue field $\kappa(x) = L/x$ is finite (see e.g. [AC V], p. 68, cor. 1). We put $Nx = |\kappa(x)|$; it is the *norm* of x.

When $d = 0$, L is a finite field, and $\operatorname{Specmax}(L)$ has only one element. If $d > 0$ (e.g. when $\operatorname{char}(k) = 0$), then $\operatorname{Specmax}(L)$ is infinite. More precisely, the Dirichlet series $z(s) = \sum_x 1/(Nx)^s$ converges for $\operatorname{Re}(s) > d$, and one has

$$z(s) \sim \log\big(1/(s-d)\big) \quad \text{when} \quad s \to d \quad (\text{with } s > d)\,. \tag{6.4.1}$$

See [Se 65], §2.7, which only contains a sketch of proof; complete details (for a slightly weaker statement) can be found in [Pi 97], App. B [2]; see also [FW 84], pp.206-207.

Let now n be an integer ≥ 1 which is invertible in L (and hence in k). Let $\chi_n : \Gamma_k \to (\mathbf{Z}/n\mathbf{Z})^*$ denote the n-th cyclotomic character of k, cf. §4.0. As in §4, we shall be interested in $\operatorname{Im}\chi_n \subset (\mathbf{Z}/n\mathbf{Z})^*$.

Theorem 7. *Let c be an element of $(\mathbf{Z}/n\mathbf{Z})^*$, and let X_c be the set of all $x \in \operatorname{Specmax}(L)$ such that $Nx \equiv c \pmod{n}$. Then :*

a) *If $c \notin \operatorname{Im}\chi_n$, then $X_c = \varnothing$.*

b) *If $c \in \operatorname{Im}\chi_n$ and $d > 0$, then X_c is Zariski-dense in* $\operatorname{Specmax}(L)$ (or in $\operatorname{Spec}(L)$, this amounts to the same). *In particular, X_c is infinite.*

A more concrete formulation of b) is that, for every non-zero $f \in L$, there exists an x with $f \notin x$ and $Nx \equiv c \pmod{n}$.

Example. Take $L = \mathbf{Z}[1/n]$. Then $\operatorname{Specmax}(L)$ is the set of all prime numbers which do not divide n, and th.7 translates into Dirichlet's theorem on the existence of primes in arithmetic progressions.

Proof of theorem 7. The group $C = \operatorname{Im}\chi_n$ is the Galois group of the cyclotomic extension $k(z_n)/k$. Let L_n be the integral closure of L in $k(z_n)$. One checks by standard arguments that the ring extension L_n/L is finite and étale. In geometric terms, $\operatorname{Spec}(L_n) \to \operatorname{Spec}(L)$ is a finite étale covering. The group C acts freely on $\operatorname{Spec}(L_n)$, with quotient $\operatorname{Spec}(L)$. For every closed point x of $\operatorname{Spec}(L)$, the Frobenius element σ_x of x is a well-defined conjugacy class of C (hence an element of C since C is commutative). Moreover, if we view C as a subgroup of $(\mathbf{Z}/n\mathbf{Z})^*$, σ_x is the image of Nx in $\mathbf{Z}/n\mathbf{Z}$. This proves a).

Suppose now that $d > 0$ and that c belongs to $C = \operatorname{Im}\chi_n$. Let $z_c(s)$ be the Dirichlet series $\sum 1/(Nx)^s$, where the sum is over the elements x of X_c. The general Chebotarev density theorem ([Se 65], [Pi 97]) gives:

$$z_c(s) \sim \frac{1}{|C|}\log(1/(s-d)) \quad \text{when } s \to d \quad \text{with} \quad s > d\,. \tag{6.4.2}$$

In particular, we have $z_c(d) = +\infty$. If the Zariski closure $\overline{X}_c$ of X_c were of dimension $< d-1$, we would have $z_c(d) < \infty$, as one sees by splitting $\overline{X}_c$ into irreducible components, and applying (6.4.1). Hence b). □

6.5. **Proof of theorem 6.** Let $A \subset G(k)$ be as in th.6. We want to prove that

$$v_\ell(A) \leq M(\ell, k, R)\,.$$

We do it in three steps:

[2] When $\operatorname{char}(k) = 0$ one can give a stronger statement, in the spirit of the Prime Number Theorem:

For every $X \geq 2$, call $\pi_L(X)$ the number of $x \in \operatorname{Specmax}(L)$ such that $Nx \leq X$. Then:

$$\pi_L(X) = (1/d)\,X^d/\log X + O(X^d/\log^2 X) \quad \text{when} \quad X \to \infty\,.$$

The general Chebotarev density theorem can also be stated (and proved) in terms of such "natural" density (standard method: use Weil-Deligne estimates to reduce everything to the known case $d = 1$).

6.5.1. **The case where k is finite.** Put $q = |k|$. It is well-known that

$$|G(k)| = q^N \prod (q^{d_i} - 1), \qquad \text{where} \quad N = |R|/2 = \sum (d_i - 1).$$

This shows that $v_\ell(A) \leq \sum v_\ell(q^{d_i} - 1)$. The integer q, viewed as an element of $\mathbf{Z}_\ell^*$, is a topological generator of $\operatorname{Im} \chi_{\ell^\infty}$. Hence every element u of $\operatorname{Im} \chi_{\ell^\infty}$ is an ℓ-adic limit of powers of q and this implies that $v_\ell(u^d - 1) \geq v_\ell(q^d - 1)$ for every $d \geq 1$. Hence the lower bound which defines $M(\ell, k, R)$ is equal to $\sum v_\ell(q^{d_i} - 1)$; this proves th.6 in the case where k is finite.

6.5.2. **The case where k is finitely generated over its prime subfield.** By 6.5.1, we may assume that k is infinite. We need a subring L of k, with field of fractions k, which has the following properties:

(a) L is normal, finitely generated over $\mathbf{Z}$ and contains $1/\ell$.

(b) G comes by base change from a semisimple group scheme $\underline{G}$ over L, in the sense of [SGA 3], XIX. 2.7.

(c) A is contained in the group $\underline{G}(L)$ of the L-points of $\underline{G}$.

Lemma 7. *There exists such an L.*

This is standard, see e.g. [EGA IV], §8.1 □

Let us now choose $(L, \underline{G})$ with properties (a), (b) and (c). For every $x \in \operatorname{Specmax}(L)$, the fiber $\underline{G}_x$ of $\underline{G}$ at x is a semisimple group over $\kappa(x)$, of type R. Moreover, the Dynkin diagram of $\underline{G}$ is finite étale over $\operatorname{Spec}(L)$, cf. [SGA 3], XXIV.3.2; since it is "constant" for the generic fiber (i.e. over k) it is constant everywhere; this shows that the $\underline{G}_x$ are of inner type. The inclusion map $i : A \to \underline{G}(L)$ gives for every x a homomorphism $i_x : A \to \underline{G}\big(\kappa(x)\big)$. Since i is injective, there is an open dense subset X_0 of $\operatorname{Specmax}(L)$ such that i_x is injective for all $x \in X_0$. We thus get:

$$v_\ell(A) \;\leq\; v_\ell\big(\underline{G}(\kappa(x))\big) = \sum v_\ell\big((Nx)^{d_i} - 1\big) \quad \text{for all} \quad x \in X_0,$$

cf. 6.5.1. Let u be any element of $\operatorname{Im} \chi_{\ell^\infty}$. By applying th.7 to the image of u in $(\mathbf{Z}/\ell^j\mathbf{Z})^*$ with $j = 1, 2, \ldots$, we find a sequence of points x_j of X_0 such that $\lim Nx_j = u$ in $\mathbf{Z}_\ell^*$. We have:

$$v_\ell(u^{d_i} - 1) = \lim_{j \to \infty} \sum v_\ell\big((Nx_j)^{d_i} - 1\big)\,,$$

and applying the formula above to each of the x_j's we obtain

$$v_\ell(A) \leq \sum v_\ell(u^{d_i} - 1) \qquad \text{for every} \quad u \in \operatorname{Im} \chi_{\ell^\infty}\,.$$

This proves th.6 in the case 6.5.2.

[Variant: One reduces the general case to the case where $\dim(L) = 1$ by using Hilbert's irreducibility theorem, as explained in [Se 81], p.2; in the case $\dim(L) = 1$, one can apply the standard Chebotarev theorem instead of the general one.]

6.5.3. **The general case.** The same argument as for lemma 7 shows that G comes by base change from a semisimple group G' over a subfield k' of k which is finitely generated over the prime subfield of k (i.e. $\mathbf{F}_p$ or $\mathbf{Q}$). Moreover, one may assume (after enlarging k' if necessary) that A is contained in $G'(k')$. The Galois group $\Gamma_{k'}$ acts on the Dynkin diagram $\operatorname{Dyn}(R)$ of G' (which is the same as the one of G). Let k'' be the Galois

extension of k' corresponding to the kernel of $\Gamma_{k'} \to \mathrm{Aut}\ \mathrm{Dyn}(R)$. Since G is of inner type over k, the field k'' is contained in k. By base change to k'', G' gives a semisimple group G'' which is of inner type and we may apply 6.5.2 to (G'', A). We get $v_\ell(A) \leq M(\ell, k'', R)$. Since k'' is contained in k, we have $M(\ell, k'', R) \leq M(\ell, k, R)$: the group $\mathrm{Im}\,\chi_{\ell^\infty}$ can only decrease by field extensions. Hence $v_\ell(A) \leq M(\ell, k, R)$. □

6.5.4. **Remark.** Surprisingly, the proof above does not really use the hypothesis that A is a subgroup of $G(k)$. It uses only that A *acts freely on* G, viewed merely as a k-variety (and not as a homogeneous space); this is indeed enough to ensure that $v_\ell(A) \leq v_\ell(G(k))$ when k is finite. Here is an example: take $G = \mathbf{SL}_2$, $\ell = 2$, $k = \mathbf{Q}$; the M-bound is 3, which means that any finite 2-subgroup of $\mathbf{SL}_2(\mathbf{Q})$ has order ≤ 8. As was said in §6.3.1, there is in fact no subgroup of order 8 in $\mathbf{SL}_2(\mathbf{Q})$. But one can make a cyclic group of order 8 act freely on the variety $\mathbf{SL}_2$: take for instance the group generated by the automorphism:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} d-c & -c-d \\ \frac{(a-b)}{2} & \frac{(a+b)}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Hence, even in this bad-looking case, the M-bound can claim to be "optimal".

6.6. **An analogue of Sylow's theorem.**

Theorem 8. *Let A and A' be two finite ℓ-subgroups of $G(k)$. Assume that $v_\ell(A)$ is equal to the* M*-bound $M(\ell, k, R)$. Then there exists $y \in G(\bar{k})$ such that $yA'y^{-1} \subset A$.*

Corollary. *If both A and A' attain the* M*-bound, then they are geometrically conjugate* (i.e. conjugate in $G(\bar{k})$). *In particular, they are isomorphic.*

Proof. We may assume that k is finitely generated over its prime subfield. If it is finite, th.8 is just a special case of Sylow's theorem. Let us assume that k is infinite, and choose $L, \underline{G}$ as in §6.5.2 with $A, A' \subset \underline{G}(L)$. Let Y be the subscheme of $\underline{G}$ made up of the points y with $yA'y^{-1} \subset A$. Let X be the set of all $x \in \mathrm{Specmax}(L)$ such that Nx, viewed as an element of $\mathbf{Z}_\ell^*$, is of the form $z_t u$ with z_t of order t and $v_\ell(u-1) = m$ (note that m is finite, cf. §4.3). It follows from th.7, applied to $n = \ell^{m+1}$, that X is Zariski-dense in $\mathrm{Spec}(L)$. If $x \in \mathrm{Specmax}(L)$, the groups A and A' inject into $\underline{G}(\kappa(x))$ (this is an easy consequence of the hypothesis that ℓ is invertible in L). If moreover x belongs to X, then the same computation as in §5.2 shows that $v_\ell\big(\underline{G}(\kappa(x))\big)$ is equal to the M-bound, hence A is an ℓ-Sylow of $\underline{G}(\kappa(x))$. By Sylow's theorem, this shows that A' is conjugate in $\underline{G}(\kappa(x))$ to a subgroup of A. In particular, the fiber at x of $Y \to \mathrm{Spec}(L)$ is non-empty. Since X is Zariski-dense, this implies that the generic fiber $Y_{/k}$ of $Y \to \mathrm{Spec}(L)$ is non-empty, i.e. that $Y(\bar{k})$ is non-empty. □

Remark. One can show that Y is smooth over L, and hence that $Y(k_s) \neq \varnothing$ which is slightly more precise than $Y(\bar{k}) \neq \varnothing$.

Exercise. Show that a family of polynomial equations with coefficients in $\mathbf{Z}$ has a solution in $\mathbf{C}$ if and only if it has a solution in $\mathbf{Z}/p\mathbf{Z}$ for infinitely many p's.

6.7. **Arbitrary semisimple algebraic groups.** In the previous sections, we have assumed that G is of inner type, i.e. that the natural homomorphism

$$\varepsilon : \Gamma_k \to \operatorname{Aut} \operatorname{Dyn}(R)$$

is trivial. Let us now look briefly at the general case, where no hypotheses on ε are made. In order to state the result which replaces th.6 we need to introduce the linear representations ε_d of Γ_k defined as follows:

Let $S = \mathbf{Q}[P_1, \ldots, P_r]$ be the $\mathbf{Q}$-algebra of W-invariant polynomials, cf. §6.1. Let $I = (P_1, \ldots, P_r)$ be the augmentation ideal of S; put $V = I/I^2$. The vector space V is of dimension r, and is graded; the dimension of its d-th component V_d is equal to the number of indices i with $d_i = d$. The group $\operatorname{Aut} \operatorname{Dyn}(R)$ acts on S, V and the V_d's; by composing this action with ε, we get for each $d > 0$ a linear representation

$$\varepsilon_d : \Gamma_k \to \operatorname{Aut}(V_d)\,.$$

Theorem 6′. *Let A be a finite subgroup of $G(k)$. Then*:

$$v_\ell(A) \leq \inf_{g \in \Gamma_k} \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(g)^d - \varepsilon_d(g))\big)$$

(The determinant is relative to the vector space $V_d \otimes \mathbf{Q}_\ell$.)

Proof (sketch). The method is the same as the one used for th.6. There are three steps:

(1) Reduction to the case where k is finitely generated over its prime subfield; this is easy.

(2) Reduction to the case where k is finite, via the general Chebotarev density theorem instead of th.7.

(3) The case where k is finite. In that case, if $q = |k|$, and if σ is the Frobenius generator of Γ_k, one has (cf. e.g. [St 68] th. 11.16)

$$v_\ell\big(G(k)\big) = \sum_d v_\ell\big(\det(q^d - \varepsilon_d(\sigma))\big) = \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(\sigma)^d - \varepsilon_d(\sigma))\big)$$

hence the desired formula:

$$(*) \qquad v_\ell(A) \leq \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(g)^d - \varepsilon_d(g))\big)$$

in the special case $g = \sigma$. By applying this to the finite extensions of k, one sees that the inequality $(*)$ is valid for all $\sigma^n, n = 1, 2, \ldots$, and hence for all $g \in \Gamma_k$, since the σ^n are dense in Γ_k. □

Remark. One may also prove th.6′ using ℓ-adic cohomology, cf. §6.8.

Example. Take R of type A_2, so that $\operatorname{Aut} \operatorname{Dyn}(R) = \{1, -1\}$ and ε may be viewed as a quadratic character of Γ_k. The V_d's are of dimension 1 for $d = 2, 3$ and are 0 otherwise. The action of $\operatorname{Aut} \operatorname{Dyn}(R)$ on V_d is trivial for all d, except $d = 3$. Hence $\varepsilon_2 = 1$, $\varepsilon_3 = \varepsilon$, and th.6′ can be rewritten as:

$$v_\ell(A) \leq \inf_{g \in \Gamma_k} \big\{ v_\ell(\chi_{\ell^\infty}(g)^2 - 1) + v_\ell\big(\chi_{\ell^\infty}(g)^3 - \varepsilon(g)\big) \big\}\,.$$

A similar result holds for the types A_r ($r > 2$), D_r (r odd) and E_6, with 2 (resp. 3) replaced by the even d_i's (resp. the odd d_i's).

6.8. **The cohomological method.** Let us consider first the general situation suggested in §6.5.4 where a finite group A acts freely on a quasi-projective k-variety X. As explained in [Il 05], §7, one can then give an upper bound for $v_\ell(A)$ in terms of the action of Γ_k on the étale cohomology of X. More precisely, let $H^i_c(X)$ denote the i-th étale cohomology group of $X_{/k_s}$, with proper support and coefficients $\mathbf{Q}_\ell$; it is a finite dimensional $\mathbf{Q}_\ell$-vector space which is 0 for $i > 2\cdot\dim(X)$. There is a natural action of Γ_k on $H^i_c(X)$, and, for each $g \in \Gamma_k$, one can define the "Lefschetz number" $\Lambda_X(g)$ by the usual formula:

$$\Lambda_X(g) = \sum_i (-1)^i \mathrm{Tr}\big(g|H^i_c(X)\big)\,.$$

One has $\Lambda_X(g) \in \mathbf{Z}_\ell$. Moreover:

Theorem 6″. $v_\ell(A) \leq \inf_{g\in\Gamma_k} v_\ell\big(\Lambda_X(g)\big)$.

Proof. See [Il 05], §7, especially cor.7.5. The proof follows the same pattern as the other proofs of the present §: one uses Chebotarev to reduce to the case where k is finite, in which case the result follows from the fact, due to Grothendieck, that, if σ is the (geometric) Frobenius generator of Γ_k, then $\Lambda_X(\sigma)$ is equal to $|X(k)|$, hence is divisible by $|A|$ since the action of A is free. (As in the proof of th.6′, one applies this, not only to σ but also to its powers σ^n, $n > 0$, and one uses the fact that the σ^n are dense in Γ_k.) □

If one applies th.6″ to $A \subset G(k)$, with A acting by left translations on $X = G$, one recovers th.6 and th.6′, thanks to the known structure of the cohomology of G, cf. e.g. [SGA $4\frac{1}{2}$], p. 230.

6.9. **The Cremona group: open problems.** Recall that the *Cremona group* $\mathbf{Cr}_r(k)$ is the group of k-automorphisms of the field $k(X_1, \ldots, X_r)$, i.e. the group of birational automorphisms (or "pseudo-automorphisms", cf. [De 70]) of the projective r-space over k. For $r = 1$, one has $\mathbf{Cr}_1(k) = \mathbf{PGL}_2(k)$. Let us assume that $r \geq 2$. As explained in [De 70], $\mathbf{Cr}_r$ is not an algebraic group, but looks like a kind of very large semisimple group of rank r (very large indeed: its "Weyl group" is the infinite group $\mathbf{GL}_r(\mathbf{Z})$). Not much is known about the finite subgroups of $\mathbf{Cr}_r(k)$ beyond the classical case $r = 2$ and k algebraically closed. Here is a question suggested by §5.1:

- Is it true that $\mathbf{Cr}_r(k)$ has no ℓ-torsion if $\varphi(t) > r$?

In the special case $k = \mathbf{Q}$, $r = 2$ or 3, this amounts to:

- Is it true that the fields $\mathbf{Q}(X_1, X_2)$ and $\mathbf{Q}(X_1, X_2, X_3)$ have no automorphism of prime order ≥ 11? (Automorphisms of order 2, 3, 5 and 7 do exist.)

It would be very interesting to attack these questions using cohomology, but I do not see how to do this. It is not even clear how to define cohomological invariants of $\mathbf{Cr}_r(\mathbf{C})$, since there is no natural topology on that group. Still, one would like to give a meaning to a sentence such as

"$\mathbf{Cr}_r(\mathbf{C})$ is connected for $r \geq 1$ and simply-connected for $r \geq 2$."

III. Construction of large subgroups

§7. Statements

We keep the notation of Lecture II: k, ℓ, χ_{ℓ^∞}, t, m, ... We consider only semisimple groups over k with a root system R which is *irreducible.* The M-bound of §6.2 will be denoted by $M(\ell, k, R)$; it only depends on the pair (ℓ, k) via the invariants t and m, and on R via the degrees $d_1, \ldots, d_r$ of W. We limit ourselves to the case $m < \infty$; see §14 for the case $m = \infty$.

A pair (G, A), where G is of inner type with root system R, and $A \subset G(k)$ is a finite group, will be called *optimal* if $v_\ell(A)$ is equal to the M-bound $M(\ell, k, R)$. (We could assume that A is an ℓ-group, but this would not be convenient for the constructions which follow.) Our goal is to prove:

Theorem 9. *If $\ell \neq 2$, an optimal pair (G, A) exists (for any k, R).*

Theorem 10. *If $\ell = 2$, an optimal pair (G, A) exists if $\operatorname{Im} \chi_{2^\infty}$ is of type* (a) *or* (b) *in the sense of* §4.2 (i.e. if $\operatorname{Im} \chi_{2^\infty}$ can be topologically generated by one element).

Theorem 11. *In the case $\ell = 2$ and type* (c)*, there exists (G, A) with*

$$v_2(A) = r_0 m + v_2(W)$$

where r_0 is the number of indices i such that d_i is even.

Note that here the M-bound is $M(2, k, R) = r_1 + r_0 m + v_2(W)$ with $r_1 = r - r_0$, cf. §6.2, prop.4. Hence $v_2(A)$ differs from $M(2, k, R)$ only by r_1. In particular, A is optimal if $r_1 = 0$. Hence:

Corollary. *If all the d_i's are even* (i.e. if $-1 \in W$)*, then an optimal pair (G, A) exists for $\ell = 2$* (and hence for all ℓ's, thanks to th.9).

This applies in particular to the exceptional types G_2, F_4, E_7 and E_8.

Remarks. (1) The simplest case where the M-bound is not attained is $k = \mathbf{Q}$, $\ell = 2$, R of type A_2, where $m = 2$, $r_0 = 1$, $r = 2$, the M-bound is 4, and it follows from [Sch 05] that $v_2(A) \leq 3$ for every finite subgroup A of $G(\mathbf{Q})$.

(2) In Theorems 9, 10 and 11, no claim is made on the structure of G except that it is of inner type and that its root system is of type R. However, if one looks closely at the proofs given in the next sections, one sees that G can be chosen to have the following properties:
- it is simply connected;
- it splits over the cyclotomic field $k(z_\ell)$ if $\ell > 2$, and over $k(i)$ if $\ell = 2$.

Simple examples (such as $k = \mathbf{Q}$, $\ell = 3$, G of type G_2) show that it is not always possible to have G split over k.

(3) If G is not chosen carefully, the group $G(k)$ may not contain any large ℓ-subgroup, even if k contains all the roots of unity. For instance, when R is of type A_1 (resp. of type E_8) it is easy (resp. it is possible) to construct a pair (G, k) such that the only torsion elements of $G(k)$ have order 1 or 2 (resp. $G(k)$ is torsion free).

(4) The three theorems above are almost obvious if the characteristic is $p \neq 0$ (especially Theorem 11 since type (c) never occurs!): one takes a

finite field k_0 contained in k which has the same invariants t and m (this is easily seen to be possible – if k is finitely generated over $\mathbf{F}_p$, one chooses the maximal finite subfield of k), and one takes for G the group deduced by base change from a split group G_0 over k_0 with root system R. If we choose for A the finite group $G_0(k_0)$, it is clear from the way we got the M-bound that $v_\ell(A) = M(\ell, k_0, R) = M(\ell, k, R)$, so that (G, A) is optimal.

In what follows, we shall assume that $\mathrm{char}(k) = 0$. Note also that we could replace k by any subfield having the same invariants t and m, for instance the intersection of k with the field of ℓ^∞-roots of unity. We could thus assume that *k is a cyclotomic number field*, if needed.

The proof of Theorem 9 will be given first for classical groups (§9), by explicit elementary constructions similar to those of Schur. The more interesting case of exceptional groups (§12) will use different methods, based on Galois twists (§10), Tits groups and braid groups (§11). The case of $\ell = 2$ will be given in §13. The last section (§14) is about $m = \infty$.

§8. Arithmetic methods ($k = \mathbf{Q}$)

These methods are not strong enough to prove the statements of §7, but they give very interesting special cases.

8.10. **Euler characteristics.** Here, the ground field is $\mathbf{Q}$. One starts from a split simply connected group scheme G over $\mathbf{Z}$ (this makes sense, cf. [SGA 3]). One may thus speak of the group $\Gamma = G(\mathbf{Z})$ of the *integral points* of G. It is a discrete subgroup of $G(\mathbf{R})$. Its Euler characteristic $\chi(\Gamma)$ ("caractéristique d'Euler-Poincaré" in French) is well-defined (see [Se 71] and [Se 79]); it is a rational number. Moreover it is proved in [Ha 71] that

$$\chi(\Gamma) = c \prod_{i=1}^{r} \frac{1}{2}\zeta(1 - d_i) = c \prod_{i=1}^{r} \frac{b_{d_i}}{2d_i}, \tag{8.10.1}$$

where b_d is the d-th Bernoulli number, ζ is the zeta function and $c = |W|/|W_K|$ where W_K is the Weyl group of a maximal compact subgroup K of $G(\mathbf{R})$. Assume that all d_i's are *even* (if not, all the terms in (8.10.1) are zero). Using standard properties of Bernoulli numbers, one can check that *the* M*-bound relative to ℓ is* $M = \sum_i v_\ell\big(\mathrm{den}\,\big(\frac{1}{2}\zeta(1-d_i)\big)\big)$, where "den" means denominator. Hence, if ℓ does not divide c, and does not divide the numerator of any $\frac{1}{2}\zeta(1 - d_i)$ (which is the case if ℓ is a so-called regular prime), one sees that *the denominator of* $\mathrm{EP}(\Gamma)$ *is divisible by* ℓ^M. But a theorem of K. Brown [Br 74] shows that this is only possible if Γ contains a finite subgroup of order ℓ^M. Hence we get an optimal pair (provided $(c, \ell) = 1$, and ℓ is regular, say).

Example. Take G of type E_8; here $c = 3^3 \cdot 5$, and the numerators of the $\frac{1}{2}\zeta(1 - d_i)$ do not cancel any denominator. Hence one obtains that a split E_8 contains an optimal A for all $\ell \neq 3, 5$, with the extra information that A can be found inside the group $\Gamma = G(\mathbf{Z})$ – but no information on what it looks like!

8.11. **Mass formulae.** In [Gr 96], B. Gross considers **Q**-forms of G such that $G(\mathbf{R})$ is *compact*; he also requires another condition which guarantees that G has a *smooth model over* $\mathbf{Z}$. This condition is fulfilled for types B, D, G_2, F_4 and E_8. He then proves a *mass formula* à la Minkowski ([Gr 96], prop.2.2):

$$\sum \frac{1}{|A_\sigma|} = \prod_{i=1}^{r} \frac{1}{2}\zeta(1-d_i)$$

where the A_σ are the $\mathbf{Z}$-points of the smooth models of G over $\mathbf{Z}$ (taken up to conjugation). Each A_σ is finite. It is then clear that, if ℓ^N is the ℓ-th part of the denominator of $\prod_{i=1}^r \frac{1}{2}\zeta(1-d_i)$, the ℓ-Sylow subgroup of one of the A_σ has order $\geq \ell^N$. If N is equal to the Minkowski bound M (which happens if ℓ does not divide the numerator of any of the $\frac{1}{2}\zeta(1-d_i)$), then such a Sylow has order ℓ^M, and we get an optimal pair. Note that there is no extra factor "c" as in (8.10.1). This works very well for G_2, F_4, E_8 (and some classical groups too, cf. [Gr 96]):

G_2 - Here the mass is $\frac{1}{4}\zeta(-1)\zeta(-5) = \dfrac{1}{2^6 3^3 7}$, and it is obtained with just one A_σ, which turns out to be isomorphic to $G_2(\mathbf{F}_2)$.

F_4 - There are two A_σ's and the mass formula is

$$\begin{aligned} \frac{1}{2^{15}\cdot 3^6\cdot 5^2\cdot 7} + \frac{1}{2^{12}\cdot 3^5\cdot 7^2\cdot 13} &= \tfrac{1}{16}\zeta(-1)\zeta(-5)\zeta(-7)\zeta(-11) \\ &= \frac{691}{2^{15}\cdot 3^6\cdot 5^2\cdot 7^2\cdot 13}. \end{aligned}$$

E_8 - Here the numerator is very large, but the denominator is exactly what is needed for the M-bound, namely:

$$2^{30}\cdot 3^{13}\cdot 5^5\cdot 7^4\cdot 11^2\cdot 13^2\cdot 19\cdot 31.$$

§9. Proof of theorem 9 for classical groups

Here $\ell \neq 2$. Recall that $\operatorname{Im}\chi_{\ell^\infty} = C_t \times \{1+\ell^m\mathbf{Z}_\ell\}$, where $m \geq 1$ and t divides $\ell - 1$. The M-bound is

$$M = \sum_{\substack{i \\ d_i \equiv 0 (\mathrm{mod}\, t)}} \big(m + v_\ell(d_i)\big).$$

We denote by K the field $k(z_\ell)$ generated by a root of unity of order ℓ. It is a cyclic extension of k, of degree t, with Galois group C_t. It contains z_{ℓ^m} but not $z_{\ell^{m+1}}$, cf. §4.1.

9.1. **The groups A_N and A_N^1.** If N is an integer ≥ 1, we denote by A_N the subgroup of $\mathbf{GL}_N(K)$ (where $K = k(z_\ell)$ as above) generated by the symmetric group S_N and the diagonal matrices whose entries are ℓ^m-th roots of unity (wreath product of S_N with a cyclic group of order ℓ^m). We have

$$v_\ell(A_N) = mN + v_\ell(N!). \tag{9.1.1}$$

The image of $\det_K : A_N \to K^*$ is $\{\pm 1\} \times \langle z_{\ell^m} \rangle$. Hence the kernel A_N^1 is such that

$$v_\ell(A_N^1) = m(N-1) + v_\ell(N!). \tag{9.1.2}$$

We are going to use A_N, and sometimes A_N^1, in order to construct optimal subgroups for the classical groups $\mathbf{SL}_n$, $\mathbf{SO}_n$ and $\mathbf{Sp}_n$; this is what Schur did in [Sch 05], §6, for the case of $\mathbf{GL}_n$.

9.2. **The case of $\mathbf{SL}_n$.** The d_i's are 2, 3,..., n. If we put $N = \left[\frac{n}{t}\right]$, we have

$$M = mN + v_\ell(N!) \qquad \text{if } t \geq 2, \tag{9.2.1}$$

$$M = m(N-1) + v_\ell(N!) \qquad \text{if } t = 1, \text{ in which case } N = n. \tag{9.2.2}$$

In the case $t \geq 2$, we take $A_N \subset \mathbf{GL}_N(K) \subset \mathbf{GL}_{Nt}(k)$, and observe that $\det_k(A_N)$ is equal to ± 1 (indeed, if $g \in A_N$, then $\det_k(g) = N_{K/k}\big(\det_K(g)\big)$ and one checks that $N_{K/k}\big(z_{\ell^m}\big) = 1$). This shows that an ℓ-Sylow of A_N is contained in $\mathbf{SL}_{Nt}(k)$ and hence in $\mathbf{SL}_n(k)$. By (9.2.1) we get an optimal pair.

In the case $t = 1$, we use the same construction with A_N^1 instead of A_N. The comparison of (9.1.2) and (9.2.2) shows that we get an optimal pair.

9.3. **The case of the orthogonal and symplectic groups, t odd.** Let us consider the case of $\mathbf{Sp}_{2n}$. The d_i's are equal to $2, 4, \ldots, 2n$. Hence, if we put $N = \left[\frac{n}{t}\right]$, the M-bound is $mN + v_\ell(N!)$. There is a natural embedding:

$$\mathbf{GL}_N \to \mathbf{Sp}_{2N} \to \mathbf{Sp}_{2n}$$

defined by $x \mapsto \begin{pmatrix} x & 0 \\ 0 & {}^t x^{-1} \end{pmatrix}$. The image of A_N by that embedding is optimal.

The same construction works for $\mathbf{SO}_{2n}$ and $\mathbf{SO}_{2n+1}$. (Note that, in all these cases, we get the *split* forms of the groups of type B_n, C_n, D_n. This is no longer true in the case t is even – nor in the cases of §12.)

9.4. **The case of the orthogonal and symplectic groups, t even.** Since t is even, the group $C_t = \mathrm{Gal}(K/k)$ contains an element σ of order 2; its image in $\mathbf{Z}_\ell^*$ is -1. Let K_0 be the subfield of K fixed by σ; we have $[K : K_0] = 2$, $[K_0 : k] = t_0$ with $t_0 = t/2$. Moreover $\sigma(z_{\ell^m})$ is equal to $(z_{\ell^m})^{-1}$; i.e. σ acts on z_{ℓ^m} just as complex conjugation does. Let us define an *hermitian form* h on K^N (where N is a given integer ≥ 1) by the standard formula

$$h(x,y) = \sum_{i=1}^{N} x_i \cdot \sigma(y_i), \qquad \text{if } x = (x_1, \ldots, x_N),\ y = (y_1, \ldots, y_N).$$

If $\mathbf{U}_N$ denotes the *unitary group* associated with h, it is clear that *the group A_N defined in §9.1 is contained in $\mathbf{U}_N(K)$.* [We use here the traditional notation $\mathbf{U}_N(K)$ for the unitary group; this is a bit misleading, since $\mathbf{U}_N$ is an algebraic group over K_0, and we are taking its K_0-points.]

Let $\delta \in K^*$ be such that $\sigma(\delta) = -\delta$, e.g. $\delta = z_\ell - z_\ell^{-1}$. We have $K = K_0 \oplus \delta \cdot K_0$, and $h(x,y)$ can be decomposed as

$$h(x,y) = q_0(x,y) + \delta \cdot b_0(x,y), \text{ with } q_0(x,y) \in K_0, \quad b_0(x,y) \in K_0.$$

Then q_0 (resp. b_0) is a non-degenerate symmetric (resp. alternating) K_0-bilinear form of rank $2N$.
Its trace $q = \mathrm{Tr}_{K_0/k}\, q_0(x,y)$ (resp. $b = \mathrm{Tr}_{K_0/k}\, b_0(x,y)$) is of rank $2Nt_0 = Nt$ over k. We thus get embeddings:

$$A_N \to \mathbf{U}_N(K) \to \mathbf{SO}_{2N}(K_0) \to \mathbf{SO}_{Nt}(k) \tag{9.4.1}$$

$$A_N \to \mathbf{U}_N(K) \to \mathbf{Sp}_{2N}(K_0) \to \mathbf{Sp}_{Nt}(k). \tag{9.4.2}$$

Now, for a given n, let us define N by $N = \left[\frac{2n}{t}\right] = \left[\frac{n}{t_0}\right]$. By (9.4.2), we get an embedding

$$A_N \to \mathbf{Sp}_{Nt}(k) \to \mathbf{Sp}_{2n}(k),$$

and one checks that it is optimal.

The same method gives an embedding of A_N into $\mathbf{SO}_{Nt}(k)$, hence into $\mathbf{SO}_{2n+1}(k)$, and this embedding is also optimal. As for $\mathbf{SO}_{2n}(k)$, one has to be more careful. The method does give an embedding of A_N into the $\mathbf{SO}_{2n}$ group relative to some quadratic form Q, but we have to ensure that such an $\mathbf{SO}_{2n}$ group is *of inner type* i.e. that $\mathrm{disc}(Q) = (-1)^n$ in k^*/k^{*2}. There are three cases:

a) If $2n > Nt$ (i.e. if t does not divide $2n$), we choose $Q = q \oplus q_1$, where q_1 has rank $2n - Nt$, and is such that $\mathrm{disc}(q) \cdot \mathrm{disc}(q_1) = (-1)^n$. We then have $A_N \subset \mathbf{SO}_{2n,Q}(k)$ and this is optimal.

b) If $2n = Nt$ and N is even, we have $\mathrm{disc}(q) = d^N$, where $d = \mathrm{disc}(K_0/k)$, hence $\mathrm{disc}(q) = 1$ in k^*/k^{*2}, which is the same as $(-1)^n$ since n is even.

c) If $2n = Nt$ and N is odd, we use an optimal subgroup A of $\mathbf{SO}_{2n-1}(k)$ relative to a quadratic form q_0 of rank $2n-1$. By adding to q_0 a suitable quadratic form of rank 1, we get a quadratic form of rank $2n$ and discriminant $(-1)^n$, as wanted. The corresponding embedding

$$A \to \mathbf{SO}_{2n-1}(k) \to \mathbf{SO}_{2n}(k)$$

is optimal. (Note that the d_i's for type D_n are $2, 4, \ldots, 2n-2$, and n. Hence, if $t \nmid n$, the M-bound for D_n is the same as the M-bound for B_{n-1}.)

§10. Galois twists

To handle exceptional groups, we have to use *twisted* inner forms instead of split ones. We shall only need the most elementary case of twisting, namely the one coming from a homomorphism $\varphi : \Gamma_k \to \mathrm{Aut}(G)$. Let us recall what this means (cf. for example [Se 64], chapter III):

Let K/k be a finite Galois extension. Let X be an algebraic variety over k, assumed to be quasi-projective (the case where X is affine would be enough). Choose a homomorphism

$$\varphi : \mathrm{Gal}(K/k) \to \mathrm{Aut}_k\, X.$$

The *twist* X_φ of X by φ is a variety over k which can be characterized as follows:

There is a K-isomorphism $\theta : X_{/K} \to X_{\varphi/K}$ such that $\gamma(\theta) = \theta \circ \varphi(\gamma)$ for every $\gamma \in \mathrm{Gal}(K/k)$.

(Here $X_{/K}$ denotes the K-variety deduced from X by the base change $k \to K$, and $\varphi(\gamma) \in \mathrm{Aut}_k\, X$ is viewed as belonging to $\mathrm{Aut}_K\, X_{/K}$.)

One shows (as a special case of Galois descent) that such a pair (X_φ, θ) exists, and is unique, up to isomorphism.

It is sometimes convenient to identify the K-points of X and X_φ via the isomorphism θ. But one should note that this is not compatible with the natural action of $\mathrm{Gal}(K/k)$ on $X(K)$ and $X_\varphi(K)$; one has

$$\gamma\big(\theta(x)\big) = \varphi(\gamma)\big(\gamma(x)\big) \qquad \text{if } \gamma \in \mathrm{Gal}(K/k), x \in X(K).$$

In other words, if we identify $X_\varphi(K)$ with $X(K)$, an element γ of $\mathrm{Gal}(K/k)$ acts on $X_\varphi(K)$ by the *twisted action* :

$$x \mapsto \varphi(\gamma)\big(\gamma(x)\big)$$

In particular, the *k-rational points of X_φ* correspond (via θ^{-1}) to the points $x \in X(K)$ such that $\gamma(x) = \varphi(\gamma^{-1})x$ for every $\gamma \in \mathrm{Gal}(K/k)$.

In what follows we apply the φ-twist to $X =$ split form of G, with $\varphi(\gamma)$ being a k-automorphism of G for every $\gamma \in \mathrm{Gal}(K/k)$. In that case, G_φ is a k-form of G; this form is inner if all $\varphi(\gamma)$ belong to $G^{\mathrm{ad}}(K)$ where G^{ad} is the adjoint group of G. The effect of the twist is to make k-rational some elements of G which were not. In order to define φ, we shall have to use the k-automorphisms of G provided by the Tits group W^*, see next section.

§11. A general construction

Here, G is a split simply connected group over k, and T is a maximal split torus of G. We put $N = N_G(T)$ and $W = N/T$ is the Weyl group.

11.1. **The Tits group.** The exact sequence $1 \to T \to N \to W \to 1$ does not split in general. However Tits ([Ti 66a], [Ti 66b]) has shown how to construct a subgroup[3] W^* of $N(k)$ having the following properties:

(1) The map $W^* \to W$ is surjective.

(2) The group $W^* \cap T$ is equal to the subgroup T_2 of T made up of the points x of T with $x^2 = 1$.

We thus have a commutative diagram, where the vertical maps are inclusions:

$$\begin{array}{ccccccccc} 1 & \to & T_2 & \to & W^* & \to & W & \to & 1 \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \to & T & \to & N & \to & W & \to & 1 \end{array}$$

We refer to Tits (*loc. cit.*) and to Bourbaki[4] ([LIE X], pp. 115–116, exerc. 12, 13) for the construction and the properties of W^*. For instance:

If G comes from a split group scheme $\underline{G}$ over $\mathbf{Z}$, then W^* is equal to $\underline{N}(\mathbf{Z})$, the group of *integral points* of the group scheme $\underline{N}$.

[3]The construction of W^* depends on more than (G, T): one needs a *pinning* ("épinglage") of (G, T) in the sense of [SGA 3], XXIII.1.1.

[4]Bourbaki works in the context of compact real Lie groups; his results can easily be translated to the algebraic setting we use here.

In the case of $\mathbf{SL}_n$, this means that one can choose for W^* the group of monomial matrices with non-zero entries ± 1 and determinant 1. For $n = 2$, W^* is the cyclic group of order 4 generated by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Note also that W^* is a quotient of the *braid group* $\mathbf{B}_W$ associated to W. (For the definition of the braid group of a Coxeter group, see e.g. [BM 97].)

11.2. **Special elements of W.** We now go back to our general notation $\ell, m, t, \ldots$ of Lecture II. Recall that the M-bound $M = M(\ell, k, R)$ is given by

$$M = \sum_{t|d_i} (m + v_\ell(d_i)), \qquad \text{cf. §6.2.} \tag{11.2.1}$$

Let $a(t)$ be the number of indices i such that $d_i \equiv 0 \pmod{t}$. We may rewrite (11.2.1) as

$$M = ma(t) + \sum_{t|d_i} v_\ell(d_i). \tag{11.2.2}$$

Note that, if no d_i is divisible by t, we have $M = 0$ and the trivial group $A = 1$ is optimal. Hence *we shall assume in what follows that* $a(t) \geq 1$.

Let now w be an element of W. We shall say that w is *special* (with respect to t and ℓ) if it has the following four properties:
(1) w has order t in W.
(2) w is the image of an element w^* of W^* such that $(w^*)^t \in T_2 \cap C(G)$, where $C(G)$ is the center of G.
(3) The characteristic polynomial of w (in the natural r-dimensional representation of W) is divisible by $(\Phi_t)^{a(t)}$, where Φ_t is the t-th cyclotomic polynomial.
(Equivalently: if z_t denotes a primitive t-th root of unity, then z_t is an eigenvalue of w of multiplicity at least $a(t)$.)
(4) Let $C_W(w)$ be the centralizer of w in W. Then:

$$v_\ell(C_W(w)) \geq \sum_{t|d_i} v_\ell(d_i).$$

Remark. The reader may wonder whether special elements exist for a given pair (t, ℓ) (with $a(t) > 0$ and $\ell \equiv 1 \pmod{t}$, of course). The answer is "no" in general: if R is of type C_3 and $t = 4$, no element of W^* has both properties (1) and (2). Fortunately, the answer is "yes" for the exceptional types $G_2, \ldots, E_8$, cf. §12.

Example : the regular case. Suppose that $w \in W$ is *regular of order* t in the sense of Springer[5] ([Sp 74], bottom of p. 170 - see also [BM 97], §3). This means that w has an eigenvector v, with eigenvalue z_t, such that v does not belong to any reflecting hyperplane. *Then w is special* (for any ℓ with $\ell \equiv 1 \pmod{t}$). Indeed:

[5]With a slight difference: Springer requires $t > 1$ and we don't; it is convenient to view $w = 1$ as a regular element of W.

Note that, if t is given, there is a very simple criterion ensuring the existence of a regular element of W of order t: the number of indices i such that $d_i \equiv 0 \pmod{t}$ should be equal to the number of i's such that $d_i \equiv 2 \pmod{t}$, cf. Lehrer-Springer [LS 99], cor.5.5.

(1) is obvious.

(2) follows from the fact, proved in [BM 97], §3, that w has a lifting $\mathbf{w}$ in the braid group $\mathbf{B}_W$ with $\mathbf{w}^t = \boldsymbol{\pi}$, where $\boldsymbol{\pi}$ has an image π in W which belongs to $T_2 \cap C(G)$. In Bourbaki's notation ([LIE X], p.116) π is the canonical element z_G of the center of G.

(3) is proved in [Sp 74], th. 4.2.

(4) is proved in [Sp 74], th. 4.2, in the stronger form $|C_W(w)| = \prod_{t|d_i} d_i$.

Special cases

$t = 1$. Here $w = 1$ and $w^* = \pi$ (one could also take $w^* = 1$).

$t = 2$. Here $w = w_0 =$ longest element of W. When -1 belongs to W, one has $w_0 = -1$ and w_0^* is *central* in W^* (because $\mathbf{w}_0$ is central in $\mathbf{B}_W$, cf. [BM 97], 1.2 and 3.4). In that case the inner automorphism of G defined by w_0^* is a "Weyl-Chevalley involution": it acts on T by $t \mapsto t^{-1}$.

11.3. **An auxiliary result.**

Lemma 7. *Suppose $w \in W$ is special of order t. Then it is possible to choose a lifting w^* of w in W^* which satisfies*

(2*) $(w^*)^t \in T_2 \cap C(G)$

and

(4*) $v_\ell\,(C_{W^*}(w^*)) \geq \sum_{t|d_i} v_\ell(d_i)$.

Proof. Let P be an ℓ-Sylow of $C_W(w)$; the groups P and $\langle w\rangle$ commute, and $P \cap \langle w\rangle = 1$ since w has order t and ℓ is prime to t (since $\ell \equiv 1 \bmod t$). Hence the group P_w generated by w and P is the direct product $P \times \langle w\rangle$. Since $\ell \neq 2$, its 2-Sylow subgroup is contained in $\langle w\rangle$. Put $C_2 = T_2 \cap C(G)$. We have an exact sequence:

$$1 \to T_2/C_2 \to W^*/C_2 \to W \to 1.$$

By property (2) of w, this exact sequence splits over $\langle w\rangle$, hence over the 2-Sylow of P_w; since the order of T_2/C_2 is a power of 2, this implies that it splits over P_w. We thus get an element w' of W^*/C_2, of order t, which lifts w, and centralizes a subgroup P' of W^*/C_2 isomorphic to P. We then choose for w^* a representative of w' in W^*; it has property (2*), moreover its centralizer contains the inverse image of P', which is canonically isomorphic to $C_2 \times P'$. By property (4) we have

$$v_\ell(P') = v_\ell(P) \geq \sum_{t|d_i} v_\ell(d_i).$$

This shows that w^* has property (4*). □

Remark. In the case where w is regular, one can do without lemma 7. Indeed the braid group construction of [BM 97] gives a lifting w^* of w having property (2*) and such that the map $C_{W^*}(w^*) \to C_W(w)$ is surjective.

11.4. **The main result.**

Proposition 5. *Suppose W contains an element w which is special with respect to t and ℓ. Then there exist an inner twist G_φ of G* (cf. §10) *and a finite ℓ-subgroup A of $G_\varphi(k)$ such that the pair (G_φ, A) is optimal in the sense of* §7.

(In particular, th.9 is true for (k, ℓ, R).)

Proof. As in §9, we put $K = k(z_\ell)$, where z_ℓ is a root of unity of order ℓ. Let $C_t = \mathrm{Gal}(K/k)$; it is a cyclic group of order t.

Choose $w^* \in W^*$ with the properties of lemma 7 and let σ be the inner automorphism of G defined by w^*. Since σ has order t, there exists an injective homomorphism:

$$\varphi : C_t \to G^{\mathrm{ad}}(k) \subset \mathrm{Aut}_k(G)$$

which maps C_t onto the subgroup $\langle\sigma\rangle$ of $\mathrm{Aut}_k(G)$ generated by σ. As explained in §10, we may then define the *φ-twist* G_φ of G, relatively to the Galois extension K/k. The group G_φ is an inner form of G; it has the same root system R. It remains to construct a finite ℓ-subgroup A of $G_\varphi(k)$ such that (G_φ, A) is optimal, i.e. $v_\ell(A) = ma(t) + \sum_{t|d_i} v_\ell(d_i)$, cf. (11.2.2).

We take for A the semi-direct product $E_m \cdot P$, with $E_m \subset T_\varphi(k)$ and $P \subset N_\varphi(k)$, where E_m and P are defined as follows:

(1) P is an ℓ-Sylow of $C_{W^*}(w^*)$. By lemma 7 we have $v_\ell(P) \geq \sum_{t|d_i} v_\ell(d_i)$. Note that the points of P are fixed by σ. Hence these points are rational over k not only in the group G but also in the group G_φ.

(2) E_m is the subgroup of $T_\varphi(k)$ made up of the elements x such that $x^{\ell^m} = 1$.

It is clear that P normalizes E_m, and that $P \cap E_m = 1$.

Lemma 8. *The group E_m contains a product of $a(t)$ copies of the group $\mathbf{Z}/\ell^m\mathbf{Z}$.*

This implies that $v_\ell(E_m) \geq ma(t)$ and hence

$$v_\ell(A) = v_\ell(E_m) + v_\ell(P) \geq ma(t) + \sum_{t|d_i} v_\ell(d_i).$$

We thus get $v_\ell(A) \geq M$ and since M is an upper bound for $v_\ell(A)$ we have $v_\ell(A) = M$. □

Proof of lemma 8. Consider first the subgroup T_{ℓ^m} of $T(k_s)$ made up of the elements x with $x^{\ell^m} = 1$. Since T is k-split, and $K = k(z_\ell) = k(z_{\ell^m})$ (cf. §4 and §9), the points of T_{ℓ^m} are rational over K. If we write $T_{\ell^m}(K)$ additively, it becomes a free $\mathbf{Z}/\ell^m\mathbf{Z}$-module of rank r and the action of a generator s of C_t is by $x \mapsto sx$, where s is identified with an element of order t in $\mathbf{Z}_\ell^*$ (i.e. $s =$"z_t" with our usual notation for roots of unity). As for the action of w^* (i.e. of w) on $T_{\ell^m}(K)$, it can be put in diagonal form since w is of order t and t divides $\ell - 1$; its diagonal elements are r elements $y_1, \ldots, y_r$ of $\mathbf{Z}/\ell^m\mathbf{Z}$, with $y_i^t = 1$. Let c be the largest integer such that $(\Phi_t)^c$ divides the characteristic polynomial of w. By property (3) of 11.2, we have $c \geq a(t)$ (in fact, $c = a(t)$, by [Sp 74], th. 3.4). This implies that the family of the y_i's contains c times each primitive t-th root of unity (viewed as element of $(\mathbf{Z}/\ell^m\mathbf{Z})^*$). In particular, there is a $\mathbf{Z}/\ell^m\mathbf{Z}$-submodule X of $T_{\ell^m}(K)$ which is free of rank c and on which w acts by $x \mapsto z_t^{-1}x$. If we twist G, T, T_{ℓ^m} by φ, the new action of $C_t = \mathrm{Gal}(K/k)$ on X is trivial (cf. end of §10). This means that X is contained in $T_\varphi(k)$, hence in E_m, which proves the lemma. □

Note the following consequence of proposition 4:

Corollary. *If W contains a t-regular element in the sense of* [Sp 74]*, then theorem 9 is true for k, ℓ, R.*

In the case $t = 1$, no twist is necessary (one takes $w = 1$, $w^* = 1$, cf. §11.2).

§12. Proof of theorem 9 for exceptional groups

In each case we will show that the Weyl group contains an element w which is special with respect to t and ℓ, so that we may apply prop.5.

12.1. **The case of G_2.** The degrees d_i are $d_1 = 2, d_2 = 6$. Since t divides one of them, t is a divisor of 6, hence is regular ([Sp 74], no. 5.4). We may then apply prop.5. □

Explicit description of w, w^*: if c is a Coxeter element of W, c is of order 6, and every lifting c^* of c in W^* has order 6. Hence, for any divisor t of 6, we may take $w = c^{6/t}$ and $w^* = (c^*)^{6/t}$.

12.2. **The case of F_4.** The d_i's are: 2, 6, 8, 12. All their divisors are regular (Springer, *loc. cit.*). One concludes as for G_2. □

12.3. **The case of E_6.** The d_i's are: 2, 5, 6, 8, 9, 12. All their divisors are regular, except $t = 5$. In that case, choose any element $w \in W$ of order 5. Since the kernel of $W^* \to W$ is a 2-group, w can be lifted to an element w^* of W^* of order 5. Conditions (1) and (2) of §11.2 are obviously satisfied. The same is true for condition (3), since $a(5) = 1$ (only one of the d_i's is divisible by 5), and w has at least one eigenvalue of order 5. As for condition (4), it is trivial, since $\ell \equiv 1 \pmod 5$ implies $\ell \geq 11$, and ℓ does not divide any of the d_i's, so that $\sum_{t|d_i} v_\ell(d_i)$ is 0. Hence w is special with respect to $(5, \ell)$. □

12.4. **The case of E_7.** The d_i's are: 2, 6, 8, 10, 12, 14, 18. By [Sp 74], *loc.cit.* all their divisors are regular except 4, 5, 8, 10, 12. If $t = 4, 5, 8$ or 12, t already occurs for E_6, with the same values of $a(t)$, namely 2, 1, 1 and 1. Hence, we have E_6-special elements w_4, w_5, w_8 and w_{12} in $W(E_6)$. One then takes their images in $W(E_7)$ by the injective map $W(E_6) \to W(E_7)$, and one checks that they are E_7-special (here again condition (4) is trivial since $v_\ell(d_i) = 0$ for all the ℓ's with $\ell \equiv 1 \pmod t$).

As for $t = 10$, one takes $w = -w_5$, which makes sense since $-1 \in W$. The element -1 (usually denoted by w_0) can be lifted to a central element ε of W^* with $\varepsilon^2 \in T_2 \cap C(G)$; this is a general property of the case $-1 \in W$ (which reflects the fact that -1 is 2-regular, see end of §11.2). Hence, if w_5^* is a lifting of w_5 of order 5, εw_5^* is a lifting of w of order 10, and this shows that w is special with respect to 10 and ℓ. □

12.5. **The case of E_8.** The d_i's are: 2, 8, 12, 14, 18, 20, 24, 30. By [Sp 74], *loc.cit.*, all their divisors are regular except 7, 9, 14, 18.

If $t = 7$ (resp. 9), one chooses $w_7 \in W$ of order 7 (resp. $w_9 \in W$ of order 9). Since 7 and 9 are odd, condition (2) of §11.2 is satisfied. The same is true for condition (3) because $a(t) = 1$, and for condition (4) because $v_\ell(d_i) = 0$ for all i.

If $t = 14$ (resp. 18), one takes $w = -w_7$ (resp. $w = -w_9$), as we did for E_7. □

§13. Proof of theorems 10 and 11

Here $\ell = 2$. There are three cases (cf. §4.2):

(a) $\operatorname{Im}\chi_{2^\infty} = 1 + 2^m\mathbf{Z}_2$ with $m \geq 2$. In that case the M-bound is $rm + v_2(W)$, and th.10 asserts that an optimal pair (G, A) exists for every type R.

(b) $\operatorname{Im}\chi_{2^\infty} = \langle -1 + 2^m \rangle$, with $m \geq 2$. The M-bound is $r_0 m + r_1 + v_2(W)$, where r_0 (resp. r_1) is the number of i's such that d_i is odd (resp. even). Here, too, th.10 asserts that an optimal pair exists.

(c) $\operatorname{Im}\chi_{2^\infty} = \langle -1, 1 + 2^m \rangle$, with $m \geq 2$.

The M-bound is the same as in case (b), but th.11 does not claim that it can be met (i.e. that an optimal pair exists); it merely says that there is a pair (G, A) with $v_2(A) = r_0 m + v_2(W)$; such a pair is optimal only when $r_1 = 0$, i.e. when -1 belongs to the Weyl group.

13.1. **Proof of theorem 10 in case (a).** We take G split and simply connected, and we choose a maximal split torus T. We use the notation (N, W, W^*) of §11. Let E be the 2-torsion subgroup of $T(k)$. Since T is isomorphic to the product of r copies of $\mathbf{G}_m$, E is isomorphic to a product of m copies of $\mathbf{Z}/2^m\mathbf{Z}$, cf. §4.2. Hence $v_2(E) = rm$. The group E is normalized by the Tits group W^*; we define A as $A = E \cdot W^*$. The exact sequence

$$1 \to E \to A \to W \to 1$$

shows that $v_2(A) = rm + v_2(W)$. Hence (G, A) is optimal.

13.2. **Cases (b) and (c).** As in §11.4, we start with a split G, with a split maximal torus T. We define N, W, W^* as usual. After choosing an order on the root system R, we may view W as a Coxeter group; let w_0 be its longest element. It has order 2, and it is regular in the sense of Springer [Sp 74]. As explained in §11.2, this implies that there is a lifting w_0^* of w_0 in W^* which has the following two properties:

(i) its square belongs to the center of G;

(ii) the natural map $C_{W^*}(w_0^*) \to C_W(w_0)$ is surjective.

Let σ be the inner automorphism of G defined by w_0^*. By (i), we have $\sigma^2 = 1$. Let $K = k(i)$ and let φ be the homomorphism of $\operatorname{Gal}(K/k)$ into $\operatorname{Aut}_k(G)$ whose image is $\{1, \sigma\}$. Let us define G_φ as the *φ-twist* of G, in the sense defined in §10. Denote by T_φ, N_φ and W_φ^* the φ-twists of T, N and W^*. We have an exact sequence

$$1 \to T_\varphi \to N_\varphi \to W_\varphi \to 1,$$

where W_φ is the φ-twist of W. Note that $W_\varphi(k)$ is equal to the centralizer $C_W(w_0)$ of w_0 in W, and similarly $W_\varphi^*(k)$ is equal to $C_{W^*}(w_0^*)$.

As in §13.1, let E be the 2-torsion subgroup of $T_\varphi(k)$. It is normalized by $C_{W^*}(w_0^*)$. Define $A \subset G_\varphi(k)$ to be the group $A = E \cdot C_{W^*}(w_0^*)$. By (ii), we have an exact sequence:

$$1 \to E \to A \to C_W(w_0) \to 1\,,$$

which shows that $v_2(A) = v_2(E) + v_2\big(C_W(w_0)\big)$. The fact that w_0 is regular of order 2 implies that

$$|C_W(w_0)| = \prod_{2|d_i} d_i,$$

hence $v_2\big(C_W(w_0)\big) = \sum v_2(d_i) = v_2(W)$. This gives:

$$v_2(A) = v_2(E) + v_2(W). \tag{13.2.1}$$

Proposition 6. *We have* :

$v_2(E) = r_1 + r_0 m$ *in case* (b)
$v_2(E) = r_0 m$ *in case* (c).

In case (b), this shows that (G_φ, A) is optimal, which proves th.10. Similarly, the fact that $v_2(A) = r_0 m + v_2(W)$ proves th.11 in case (c).

13.3. **Proof of proposition 6.** We need to describe explicitly the torus T_φ. To do so, let us first define the following two tori:

$\mathbf{G}_m^\sigma$ = 1-dimensional torus deduced from $\mathbf{G}_m$ by Galois twist relatively to K/k. Its group of k-points is $K_1^* = \operatorname{Ker} N_{K/k} : K^* \to k^*$.

$R_{K/k}\mathbf{G}_m$ = 2-dimensional torus deduced from $\mathbf{G}_m$ by Weil's restriction of scalars relatively to K/k. Its group of k-points is K^*.

Lemma 9. *The torus T_φ is isomorphic to the product of r_1 copies of $R_{K/k}\mathbf{G}_m$ and $r_0 - r_1$ copies of $\mathbf{G}_m^\sigma$.*

Proof. The character group $X = \operatorname{Hom}(T, \mathbf{G}_m)$ is free of rank r, with basis the fundamental weights $\omega_1, \ldots, \omega_r$. This gives a decomposition of T as

$$T = T_1 \times T_2 \times \cdots \times T_r\,,$$

where each T_i is canonically isomorphic to $\mathbf{G}_m$. Let $\tau = -w_0$ be the opposition involution of the root system R; it permutes $\omega_1, \ldots, \omega_r$ with r_1 orbits of order 2, and $r_0 - r_1$ orbits of order 1. (This follows from the fact that -1 is an eigenvalue of w_0 of multiplicity r_0.) The involution τ permutes the tori T_j. If an index j is fixed by τ, then w_0 acts on T_j by $t \mapsto t^{-1}$ and the twisted torus $(T_j)_\varphi$ is isomorphic to $\mathbf{G}_m^\sigma$; similarly, if τ permutes j and j', the torus $(T_j \times T_{j'})_\varphi$ is isomorphic to $R_{K/k}\mathbf{G}_m$. This proves lemma 9. □

End of the proof of prop.6. The 2-torsion subgroup of $\mathbf{G}_m^\sigma(k) = K_1^*$ is cyclic of order 2^m; the 2-torsion subgroup of $R_{K/k}\mathbf{G}_m(k) = K^*$ is cyclic of order 2^{m+1} in case (b) and of order 2^m in case (c). We get what we wanted, namely:

case (b): $v_2(E) = r_1(m+1) + (r_0 - r_1)m = r_0 m + r_1$
case (c): $v_2(E) = r_1 m + (r_0 - r_1)m = r_0 m$.

This completes the proof of prop.6, and hence of th.10 and th.11. □

13.4. **Remarks on the non simply connected case.** The proof above could have been given without assuming that the split group G is simply connected. The main difference is in lemma 9: in the general case, the torus T_φ is a product of three factors (instead of two):

$$T_\varphi = (\mathbf{G}_m)^\alpha \times (\mathbf{G}_m^\sigma)^\beta \times (R_{K/k}\mathbf{G}_m)^\gamma\,,$$

where α, β, γ are integers, with $\beta + \gamma = r_0$ and $\alpha + \gamma = r_1$. This gives the following formulae for $v_2(E)$:

case (b) : $v_2(E) = \alpha + \beta m + \gamma(m+1) = r_1 + r_0 m$

case (c) : $v_2(E) = \alpha + \beta m + \gamma m = \alpha + r_0 m$.

In case (b) one finds the same value for $v_2(A)$, namely the M-bound. In case (c) one finds a result which is intermediate between the M-bound $r_1 + r_0 m + v_2(W)$ and the value $r_0 m + v_2(W)$ given by th.11.

Examples (assuming we are in case (c)).

- *Type A_r, r even.* One finds that α is always 0, so that one does not gain anything by choosing non simply connected groups. Indeed, in that case, it is possible to prove, by a variant of Schur's method, that the value of $v_2(A)$ given by th.11 is best possible.

- *Type A_r, r odd ≥ 3.* Here $r_1 = (r-1)/2$. One finds that $\alpha = 0$ if $r \equiv 1 \pmod 4$, but that α can be equal to 1 if $r \equiv 3 \pmod 4$. When $r = 3$, we thus get $\alpha = r_1$; this shows that the M-bound is best possible for type A_3.

- *Type D_r, r odd.* Here $r_1 = 1$, and if one chooses G neither simply connected nor adjoint, one has $\alpha = 1$. This means that the orthogonal group $\mathbf{SO}_{2r}$ has an inner k-form which contains an optimal A. (Note the case $r = 3$, where $D_3 = A_3$.)

- *Type E_6.* Here $r_1 = 2$, and one has $\alpha = 0$ both for the simply connected group and for the adjoint group (indeed, α is 0 for every adjoint group). I do not know whether the bound of th.11 is best possible in this case.

§14. The case $m = \infty$

14.1. **Statements.** We keep the notation $(G, R, W, d_i, \ell, t, m)$ of §4 and §6; as before, we assume that G is of inner type.

We consider the case $m = \infty$, i.e. the case where *the image of χ_{ℓ^∞} is finite*; that image is then cyclic of order t, cf. §4.

Let $a(t)$ be the number of i's such that $d_i \equiv 0 \pmod t$. If $a(t) = 0$, then $G(k)$ is ℓ-torsion free, cf. §6.2, cor.to prop. 4. In what follows, we shall thus assume that $a(t) \geq 1$. In that case, $G(k)$ may contain infinite ℓ-subgroups (we say that a group is an *ℓ-group* if every element of that group has order a power of ℓ). The following two theorems show that $a(t)$ controls the size of such a subgroup:

Theorem 12. *Let A be an ℓ-subgroup of $G(k)$. Then A contains a subgroup of finite index isomorphic to the ℓ-group $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^a$, with $a \leq a(t)$.*

(Note that $\mathbf{Q}_\ell/\mathbf{Z}_\ell$ is the union of an increasing sequence of cyclic groups of order $\ell, \ell^2, \ldots$; it is the analogue of $\mathbf{Z}/\ell^m\mathbf{Z}$ for $m = \infty$.)

The bound $a \leq a(t)$ of th.12. is optimal. More precisely:

Theorem 13. *There exist a semisimple group G of inner type, with root system R, and an ℓ-subgroup A of $G(k)$, such that A is isomorphic to the product of $a(t)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.*

14.2. **Proof of theorem 12.** We need a few lemmas:

Lemma 10. *Any finitely generated ℓ-subgroup of $G(k)$ is finite.*

Proof. Let B be a finitely generated ℓ-subgroup of $G(k)$. We may embed B in $\mathbf{GL}_n(k)$ for n large enough. By a known result (see §1.2) there exists a subgroup B' of B, of finite index, which is torsion-free if $\mathrm{char}(k) = 0$, and has only p-torsion if $\mathrm{char}(k) = p$. Since B' is an ℓ-group, this means that $B' = 1$, hence B is finite. □

Lemma 11. *There exists a maximal k-torus of G which is normalized by A.* (Recall that A is an ℓ-subgroup of $G(k)$.)

Proof. Let F be the set of all finite subgroups of A, ordered by inclusion. Lemma 10 implies that, if B_1 and B_2 belong to F, so does $\langle B_1, B_2\rangle$. Let X be the k-variety parametrizing the maximal tori of G; it is a homogeneous space of G. If $B \in F$, let X^B be the subvariety of X fixed by B; a point of X^B corresponds to a maximal torus of G normalized by B. By the noetherian property of the scheme X, one may choose $B_0 \in F$ such that X^{B_0} is minimal among the X^B 's. If $B \in F$, then $X^{\langle B_0, B\rangle}$ is contained in X^{B_0}, hence equal to X^{B_0}. This shows that X^{B_0} is contained in all the X^B 's, i.e. that every maximal torus which is normalized by B_0 is normalized by all the B's, hence by A. By the corollary to th.3″ of §3.3 (applied to the finite ℓ-group B_0) there exists such a torus which is defined over k. □

Lemma 12. *Let $u \in \mathbf{M}_r(\mathbf{Z}_\ell)$ be an $r \times r$ matrix with coefficients in $\mathbf{Z}_\ell$, which we view as an endomorphism of $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^r$. Then* $\mathrm{Ker}(u)$ *has a subgroup of finite index isomorphic to the product of $r - \mathrm{rank}(u)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.*

In other words, the "corank" of $\mathrm{Ker}(u)$ is equal to $r - \mathrm{rank}(u)$.

Proof. Same as that of lemma 4 of §5.2: by reduction to the case where u is a diagonal matrix. □

Lemma 13. *Let z_t be a primitive t-th root of unity, and let w be an element of W. The multiplicity of z_t as an eigenvalue of w is $\leq a(t)$.*

Proof. See [Sp 74], th.3.4(i) where it is deduced from the fact that the polynomial $\det(t - w)$ divides $\prod_i (t^{d_i} - 1)$. □

Lemma 14. *Let T be a maximal k-torus of G, and let $T(k)_\ell$ be the ℓ-torsion subgroup of $T(k)$. We have* $\mathrm{corank}\ T(k)_\ell \leq a(t)$.

As above, the "corank" of a commutative ℓ-group is the largest n such that the group contains the product of n copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.

Proof. As in §5.2, let $Y(T) = \mathrm{Hom}_{k_s}(\mathbf{G}_m, T)$ be the group of cocharacters of T. The action of the Galois group Γ_k on $Y(T)$ gives a homomorphism

$$\rho : \Gamma_k \to \mathrm{Aut}\, Y(T) \simeq \mathbf{GL}_r(\mathbf{Z})$$

and the image of ρ is contained in the Weyl group W (this is still another way of saying that G is of inner type). The group Γ_k acts on $T(k_s)_\ell \simeq (\mathbf{Q}_\ell/\mathbf{Z}_\ell)^r$ by $\rho \otimes \chi$, where $\chi = \chi_{\ell^\infty}$. Let us now choose $g \in \Gamma_k$ such that $\chi(g) = z_t^{-1}$, where z_t is an element of order t of $\mathbf{Z}_\ell^*$, and let $w = \rho(g)$. The element g acts on $T(k_s)_\ell$ by wz_t^{-1}. Let T_g be the kernel of $g - 1$ on $T(k_s)_\ell$. By lemma 12, we have $\mathrm{corank}\ (T_g) = r - \mathrm{rank}(g-1)$, which is equal to the multiplicity of z_t as an eigenvalue of w; using lemma 13, we get $\mathrm{corank}(T_g) \leq a(t)$, and since $T(k)_\ell$ is contained in T_g, we have $\mathrm{corank}(T(k)_\ell) \leq a(t)$. □

End of the proof of th.12. By lemma 11, there is a maximal k-torus T of G which is normalized by A. Let $A^\circ = A \cap T(k)$. Then A° is an abelian subgroup of A of finite index. Since A° is contained in $T(k)_\ell$, lemma 14 shows that A° is isomorphic to the product of a finite group with a product of at most $a(t)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$. □

14.3. **Proof of theorem 13.** We follow the same strategy as for theorem 9, 10 and 11. There are three cases:

14.3.1. **Classical groups** ($\ell \neq 2$). We change slightly the definitions of §9.1: we define A_N as the subgroup of $\mathbf{GL}_N(K)$, with $K = k(z_\ell)$, made up of the diagonal matrices of order a power of ℓ; it is isomorphic to $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^N$.

For any given $n \geq 2$, we put $N = [n/t]$ and we get embeddings

$$A_N \to \mathbf{GL}_N(K) \to \mathbf{GL}_{Nt}(k) \to \mathbf{GL}_n(k).$$

If $t > 1$, one checks that the k-determinant of every element of A_N is 1; we thus get an embedding $A_N \to \mathbf{SL}_n(k)$ which has the required properties since $N = a(t)$ in that case. When $t = 1$, we replace A_N by the subgroup of its elements of k-determinant 1, and we also get what we want. This solves the case of type A_r. Types B_r, C_r and D_r are then treated by the methods of §9.3 and §9.4.

14.3.2. **Exceptional groups** ($\ell \neq 2$). One replaces prop.5 of §11.4 by a statement giving the existence of $A \subset G_\varphi(k)$ with $A \simeq (\mathbf{Q}_\ell/\mathbf{Z}_\ell)^{a(t)}$. The proof is the same. One then proceeds as in §12.

14.3.3. **The case** $\ell = 2$. Same method as in §13. □

References

[A V] N. Bourbaki, *Algèbre, Chapitre* V, Masson, Paris, 1981.

[AC N] N. Bourbaki, *Algèbre Commutative, Chapitre* N, Hermann-Masson, Paris, 1961–1998.

[Bl 04] H. Blichfeldt, *On the order of linear homogeneous groups,* Trans. Amer. Math. Soc. **5** (1904), 310–325.

[BM 97] M. Broué and J. Michel, *Sur certains éléments réguliers des groupes de Weyl et les variétés de Deligne–Lusztig associées*, in Finite Reductive Groups: Related Structures and Representations, M. Cabanes (*edit.*), Progress in Math. **141**, Birkhäuser– Boston, 1997, 73–139.

[Bo 69] A. Borel, *Groupes arithmétiques*, Hermann, Paris 1969.

[Bo 91] A. Borel, *Linear Algebraic Groups*, second edition, Springer-Verlag, 1991.

[Br 74] K. Brown, *Euler characteristics of discrete groups and G-spaces,* Invent. math. **27** (1974), 229–264.

[Br 01] M. Broué, *Reflection groups, braid groups, Hecke algebras, finite reductive groups,* in Current Developments in Mathematics 2000, International Press, 2001, 1-107.

[BS 53] A. Borel and J-P. Serre, *Sur certains sous-groupes des groupes de Lie compacts*, Comm.Math.Helv. **27** (1953), 128–139 (= A. Borel, Coll. Works, vol.I, n°24).

[Bu 11] W. Burnside, *Theory of Groups of Finite Order,* second edition, Cambridge Univ.Press. 1911; reprinted by Dover Publ., 1955.

[De 70] M. Demazure, *Sous-groupes algébriques de rang maximum du groupe de Cremona,* Ann.scient.E.N.S. (4) **3** (1970), 507–588.

[EGA IV] A. Grothendieck, *Eléments de Géométrie Algébrique* (rédigés avec la collaboration de J. Dieudonné), Chap.IV, Etude Locale des Schémas et des Morphismes de Schémas (Troisième Partie), Publ.Math.I.H.E.S. **28** (1966).

[Fe 97] W. Feit, *Finite linear groups and theorems of Minkowski and Schur*, Proc. A.M.S. **125** (1997), 1259-1262.

[FW 84] G. Faltings, G. Wüstholz et al, *Rational Points*, Seminar Bonn-Wuppertal 1983/1984, Vieweg, Braunschweig, 1984.

[GL 06] R.M. Guralnick and M. Lorenz, *Orders of finite groups of matrices*, Contemp.Math., to appear.

[GMS 03] S. Garibaldi, A. Merkurjev and J.-P. Serre, *Cohomological Invariants in Galois Cohomology,* A.M.S. Lect. Series **28** (2003).

[Gr 96] B. H. Gross, *Groups over* $\mathbf{Z}$*,* Invent.math. **124** (1996), 263–279.

[Ha 71] G. Harder, *A Gauss-Bonnet formula for discrete arithmetically defined groups,* Ann.Sci. E.N.S. (4) **4** (1971), 409–455.

[Il 06] L. Illusie, *Miscellany on traces in ℓ-adic cohomology: a survey*, Jap. J.Math., (new series), **1** (2006), 107–136.

[LIE N] N. Bourbaki, *Groupes et Algèbres de Lie, Chapitre* N, Hermann-Masson, Paris 1972-1982.

[LS 99] G.I. Lehrer and T.A. Springer, *Reflection subquotients of unitary reflection groups*, Canadian J. Math. **51** (1999), 1175-1193.

[Mi 87] H. Minkowski, *Zur Theorie der positiven quadratischen Formen,* J.Crelle **101** (1887), 196–202 (= Ges.Abh., Band I, n°VI).

[Pi 97] R. Pink, *The Mumford-Tate conjecture for Drinfeld-modules,* Publ. Res. Inst. Math. Sci. **33** (1997), 393–425.

[Ro 58] P. Roquette, *Realisierung von Darstellungen endlicher nilpotenter Gruppen*, Arch.Math. **9** (1958), 241-250.

[Sch 05] I. Schur, *Über eine Klasse von endlichen Gruppen linearer Substitutionen*, Sitz.Preuss.Akad.Wiss. Berlin (1905), 77–91 (= Ges.Abh., Band I, n° 6).

[Se 64] J.-P. Serre, *Cohomologie Galoisienne*, Lect.Notes in Math. **5**, Springer- Verlag, 1964; fifth revised edition, 1994; English translation: *Galois Cohomology*, corrected second printing, Springer–Verlag, 2002.

[Se 65] J.-P. Serre, *Zeta and L functions*, in Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ. 1963), 82–92, Harper and Row, New York, 1965 (= Oe.64).

[Se 71] J.-P. Serre, *Cohomologie des groupes discrets,* Ann.of Math. Studies **70**, 77–169, Princeton, 1971 (= Oe.88).

[Se 79] J.-P. Serre, *Arithmetic groups,* in Homological Group Theory, C.T. Wall edit., LMS Lect.Notes Series **36**, Cambridge Univ.Press (1979), 105–136 (= Oe.120).

[Se 81] J-P. Serre, *Lettres à Ken Ribet du* 1/1/1981 *et du* 29/1/81, reproduced in Coll. Papers IV, 1–20 (= Oe.133).

[Se 93] J.-P. Serre, *Gèbres*, L'Ens.Math. (2) **39** (1993), 33–85 (= Oe.160).

[Se 00] J.-P. Serre, *Local Algebra,* Springer-Verlag, 2000.

[SGA 3] M. Demazure and A. Grothendieck, *Schémas en Groupes*, Lect.Notes in Math. **151-153**, Springer-Verlag, 1970.

[SGA $4\frac{1}{2}$] P. Deligne *et al*, *Cohomologie Étale*, Lect.Notes in Math. **569**, Springer-Verlag, 1977.

[Sp 74] T. A. Springer, *Regular elements of finite reflection groups*, Invent.math. **25** (1974), 159–198.

[SS 68] T. A. Springer and R. Steinberg, *Conjugacy Classes*, in Seminar on Algebraic Groups and Related Finite Groups, Lect.Notes in Math. **131**, Springer-Verlag, 1970 (= R. Steinberg, Coll.Papers, n°25).

[St 67] R. Steinberg, *Lectures on Chevalley Groups,* Yale University, 1967.

[St 68] R. Steinberg, *Endomorphisms of linear algebraic groups,* A.M.S.Memoirs, **80**, 1968 (= Coll.Papers, n°23).

[SZ 96] A. Silverberg and Yu.G. Zarhin, *Variations on a theme of Minkowski and Serre*, J.Pure Applied Algebra **111** (1996), 285-302.

[Th 60-64] J.G. Thompson, *Normal p-complements for finite groups*, Math.Zeit. **72** (1960), 332–354 and J. Algebra **1** (1964), 43–46.

[Ti 66a] J. Tits, *Normalisateurs de tores.* I. *Groupes de Coxeter étendus*, J. Algebra **1** (1966), 96–116.

[Ti 66b] J. Tits, *Sur les constantes de structure et le théorème d'existence des algèbres de Lie semi-simples*, Publ.Math. I.H.E.S. **31** (1966), 21–58.

J.-P. Serre
Collège de France
3, rue d'Ulm
F-75005 PARIS.

Group Representation Theory, eds. M. Geck, D. Testerman, J. Thévenaz,
EPFL Press, Lausanne 2007, pp. 405-450

Bounds for the orders of the finite subgroups of $G(k)$

Jean-Pierre SERRE

Introduction

The present text reproduces - with a number of additions - a series of three two-hour lectures given at the Ecole Polytechnique Fédérale de Lausanne (E.P.F.L.) on May 25-26-27, 2005.

The starting point is a classical result of Minkowski, dating from 1887, which gives a multiplicative upper bound for the orders of the finite subgroups of $\mathbf{GL}_n(\mathbf{Q})$. The method can easily be extended to other algebraic groups than $\mathbf{GL}_n$, and the field $\mathbf{Q}$ can be replaced by any number field. What is less obvious is that:

a) one can work over an arbitrary ground field;

b) in most cases one may construct examples showing that the bound thus obtained is optimal.

This is what I explain in the lectures.

Lecture I is historical: Minkowski (§1), Schur (§2), Blichfeldt and others (§3). The results it describes are mostly well-known, so that I did not feel compelled to give complete proofs.

Lecture II gives upper bounds for the order of a finite ℓ-subgroup of $G(k)$, where G is a reductive group over a field k, and ℓ is a prime number. These bounds depend on G via its root system, and on k via the size of the Galois group of its ℓ-cyclotomic tower (§4). One of these bounds (called here the S-bound, cf. §5) is a bit crude but is easy to prove and to apply. The second one (called the M-bound) is the most interesting one (§6). Its proof follows Minkowski's method, combined with Chebotarev's density theorem (for schemes of any dimension, not merely dimension 1); it has a curious cohomological generalization cf. §6.8. The last subsection (§6.9) mentions some related problems, not on semisimple groups, but on Cremona groups; for instance: does the field $\mathbf{Q}(X, Y, Z)$ have an automorphism of order 11 ?

Lecture III gives the construction of "optimal" large subgroups. The case of the classical groups (§9) is not difficult. Exceptional groups such as E_8 are a different matter; to handle them, we shall use Galois twists, braid groups and Tits groups, cf. §§10-12.

Acknowledgements. A first draft of these notes, made by D. Testerman and R. Corran, has been very useful; and so has been the generous help of D. Testerman with the successive versions of the text. My thanks go to both of them, and to the E.P.F.L. staff for its hospitality. I also thank M. Broué and J. Michel for several discussions on braid groups.

J-P. Serre April 2006

Table of Contents

I. History: Minkowski, Schur, ...

§1. Minkowski

Reference: [Mi 87].

1.1. **Statements.** We shall use the following notation:

ℓ is a fixed prime number; when we need other primes we usually denote them by p;

the ℓ-adic valuation of a rational number x is denoted by $v_\ell(x)$; one has $v_\ell(\ell) = 1$, and $v_\ell(x) = 0$ if x is an integer with $(x, \ell) = 1$;

the number of elements of a finite set A is denoted by $|A|$; we write $v_\ell(A)$ instead of $v_\ell(|A|)$; if A is a group, $\ell^{v_\ell(A)}$ is the order of an ℓ-Sylow of A;

if x is a real number, its integral part ("floor") is denoted by $[x]$.

We may now state Minkowski's theorem ([Mi 87]):

Theorem 1. *Let n be an integer ≥ 1, and let ℓ be a prime number. Define*:

$$M(n,\ell) = \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \left[\frac{n}{\ell^2(\ell-1)}\right] + \cdots$$

Then:
(i) *If A is a finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$, we have $v_\ell(A) \leq M(n,\ell)$.*
(ii) *There exists a finite ℓ-subgroup A of $\mathbf{GL}_n(\mathbf{Q})$ with $v_\ell(A) = M(n,\ell)$.*

The proof will be given in §1.3 and §1.4.

Remarks.

1) Let us define an integer $M(n)$ by:

$$M(n) = \prod_\ell \ell^{M(n,\ell)}.$$

Part (i) of th.1 says that the order of any finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$ *divides* $M(n)$, and part (ii) says that $M(n)$ is the smallest integer having this property. Hence $M(n)$ is a sharp multiplicative bound for $|A|$.

Here are the values of $M(n)$ for $n \leq 8$:

$M(1) = 2$
$M(2) = 2^3 \cdot 3 = 24$
$M(3) = 2^4 \cdot 3 = 48$
$M(4) = 2^7 \cdot 3^2 \cdot 5 = 5760$
$M(5) = 2^8 \cdot 3^2 \cdot 5 = 11520$
$M(6) = 2^{10} \cdot 3^4 \cdot 5 \cdot 7 = 2903040$
$M(7) = 2^{11} \cdot 3^4 \cdot 5 \cdot 7 = 5806080$
$M(8) = 2^{15} \cdot 3^5 \cdot 5^2 \cdot 7 = 1393459200.$

Note that

$$M(n)/M(n-1) = \begin{cases} 2 & \text{if } n \text{ is odd} \\ \text{denominator of } b_n/n & \text{if } n \text{ is even,} \end{cases}$$

where b_n is the n-th Bernoulli number. (The occurence of the Bernoulli numbers is natural in view of the mass formulae which Minkowski had proved a few years before.)

2) One may ask whether there is a finite subgroup A of $\mathbf{GL}_n(\mathbf{Q})$ of order $M(n)$. It is so for $n = 1$ and $n = 3$ and probably for no other value of n (as Burnside already remarked on p.484 of [Bu 11]). Indeed, some incomplete arguments of Weisfeiler and Feit would imply that the upper bound of $|A|$ is $2^n \cdot n!$ if $n > 10$, which is much smaller than $M(n)$. See the comments of Guralnick-Lorenz in [GL 06], §6.1.

Exercise. Let $\left[\frac{n}{\ell-1}\right] = \sum a_i \ell^i, 0 \leq a_i \leq \ell - 1$, be the ℓ-adic expansion of $\left[\frac{n}{\ell-1}\right]$. Show that $M(n,\ell) = \sum a_i \frac{\ell^{i+1}-1}{\ell-1} = \sum M(a_i \ell^i(\ell-1), \ell)$.

1.2. **Minkowski's lemma.** Minkowski's paper starts with the following often quoted lemma:

Lemma 1. *If $m \geq 3$, the kernel of $\mathbf{GL}_n(\mathbf{Z}) \to \mathbf{GL}_n(\mathbf{Z}/m\mathbf{Z})$ is torsion free.*

Proof. Easy exercise ! One may deduce it from general results on formal groups over local rings, cf. Bourbaki [LIE III], §7. Many variants exist. For instance:

Lemma 1′. *Let R be a local ring with maximal ideal $\mathfrak{m}$ and residue field $k = R/\mathfrak{m}$. If ℓ is a prime number distinct from $\mathrm{char}(k)$, the kernel of the map $\mathbf{GL}_n(R) \to \mathbf{GL}_n(k)$ does not contain any element of order ℓ.*

Proof. Suppose $x \in \mathbf{GL}_n(R)$ has order ℓ and gives 1 in $\mathbf{GL}_n(k)$. Write $x = 1 + y$; all the coefficients of the matrix y belong to $\mathfrak{m}$. Since $x^\ell = 1$, we have

$$\ell \cdot y + \binom{\ell}{2} \cdot y^2 + \cdots + \ell \cdot y^{\ell-1} + y^\ell = 0,$$

which we may write as $y \cdot u = 0$, with $u = \ell + \binom{\ell}{2} y + \cdots + y^{\ell-1}$. The image of u in $\mathbf{GL}_n(k)$ is ℓ, which is invertible. Hence u is invertible, and since $y \cdot u$ is 0, this shows that $y = 0$. □

Several other variants can be found in [SZ 96].

Remark. A nice consequence of lemma 1′ is the following result of Malcev and Selberg ([Bo 69], §17):

(*) *Let Γ be a finitely generated subgroup of $\mathbf{GL}_n(K)$, where K is a field of characteristic 0. Then Γ has a torsion free subgroup of finite index.*

Sketch of proof (for more details, see Borel, *loc.cit.*). Let S be a finite generating subset of Γ, and let L be the ring generated by the coefficients of the elements of $S \cup S^{-1}$. We have $\Gamma \subset \mathbf{GL}_n(L)$. Let $\mathfrak{m}$ be a maximal ideal of L; the residue field $k = A/\mathfrak{m}$ is finite ([AC V], p.68, cor.1 to th.3); let p be its characteristic. The kernel Γ_1 of $\Gamma \to \mathbf{GL}_n(k)$ has finite index in Γ; by lemma 1′ (applied to the local ring $R = L_\mathfrak{m}$), Γ_1 does not have any torsion except possibly p-torsion. By choosing another maximal ideal of L, with a different residue characteristic, one gets a torsion free subgroup of finite index of Γ_1, and hence of Γ. □

Remark. When K has characteristic $p > 0$ the same proof shows that Γ has a subgroup of finite index which is "p'-torsion free", i.e. such that its elements of finite order have order a power of p.

1.3. **Proof of theorem 1 (i).** Let A be a finite subgroup of $\mathbf{GL}_n(\mathbf{Q})$; we have to show that $v_\ell(A) \leq M(n, \ell)$. Note first:

1.3.1. *The group A is conjugate to a subgroup of $\mathbf{GL}_n(\mathbf{Z})$.*

This amounts to saying that there exists an A-stable lattice in $\mathbf{Q}^n$, which is clear: just take the lattice generated by the A-transforms of the standard lattice $\mathbf{Z}^n$.

1.3.2. *There is a positive definite quadratic form on $\mathbf{Q}^n$, with integral coefficients, which is invariant by A.*

Same argument: take the sum of the A-transforms of $x_1^2 + \cdots + x_n^2$, and multiply it by a suitable non-zero integer, in order to cancel any denominator.

Let us now proceed with the proof of $v_\ell(A) \leq M(n,\ell)$. We do it in two steps:

1.3.3. *The case $\ell > 2$.*

By 1.3.1, we may assume that A is contained in $\mathbf{GL}_n(\mathbf{Z})$. Let p be a prime number $\neq 2$. By lemma 1, the map $A \to \mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is injective. Hence

$$v_\ell(A) \leq a(p) = v_\ell\big(\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})\big).$$

The order of $\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is $p^{n(n-1)/2}(p-1)(p^2-1)\dots(p^n-1)$. Let us assume that $p \neq \ell$. Then we have

$$a(p) = \sum_{i=1}^{n} v_\ell(p^i - 1).$$

We now choose p in such a way that $a(p)$ is as small as possible. More precisely, we choose p such that:

(*) *The image of p in $(\mathbf{Z}/\ell^2\mathbf{Z})^*$ is a generator of that group.*

This is possible by Dirichlet's theorem on the existence of primes in arithmetic progressions (of course, one should also observe that $(\mathbf{Z}/\ell^2\mathbf{Z})^*$ is cyclic.)

Once p is chosen in that way, then $p^i - 1$ is divisible by ℓ only if i is divisible by $\ell - 1$; moreover, one has $v_\ell(p^{\ell-1} - 1) = 1$ because of (*), and this implies that $v_\ell(p^i - 1) = 1 + v_\ell(i)$ if i is divisible by $\ell - 1$. (This is where the hypothesis $\ell > 2$ is used.) One can then compute $a(p)$ by the formula above. The number of indices $i \leq n$ which are divisible by $\ell - 1$ is $\left[\frac{n}{\ell-1}\right]$. We thus get:

$$\begin{aligned} a(p) &= \left[\frac{n}{\ell-1}\right] + \sum_{1 \leq j \leq \left[\frac{n}{\ell-1}\right]} v_\ell(j) = \left[\frac{n}{\ell-1}\right] + v_\ell\Big(\left[\frac{n}{\ell-1}\right]!\Big) \\ &= \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \cdots = M(n,\ell). \end{aligned}$$

This proves th.1 (i) in the case $\ell \neq 2$.

1.3.4. *The case $\ell = 2$.*

When $\ell = 2$, the method above does not give the right bound as soon as $n > 1$. One needs to replace $\mathbf{GL}_n$ by an orthogonal group. Indeed, by 1.3.1 and 1.3.2, we may assume, not only that A is contained in $\mathbf{GL}_n(\mathbf{Z})$, but also that it is contained in the orthogonal group $\mathbf{O}_n(q)$, where q is a non-degenerate quadratic form with integral coefficients. Let D be the discriminant of q, and let us choose a prime number $p > 2$ which does not divide D. The image of A in $\mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$ is contained in the orthogonal

group $\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})$ relative to the reduction of q mod p. If we put $r = [n/2]$, the order of $\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})$ is known to be:

$$2 \cdot p^{r^2}(p^2-1)(p^4-1)\dots(p^{2r}-1) \quad \text{if } n \text{ is odd.}$$

and

$$2 \cdot p^{r(r-1)}(p^2-1)(p^4-1)\dots(p^{2r}-1)/(p^r+\varepsilon) \quad \text{if } n \text{ is even,}$$

with $\varepsilon = \pm 1$ equal to the Legendre symbol at p of $(-1)^r D$.

If we choose $p \equiv \pm 3 \pmod 8$, we have $v_2(p^{2i}-1) = 3 + v_2(i)$, and $v_2(p^r+\varepsilon) \geq 1$. If n is odd, this gives

$$v_2(\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})) = 1 + 3r + v_2(r!) = n + r + \left[\frac{r}{2}\right] + \left[\frac{r}{4}\right] + \dots = M(n,2),$$

and, if n is even:

$$v_2(\mathbf{O}_n(\mathbf{Z}/p\mathbf{Z})) \leq 3r + v_2(r!) = M(n,2).$$

Hence $v_2(A)$ is at most equal to $M(n,2)$. □

Remark. There are several ways of writing down this proof. For instance:

- There is no need to embed A in $\mathbf{GL}_n(\mathbf{Z})$. It sits in $\mathbf{GL}_n(\mathbf{Z}[1/N])$ for a suitable $N \geq 1$, and this allows us to reduce mod p for all p's not dividing N.

- Minkowski's lemma is not needed either: we could replace it by the trivial fact that a matrix which is different from 1 is not congruent to 1 mod p for all large enough p's.

- Even when $\ell > 2$, we could have worked in $\mathbf{O}_n$ instead of $\mathbf{GL}_n$; that is what Minkowski does.

- When $\ell = 2$ the case n even can be reduced to the case n odd by observing that, if $A \subset \mathbf{GL}_n(\mathbf{Q})$, then $A \times \{\pm 1\}$ embeds into $\mathbf{GL}_{n+1}(\mathbf{Q})$, and $M(n+1,2)$ is equal to $1 + M(n,2)$.

1.4. **Proof of theorem 1 (ii).** The symmetric group S_ℓ has a faithful representation $S_\ell \to \mathbf{GL}(V_1)$ where V_1 is a $\mathbf{Q}$-vector space of dimension $\ell - 1$. Put $r = \left[\frac{n}{\ell-1}\right]$, and let $V = V_1 \oplus \dots \oplus V_r$ be the direct sum of r copies of V_1. Let S be the semi-direct product of S_r with the product $(S_\ell)^r$ of r copies of S_ℓ ("wreath product"). The group S has a natural, and faithful, action on V. We may thus view S as a subgroup of $\mathbf{GL}_{r(\ell-1)}(\mathbf{Q})$, hence also of $\mathbf{GL}_n(\mathbf{Q})$, since $n \geq r(\ell-1)$. We have

$$v_\ell(S) = r + v_\ell(r!) = \left[\frac{n}{\ell-1}\right] + \left[\frac{n}{\ell(\ell-1)}\right] + \dots = M(n,\ell).$$

An ℓ-Sylow A of S satisfies the conditions of th.1 (ii). □

Example. When $\ell = 2$ the group S defined above is the "hyper-octahedral group", i.e. the group of automorphisms of an n-cube (= the Weyl group of a root system of type B_n); in ATLAS notation, it may be written as $2^n \cdot S_n$.

1.5. **A conjugacy theorem.** The finite ℓ-subgroups of $\mathbf{GL}_n(\mathbf{Q})$ have the following Sylow-like property:

Theorem 1′. *Let A and A' be two finite ℓ-subgroups of $\mathbf{GL}_n(\mathbf{Q})$. Assume that A has the maximal order allowed by th.*1*. Then A' is conjugate to a subgroup of A.*

Corollary. *If $|A| = |A'| = \ell^{M(n,\ell)}$, then A and A' are conjugate in $\mathbf{GL}_n(\mathbf{Q})$.*

Proof of theorem 1′. See Bourbaki, [LIE III], §7, exerc.6 f) where only the case $\ell > 2$ is given, and Feit [Fe 97] who does the case $\ell = 2$. Let us sketch Bourbaki's method (which we shall use in §6.6 in a more general setting):

We may assume that A and A' are contained in $\mathbf{GL}_n(\mathbf{Z})$. Choose a prime p as in 1.3.3, and reduce mod p. The groups A and A' then become ℓ-subgroups of $G_p = \mathbf{GL}_n(\mathbf{Z}/p\mathbf{Z})$, and A is an ℓ-Sylow of G_p. By Sylow's theorem applied to G_p, one finds an injection $i : A' \to A$ which is induced by an inner automorphism of G_p. The two linear representations of A':

$$A' \to \mathbf{GL}_n(\mathbf{Q}) \quad \text{and} \quad A' \xrightarrow{i} A \to \mathbf{GL}_n(\mathbf{Q})$$

become isomorphic after reduction mod p. Since $p \neq \ell$, a standard argument shows that they are isomorphic over $\mathbf{Q}$, which proves th.1′ in that case. The case $\ell = 2$ can be handled by a similar, but more complicated, argument: if n is odd, one uses orthogonal groups as in 1.3.4, and one reduces the case n even to the case n odd by the trick mentioned at the end of §1.3. □

Exercise. Let $A(n)$ be a maximal 2-subgroup of $\mathbf{GL}_n(\mathbf{Q})$. Show that the $A(n)$'s can be characterized by the following three properties:

$$\begin{aligned} A(1) &= \{\pm 1\}. \\ A(2n) &= \big(A(n) \times A(n)\big) \cdot \{\pm 1\} \text{ (wreath product) if } n \text{ is a power of 2.} \\ A(n) &= A(2^{m_1}) \times \cdots \times A(2^{m_k}) \text{ if } n = 2^{m_1} + \cdots + 2^{m_k} \text{ with } m_1 < \cdots < m_k. \end{aligned}$$

§2. Schur

Ten years after [Mi 87], Frobenius founded the theory of characters of finite groups. It was then (and still is now) very tempting to use that theory to give a different proof of Minkowski's results. The first people to do so were Schur ([Sch 05]) and Burnside ([Bu 11], Note G). Schur's paper is especially interesting. He works first over $\mathbf{Q}$, as Minkowski did, and uses a very original argument in character theory, see §2.1 below. He then attacks the case of an arbitrary number field, where he gets a complete answer, see §2.2.

2.1. **Finite linear groups with rational trace.** What Schur proves in §1 of [Sch 05] is:

Theorem 2. *Let A be a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$. Assume that the traces of the elements of A lie in $\mathbf{Q}$. Then $v_\ell(A) \leq M(n,\ell)$, where $M(n,\ell)$ is as in th.*1*.*

The condition on the traces is obviously satisfied if A is contained in $\mathbf{GL}_n(\mathbf{Q})$. Hence th.2 is a generalization of th.1. (As a matter of fact, it is a genuine generalization only when $\ell = 2$; indeed, when $\ell > 2$, it is known, cf. [Ro 58], that a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$ with rational trace is conjugate to a subgroup of $\mathbf{GL}_n(\mathbf{Q})$.)

Proof. We start from the following general fact, which is implicit in [Sch 05] (and is sometimes called "Blichfeldt's lemma"):

Proposition 1. *Let G be a finite subgroup of $\mathbf{GL}_n(\mathbf{C})$ and let X be the subset of $\mathbf{C}$ made up of the elements $\mathrm{Tr}(g)$ for $g \in G, g \neq 1$. Let $N = \prod(n - x)$ be the product of the $n - x$, for $x \in X$. Then N is a non-zero integer which is divisible by $|G|$.*

(Hence the knowledge of the set X gives a multiplicative bound for the order of G.)

Proof. Let $m = |G|$, and let z be a primitive m-th root of unity. The elements of X are sums of powers of z; hence they belong to the ring of integers of the cyclotomic field $K = \mathbf{Q}(z)$. This already shows that N is an algebraic integer. If s is an element of $\mathrm{Gal}(K/\mathbf{Q})$, one has $s(z) = z^a$ for some $a \in (\mathbf{Z}/m\mathbf{Z})^*$. If $x = \mathrm{Tr}(g)$, with $g \in G$, then $s(x) = \mathrm{Tr}(g^a)$, hence $s(x)$ belongs to X. This shows that X is stable under the action of $\mathrm{Gal}(K/\mathbf{Q})$; hence N is fixed by $\mathrm{Gal}(K/\mathbf{Q})$; this proves that N belongs to $\mathbf{Z}$.

The factors of N are $\neq 0$. Indeed, $\mathrm{Tr}(g)$ is equal to the sum of n complex numbers z_i with $|z_i| = 1$, hence can be equal to n only if all the z_i are equal to 1, which is impossible since $g \neq 1$. This shows that $N \neq 0$ (one could also prove that N is positive, but we shall not need it).

It remains to see that N is divisible by $|G|$. It is well-known that, if χ is a generalized character of G, the sum $\sum_{g\in G} \chi(g)$ is divisible by $|G|$. Let us apply this to the function $g \mapsto \chi(g) = \prod_{x\in X} \big(\mathrm{Tr}(g) - x\big)$, which is a $\mathbf{Z}$-linear combination of the characters $g \mapsto \mathrm{Tr}(g)^m, m \geq 0$. Since $\chi(g) = 0$ for $g \neq 1$ and $\chi(1) = N$, the sum of the $\chi(g)$ is equal to N. Hence N is divisible by $|G|$. □

The next lemma gives an information on the $\mathrm{Tr}(g)$'s:

Lemma 2. *Let A be as in th.2. If $g \in A$, then $\mathrm{Tr}(g)$ may be written as $n - \ell y$ with $y \in \mathbf{Z}$ and $0 \leq y \leq n/(\ell - 1)$.*

Proof. Each eigenvalue of g is of order ℓ^α for some $\alpha \geq 0$, and all the eigenvalues with the same α have the same multiplicity. By splitting $\mathbf{C}^n$ according to the α's, one is reduced to the following three cases:
(1) $g = 1$ and $n = 1$. Here $\mathrm{Tr}(g) = 1$ and we take $y = 0$.
(2) g has order ℓ and $n = \ell - 1$. Here $\mathrm{Tr}(g) = -1$, and $y = 1$.
(3) g has order ℓ^α with $\alpha > 1$ and $n = \ell^{\alpha-1}(\ell - 1)$. Here $\mathrm{Tr}(g) = 0$ and $y = \ell^{\alpha-2}(\ell - 1)$.
In each case we have $0 \leq y \leq n/(\ell - 1)$. □

End of the proof of theorem 2. We apply prop.1 to $G = A$. By lemma 2, each factor $n - x$ of N can be written as ℓy with $1 \leq y \leq d = [n/\ell - 1)]$.

This shows that N divides the product $\ell^d \cdot d!$ and we have

$$v_\ell(N) < d + v_\ell(d!) = [n/(\ell-1)] + [n/\ell(\ell-1)] + \cdots = M(n,\ell).$$

Since $|G|$ divides N, this proves th.2. □

Remark. One may ask whether th.2 can be complemented by a conjugacy theorem analogous to th.1′ of §1.5. The answer is of course "yes" if $\ell > 2$ (because of th.1′), but it is "no" for $\ell = 2$: the dihedral group D_4 and the quaternion group Q_8 are non-conjugate 2-subgroups of $\mathbf{GL}_2(\mathbf{C})$, with rational trace, which have the maximal order allowed by th.2, namely 8.

2.2. **Replacing Q by an arbitrary number field.** This is what Schur does in §§2-6 of [Sch 05]. Before stating his result, some notation is necessary:

- k is a number field, viewed as a subfield of $\mathbf{C}$.
- For each $a \geq 1$, z_a denotes a primitive a-th root of unity.
- (assuming $\ell > 2$). We put $t = [k(z_\ell) : k]$ and we denote by m the maximal a such that $k(z_\ell)$ contains $z_{\ell a}$ (this notation coincides with Schur's, and it will be extended to arbitrary fields in §4 of Lect.II). We put

$$M_k(n,\ell) = m \cdot \left[\frac{n}{t}\right] + \left[\frac{n}{\ell t}\right] + \left[\frac{n}{\ell^2 t}\right] + \dots$$

- (assuming $\ell = 2$). We put $t = [k(i) : k]$ and we define m as explained in §4.2 (warning: t and m do not always coincide with Schur's t_2 and m_2). We put:

$$M_k(n,2) = n + (m'-1)\left[\frac{n}{t}\right] + \left[\frac{n}{2t}\right] + \left[\frac{n}{4t}\right] + \dots,$$

where m' is equal to $m+1$ in case (b) of §4.2 and is equal to m in the other cases.

The main result of [Sch 05] is:

Theorem 2′. *Let A be a finite ℓ-subgroup of $\mathbf{GL}_n(\mathbf{C})$ such that $\mathrm{Tr}(g)$ belongs to k for every $g \in A$. Then $v_\ell(A) \leq M_k(n,\ell)$.*

Note that, when $k = \mathbf{Q}$, the integer $M_k(n,\ell)$ is equal to Minkowski's $M(n,\ell)$; hence th.2′ is a generalization of th.2.

Proof. I shall not give all the details of Schur's proof, but just explain its main steps. For more information, see [Sch 05] (and also [GL 06] for the case $\ell > 2$).

One of the inputs of the proof is the following result, which had just been proved by Blichfeldt ([Bl 04] - see also §3 below):

2.2.1. *Every linear representation of A is monomial.*

Hence one can decompose the vector space $\mathbf{C}^n$ as a direct sum of n lines $D_1, \dots, D_n$ which are permuted by A. This gives a homomorphism $A \to S_n$; its kernel A' is a normal abelian subgroup of A. Hence:

2.2.2. *The group A has a normal abelian subgroup A' such that $(A : A')$ divides $n!$.*

This led Schur to investigate the case where A is abelian. He proved:

2.2.3. *If A is as in th.2′, and is abelian, then* :

$$v_\ell(A) \leq \begin{cases} m \cdot \left[\frac{n}{t}\right] & \text{if } \ell > 2 \\ (m'-t) \cdot \left[\frac{n}{t}\right] + n & \text{if } \ell = 2. \end{cases}$$

Sketch of proof. Since A is abelian, and the traces of its elements belong to k, it is conjugate to a subgroup of $\mathbf{GL}_n(k)$. Let R be the k-subalgebra of $\mathbf{M}_n(k)$ generated by A. We may write R as a product $\prod K_i$, where the K_i are cyclotomic extensions of k, of the form $k(z_{\ell^{a_i}})$, with $a_i \geq 0$. Let $n_i = [K_i : k]$; then $\sum n_i \leq n$. The image of A in K_i^* is a cyclic group of order ℓ^{a_i}. If $\ell > 2$, it is not difficult to see that $a_i \leq m \cdot \left[\frac{n_i}{t}\right]$ for every i. Adding up, we find $\sum a_i \leq m \cdot \left[\frac{n}{t}\right]$, and since $v_\ell(A) \leq \sum a_i$, we get the inequality (2.2.3). The case $\ell = 2$ is similar. □

Once this is done, the case $\ell = 2$ follows. Indeed (2.2.2) and (2.2.3) give $v_2(A) \leq v_2(A') + v_2(n!) \leq n + (m'-t) \cdot \left[\frac{n}{t}\right] + v_2(n!)$, and this is equivalent to $v_2(A) \leq M_k(n, 2)$. The case $\ell > 2$ requires more work, cf. [Sch 05], §5. □

Remarks

1) The bound $v_\ell(A) \leq M_k(n, \ell)$ is *optimal*; this is proved by the same explicit constructions as in §1.4, cf. [Sch 05], §6.

2) As we already pointed out in §2.1, the hypothesis $\mathrm{Tr}(A) \subset k$ implies, when $\ell > 2$, that A is conjugate to a subgroup of $\mathbf{GL}_n(k)$. One may then use Minkowski's method, as will be explained in §6 for semisimple algebraic groups (of course $\mathbf{GL}_n$ is not semisimple, but the method applies with almost no change – the invariant degrees d_i of §6 have to be replaced by $1, 2, \ldots, n$). The bound found in that way coincides with Schur's.

For $\ell = 2$, if one does not assume that A can be embedded in $\mathbf{GL}_n(k)$, I do not see how to apply either Minkowski's method or the cohomological method of §6.8. This raises interesting questions. For instance, consider a finite subgroup A of $E_8(\mathbf{C})$, and suppose that the conjugacy classes of the elements of A are $\mathbf{Q}$-rational. Is it true that $v_2(A) \leq 30$, $v_3(A) \leq 13, \ldots$, as would be the case if A were contained in the rational points of a $\mathbf{Q}$-form of E_8, cf. §6.3.2 ?

§3. Blichfeldt and others

Blichfeldt's theorem (§3.1 below) has already been used in §2.2. The results of §3.3 will be applied in §5.4, in order to prove what I call the "S-bound".

3.1. **Blichfeldt's theorem.** Recall that a finite group A is called *supersolvable* if it has a composition series

$$1 = A_0 \subset A_1 \subset \cdots \subset A_m = A$$

where the A_i are normal in A (and not merely in A_{i+1}) and the quotients A_i/A_{i-1} are cyclic. One has

$$\text{nilpotent} \Rightarrow \text{supersolvable} \Rightarrow \text{solvable}.$$

In particular, an ℓ-group is supersolvable.

One proves easily:

(*) *If A is supersolvable and non abelian, there exists an abelian normal subgroup A' of A which is not contained in the center of A.*

Recall also that a linear representation V of a group A is called *monomial* if one can split V as a direct sum of lines which are permuted by A. When V is irreducible, this amounts to saying that V is induced by a 1-dimensional representation of a subgroup of A.

We may now state Blichfeldt's theorem ([Bl 04], see also [Bu 11], §258):

Theorem 3. *Every complex linear representation of a supersolvable finite group is monomial.*

(As a matter of fact, Blichfeldt was only interested in the case where A is nilpotent.)

Proof. The argument is now standard. We may assume that the given representation V is irreducible and faithful. If A is abelian, we have $\dim V = 1$ and there is nothing to prove. If not, we choose A' as in (*) above, and we split V as $V = \oplus V_\chi$, where χ runs through the 1-dimensional characters of A', and V_χ is the corresponding eigenspace; let V_ψ be a non-zero V_χ; it is distinct from V (otherwise, A' would be central), and every non-zero V_χ is an A-transform of V_ψ (because V is irreducible). Call B the subgroup of A stabilizing V_ψ. We have $A' \subset B \subset A$, and $|B| < |A|$. Using induction on $|A|$, we may assume that th.3 is true for B; this gives a splitting of V_ψ as a direct sum of lines which are stable under B. By transforming them by A, we get the desired splitting of V. □

3.2. **Borel-Serre.** In [BS 53], Borel and I proved:

Theorem 3′. *Let G be a compact real Lie group, and let A be a finite supersolvable subgroup of G. There exists a maximal torus T of G which is normalized by A.*

Remark. When one applies th.3′ to $G = \mathbf{U}_n(\mathbf{C})$, one recovers th.3. Hence th.3′ may be viewed as a generalization of Blichfeldt's theorem.

Proof of theorem 3′ (sketch).

Lemma 3. *Let $\mathfrak{g}$ be a finite dimensional Lie algebra over a field of characteristic 0, and let s be an automorphism of prime order of $\mathfrak{g}$. If s has no fixed point $\neq 0$, then $\mathfrak{g}$ is nilpotent.*

(Note the analogy with a - much deeper - theorem of Thompson [Th 60-64]: if a finite group G has an automorphism of prime order with no non-trivial fixed point, then G is nilpotent.)

Proof of lemma 3. By extending scalars, we may assume that the ground field is algebraically closed. Let p be the order of s, and let z be a primitive p-th root of unity. Let $\mathfrak{g}_i$ be the kernel of $s - z^i$ in $\mathfrak{g}$. We have

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_{p-1}\,,$$

and the hypothesis made on s means that $\mathfrak{g}_0 = 0$. One then shows that $\mathrm{ad}(x)^{p-1} = 0$ for every x belonging to one of the $\mathfrak{g}_i$'s. This implies that the

Killing form of $\mathfrak{g}$ is 0, hence that $\mathfrak{g}$ is solvable (Cartan's criterion). The fact that $\mathfrak{g}$ is nilpotent follows easily. (For more details, see §4 of [BS 54].) □

Once this is done, th.3′ is proved by an induction argument similar to the one used in the proof of Blichfeldt's theorem, cf. [BS 53], §3. □

3.3. **Steinberg and Springer-Steinberg.** We now come to the setting of linear algebraic groups. Let k be a field, and let G be an algebraic group over k. We shall assume in what follows that G is linear and smooth over k; the connected component of the identity of G is denoted by G°. Recall that G is said to be *reductive* if it is connected and if its unipotent radical (over an algebraic closure of k) is trivial, cf. [Bo 91], §11.21. If $k = \mathbf{C}$, such groups correspond (by a standard dictionary, cf. [Se 93], §5) to the connected compact Lie groups. [In the literature, a group G such that G° is reductive is sometimes called "reductive"; this is reasonable in characteristic 0, but not otherwise. Here we prefer that "reductive" implies "connected".]

Theorem 3′ has the following analogue:

Theorem 3″. *Let A be a finite supersolvable group of order prime to* $\mathrm{char}(k)$ *and let G be a reductive group over k on which A acts by k-automorphisms. Then there exists a maximal torus T of G, defined over k, which is stable under the action of A.*

(When $k = \mathbf{C}$, this is equivalent to th.3′, thanks to the dictionary mentioned above.)

Corollary. *If A is a finite supersolvable subgroup of $G(k)$, of order prime to* $\mathrm{char}(k)$*, there is a maximal k-torus T of G whose normalizer N is such that A is contained in $N(k)$.*

(Recall that, if X is a k-variety, $X(k)$ is the set of k-points of X.)

Proof of theorem 3″. When k is algebraically closed, this is proved in [SS 68], I.5.16, with the help of several results from [St 68]. For an arbitrary field k, the same proof works with very little change. One starts with the following basic result of Steinberg ([St 68], th.7.2):

Proposition 2. *Assume k is algebraically closed. Let $s : G \to G$ be a surjective homomorphism. Then there exists a Borel subgroup B of G such that $s(B) = B$.*

When s has finite order prime to $\mathrm{char}(k)$, one can say much more:

Proposition 3. *Let s be an automorphism of G of finite order prime to* $\mathrm{char}(k)$*, and let G^s be the subgroup of G fixed by s. Then*:

a) *The connected component of G^s is reductive.*

b) *One has* $\dim G^s > 0$ *if G is not a torus.*

c) *If k is algebraically closed, there exists a Borel subgroup B of G and a maximal torus T of B such that $s(B) = B$ and $s(T) = T$.*

Proof (*sketch*). We may assume k is algebraically closed, since assertions a) and b) are "geometric". A proof of a) is given in [St 68], cor.9.4. A proof of c) is given in [SS 68], I.2.9, as an application of prop.2. Assertion b) follows from c) by the following method of Steinberg: one observes that a pair (B, T)

with $B \supset T$, determines *canonically* a homomorphism $h : \mathbf{G}_m \to T$ (indeed B gives a basis of the root system of (G, T), and one takes for h twice the sum of the corresponding coroots). Moreover, h is non-trivial if G is not a torus. The canonicity of h implies that it is fixed by s. Hence G^s contains $\mathrm{Im}(h)$. □

End of the proof of th.3″. By induction on $|A| + \dim G$. When $A = 1$, one takes for T any maximal k-torus of G; by a theorem of Grothendieck, there is such a torus (cf. [Bo 91], th.18.2). We may thus assume $A \neq 1$. In that case A contains a cyclic subgroup $< s >$, non-trivial, which is normal. We may also assume that G is semisimple and that A acts faithfully. Let G_1 be the connected component of G^s; we have $\dim G_1 > 0$, cf. prop.3 b). The group A/A' acts on G_1. By the induction assumption, there is a maximal torus T_1 of G_1, defined over k, which is stable under the action of A/A', hence of A. Let G_2 be the centralizer of T_1 in G. It is a reductive group of the same rank as G. We have $\dim G_2 < \dim G$, since T_1 is not contained in the center of G. Moreover, G_2 is stable under the action by A. By applying the induction assumption to the pair (G_2, A) we get a maximal k-torus T of G_2 which is A-stable. Since G_2 and G have the same rank, T is a maximal torus of G. □

II. Upper bounds

Let G be a reductive group over a field k, and let ℓ be a prime number, different from $\mathrm{char}(k)$. Let A be a finite subgroup of $G(k)$. We want to give an upper bound for $v_\ell(A)$, in terms of invariants of G, k and ℓ. We give two such bounds. The first one (§5) is less precise, but very easy to apply; we call it the S-bound (S for Schur). The other bound (§6) is the M-bound (M for Minkowski). Both bounds involve some cyclotomic invariants of k, which are defined in §4 below.

§4. The invariants t and m

4.0. **Cyclotomic characters.** Let $\bar{k}$ be an algebraic closure of k, and let k_s be the separable closure of k in $\bar{k}$. For each $n \geq 1$ prime to $\mathrm{char}(k)$, let $\mu_n \subset k_s^*$ be the group of n-th roots of unity and let z_n be a generator of μ_n.

The Galois group $\Gamma_k = \mathrm{Gal}(k_s/k)$ acts on $\langle z_n \rangle = \mu_n$. This action defines a continuous homomorphism

$$\chi_n : \Gamma_k \to \mathrm{Aut}(\mu_n) = (\mathbf{Z}/n\mathbf{Z})^*,$$

which is called the *n-th cyclotomic character of k*.

This applies in particular to $n = \ell^d$ $(d = 0, 1, \dots)$; by taking inverse limits we get the ℓ^∞-cyclotomic character

$$\chi_{\ell^\infty} : \Gamma_k \to \mathbf{Z}_\ell^* = \varprojlim (\mathbf{Z}/\ell^d\mathbf{Z})^*,$$

where $\mathbf{Z}_\ell$ is the ring of ℓ-adic integers. What matters for us is the image $\mathrm{Im}\ \chi_{\ell^\infty}$, which is a closed subgroup of $\mathbf{Z}_\ell^*$. To discuss its structure, it is convenient to separate the cases $\ell \neq 2$ and $\ell = 2$.

4.1. **The case $\ell \neq 2$.** We have

$$\mathbf{Z}_\ell^* = C_{\ell-1} \times \{1 + \ell \cdot \mathbf{Z}_\ell\}$$

where $C_{\ell-1}$ is cyclic of order $\ell - 1$ (i.e. $C_{\ell-1}$ is the group $\mu_{\ell-1}$ of the ℓ-adic field $\mathbf{Q}_\ell$; it is canonically isomorphic to $\mathbf{F}_\ell^*$). As for $1 + \ell \cdot \mathbf{Z}_\ell$, it is procyclic, generated by $1 + \ell$, and isomorphic to the additive group $\mathbf{Z}_\ell$; its closed subgroups are the groups $1 + \ell^d \cdot \mathbf{Z}_\ell = \langle 1 + \ell^d \rangle$, $d = 1, 2, \ldots, \infty$, with the convention $\ell^\infty = 0$.

Since $\ell - 1$ and ℓ are relatively prime, the subgroup $\operatorname{Im} \chi_{\ell^\infty}$ of $\mathbf{Z}_\ell^*$ decomposes as a direct product:

$$\operatorname{Im} \chi_{\ell^\infty} = C_t \times \{1 + \ell^m \cdot \mathbf{Z}_\ell\} \ ,$$

where t is a divisor of $\ell - 1$, C_t is cyclic of order t and $m = 1, 2, \ldots$ or ∞.

Remark. An alternative definition of the invariants t and m is:

$$\begin{aligned} t &= [k(z_\ell) : k] = k\text{-degree of } z_\ell \\ m &= \text{upper bound of the } d \geq 1 \text{ such that } z_{\ell^d} \text{ is contained in } k(z_\ell). \end{aligned}$$

Examples. If $k = \mathbf{Q}$ or $\mathbf{Q}_\ell$, χ_{ℓ^∞} is surjective and we have $t = \ell - 1$, $m = 1$. If $k = k_s$, then χ_{ℓ^∞} is trivial and $t = 1$, $m = \infty$. If k is finite with q elements, $\operatorname{Im} \chi_{\ell^\infty}$ is the closed subgroup of $\mathbf{Z}_\ell^*$ generated by q and we have:

$$\begin{aligned} t &= \text{order of } q \text{ in } \mathbf{F}_\ell^* \\ m &= v_\ell(q^t - 1) = v_\ell\left(q^{\ell-1} - 1\right). \end{aligned}$$

4.2. **The case $\ell = 2$.** Here $\mathbf{Z}_2^* = C_2 \times \{1 + 4 \cdot \mathbf{Z}_2\}$, where $C_2 = \{1, -1\}$ and the multiplicative group $1 + 4 \cdot \mathbf{Z}_2$ is isomorphic to the additive group $\mathbf{Z}_2$. There are three possibilities for $\operatorname{Im} \chi_{2^\infty}$:

(a) $\operatorname{Im} \chi_{2^\infty} = 1 + 2^m \cdot \mathbf{Z}_2 = \langle 1 + 2^m \rangle$, with $m = 2, \ldots, \infty$. We put $t = 1$.
(b) $\operatorname{Im} \chi_{2^\infty} = \langle -1 + 2^m \rangle$, with $m = 2, \ldots, \infty$. We put $t = 2$.
(c) $\operatorname{Im} \chi_{2^\infty} = C_2 \times \{1 + 2^m \cdot \mathbf{Z}_2\} = \langle -1, 1 + 2^m \rangle$, $m = 2, \ldots, \infty$. We put $t = 2$.

If $m < \infty$, these types are distinct. If $m = \infty$, types (b) and (c) coincide; in that case $\operatorname{Im} \chi_{2^\infty}$ is equal to C_2.

Remark. We have $t = [k(i) : k]$ with the usual notation $i = z_4$. Hence case (a) means that -1 is a square in k, and in that case m is the largest $d \geq 2$ such that $z_{2^d} \in k$.

If $t = 2$, case (c) is characterized by the fact that -1 belongs to $\operatorname{Im} \chi_{2^\infty}$. As for m, it is given by:

$$\begin{aligned} m &= -1 + \text{ upper bound of the } d \geq 2 \text{ such that } z_{2^d} \in k(i) \text{ in case (b)} \\ m &= \text{upper bound of the } d \geq 2 \text{ such that } z_{2^d} \in k(i) \text{ in case (c)}. \end{aligned}$$

Examples. If $k = \mathbf{Q}$ or $\mathbf{Q}_2$, we have type (c) with $t = 2, m = 2$. If $k = \mathbf{R}$, we have types (b) and (c) with $m = \infty$. If k is separably closed, we have type (a) with $t = 1$ and $m = \infty$.

When $\operatorname{char}(k) \neq 0$, type (c) is impossible unless $m = \infty$. If k is finite with q elements, we have type (a) with $m = v_2(q - 1)$ if $q \equiv 1 \pmod 4$ and type (b) with $m = v_2(q + 1)$ if $q \equiv -1 \pmod 4$.

4.3. **The case of finitely generated fields.** Let k_0 be the prime subfield of k, i.e. $\mathbf{Q}$ if $\mathrm{char}(k) = 0$ or $\mathbf{F}_p$ if $\mathrm{char}(k) = p > 0$. Suppose that k is *finitely generated over* k_0. Then *the invariant* m *is finite*, i.e. $\mathrm{Im}\,\chi_{\ell^\infty}$ is infinite.

Indeed, if not, there would be a finite extension k' of k containing the group μ of all the ℓ^d-th roots of unity ($d = 1, 2, \dots$). Let $K = k_0(\mu)$ be the extension of k_0 generated by μ. Then:

(a) K is algebraic over k_0
(b) K is finitely generated over k_0 (because it is contained in k', cf. [A V], §14, cor. 3 to prop. 17).

Hence K is either a finite field or a number field, which is absurd since such a field only contains finitely many roots of unity.

§5. The S-bound

We start with the case of tori:

5.1. **The S-bound for a torus: statements.**

Theorem 4. *Let T be a torus over k, and let A be a finite subgroup of $T(k)$. Then*

$$v_\ell(A) \,\le\, m \left[\frac{\dim T}{\varphi(t)}\right],$$

where m and t are defined as in §4 above and φ is Euler's totient function.

The bound given by th.4 is optimal. More precisely:

Theorem 4′. *Assume $m < \infty$. For every $n \ge 1$ there exist a k-torus T of dimension n and a finite subgroup A of $T(k)$ such that $v_\ell(A) = m \cdot [n/\varphi(t)]$.*

Example. Take $k = \mathbf{Q}$ and $\ell = 2$, so that $t = m = 2$. Then th.4 says that any finite 2-subgroup of $T(\mathbf{Q})$ has order $\le 4^{\dim T}$, and th.4′ says that this bound can be attained.

5.2. **Proof of theorem 4.**

Lemma 4. *Let $u \in \mathbf{M}_n(\mathbf{Z}_\ell)$ be an $n \times n$ matrix with coefficients in $\mathbf{Z}_\ell$, which we view as an endomorphism of $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^n$. Then*

$$v_\ell\big(\ker(u)\big) = v_\ell\big(\det(u)\big).$$

Proof. This is clear if u is a diagonal matrix, and one reduces the general case to the diagonal one by multiplying u on the right and on the left by invertible matrices. □

Now let n be the dimension of the torus T. Let $Y(T) = \mathrm{Hom}_{k_s}(\mathbf{G}_m, T)$ be the group of cocharacters of T. The action of Γ_k on $Y(T)$ gives a homomorphism $\rho : \Gamma_k \to \mathrm{Aut}\,\big(Y(T)\big) \cong \mathbf{GL}_n(\mathbf{Z})$. If we identify T with $\mathbf{G}_m \times \cdots \times \mathbf{G}_m$ (over k_s) by choosing a basis of $Y(T)$, the ℓ^∞-division points of $T(k_s)$ form a group isomorphic to $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^n$ and the action of $g \in \Gamma_k$ on that group is by $\rho(g)\chi(g)$, where $\chi = \chi_{\ell^\infty}$.

Lemma 5. *Let A be a finite subgroup of $T(k)$. For every $g \in \Gamma_k$ we have*

$$v_\ell(A) \ \le \ v_\ell\big(\det(\rho(g)\,\chi(g)-1)\big) = v_\ell\big(\det(\rho(g^{-1})-\chi(g)\big).$$

Proof. By replacing A by its ℓ-Sylow subgroup, we may assume that A is an ℓ-group, hence is contained in the ℓ-division points of $T(k_s)$. Since the points of A are rational over k, they are fixed by g, i.e. they belong to the kernel of $g-1$. The inequality then follows from lemma 4, applied to $u = \rho(g)\,\chi(g) - 1$. □

We now choose $g \in \Gamma_k$ such that the inequality of lemma 5 gives that of th.4. Here is the choice:

$$\chi(g) = z_t u, \qquad \text{where } z_t \in \mathbf{Z}_\ell^* \text{ has order } t, \text{ and } v_\ell(1-u) = m.$$

(This works for $\ell = 2$ as well as for $\ell \neq 2$, thanks to the definition of t in §4.1 and §4.2. Note that in all cases but $\ell = 2$, type (c), $\chi(g)$ is a topological generator of $\operatorname{Im}\chi$.)

We have $\rho(g) \in \mathbf{GL}_n(\mathbf{Z})$, and $\rho(g)$ is of finite order (because the image of $\rho : \Gamma_k \to \mathbf{GL}_n(\mathbf{Z})$ is finite). Hence the characteristic polynomial F of $\rho(g^{-1})$ is a product of cyclotomic polynomials:

$$F = \prod \Phi_{d_j}, \qquad \text{with } \sum \varphi(d_j) = n. \tag{5.2.1}$$

The inequality of lemma 5 gives

$$v_\ell(A) \le \sum v_\ell\big(\Phi_{d_j}(z_t u)\big).$$

We thus need to compute $v_\ell\big(\Phi_d(z_t u)\big)$ for every $d \ge 1$. The result is:

Lemma 6. *We have*

$$v_\ell\big(\Phi_d(z_t u)\big) = \begin{cases} m & \text{if } d = t \\ 1 & \text{if } d = t\cdot\ell^\alpha,\ \alpha \ge 1 \quad \text{or} \quad \alpha = -1 \ (\text{if } t = 2 = \ell) \\ 0 & \text{otherwise.} \end{cases}$$

Proof. (We restrict ourselves to the case $\ell \neq 2$. The case $\ell = 2$ is analogous but slightly different.)

We have $\Phi_d(z_t u) = \prod(z_t u - z)$ where z runs through the primitive d-th roots of unity in $\overline{\mathbf{Q}}_\ell$. Write d as $d = \delta\cdot\ell^\alpha$ with $(\delta,\ell) = 1$ and $\alpha \ge 0$. The images of the z's in the residue field $\overline{\mathbf{F}}_\ell$ of $\overline{\mathbf{Q}}_\ell$ are primitive δ-th roots of unity. If $\delta \neq t$, none of them is equal to the image of $z_t u$, which has order t. In that case, all the $z_t u - z$ are units in $\overline{\mathbf{Q}}_\ell$ hence have valuation 0 and we have $v_\ell\big(\Phi_d(z_t u)\big) = 0$. If $\delta = t$, i.e. $d = t\cdot\ell^\alpha$ with $\alpha \ge 0$, there are two cases:

(a) $\alpha = 0$, i.e. $d = t$. In that case, one of the z's is equal to z_t and we have $v_\ell(z_t u - z) = v_\ell(u-1) = m$; the other z's contribute 0.

(b) $\alpha \ge 1$. Here z can be written as $z'\cdot z''$ where z' runs through the t-th primitive roots of 1, and z'' through the ℓ^α-th primitive roots of 1. The valuation of $z - z_t u$ is 0 unless $z' = z_t$, in which case $v_\ell(z - z_t u) = v_\ell(z'' - u)$. It is well-known that $v_\ell(z''-1) = \frac{1}{(\ell-1)\ell^{\alpha-1}}$. Since $v_\ell(u-1) = m$, which is strictly larger, we have

$$v_\ell(z''-u) = v_\ell\big((z''-1)-(u-1)\big) = \frac{1}{(\ell-1)\ell^{\alpha-1}} = \frac{1}{\varphi(\ell^\alpha)}.$$

Since the number of the z'' is $\varphi(\ell^\alpha)$, we thus get $v_\ell\big(\Phi_d(z_t u)\big) = 1$, as claimed. □

We can now prove theorem 4: With the notation of (5.2.1), denote by r_1 the number of j's with $d_j = t$, and by r_2 the number of j's with $d_j = t\cdot\ell^{\alpha_j}$, $\alpha_j \geq 1$, or $\alpha_j = -1$ in case $\ell = 2, t = 2$. Using lemmas 5 and 6 we get

$$v_\ell(A) \leq r_1 m + r_2$$

and of course

$$r_1\varphi(t) + \sum \varphi(t\cdot\ell^{\alpha_j}) \leq n = \dim T.$$

Since $\varphi(t\cdot\ell^{\alpha_j}) \geq \varphi(t)(\ell-1)$ this shows that $r_1\varphi(t) + r_2\varphi(t)(\ell-1) \leq n$. Hence $r_1 + r_2(\ell-1) \leq [n/\varphi(t)]$, and we have:

$$v_\ell(A) \leq r_1 m + r_2 \;\leq\; r_1 m + r_2(\ell-1)m \;\leq\; m[n/\varphi(t)]\,,$$

which concludes the proof. □

Remark. Since $(\ell-1)m > 0$ in all cases (even if $\ell = 2$), the above proof shows that $v_\ell(A)$ can be equal to $m[n/\varphi(t)]$ only when $r_2 = 0$. In other words:

Complement to theorem 4. *Assume $v_\ell(A) = m[n/\varphi(k)]$, where $n = \dim T$. If $g \in \Gamma_k$ is such that $\chi(g) = z_t u$, with $v_\ell(u-1) = m$ as above, the characteristic polynomial of $\rho(g)$ is divisible by $(\Phi_t)^N$, with $N = [n/\varphi(k)]$.*
(In other words, the primitive t-th roots of unity are eigenvalues of $\rho(g)$ with multiplicity N.)

When $t = 1$ or 2 (i.e. when $\varphi(t) = 1$), this can be used to determine the structure of an "optimal" T:

Corollary. *Assume $t = 1$ or 2, and $v_\ell(A) = mn$. Then :*
(i) *If $t = 1$, the torus T is split* (i.e. isomorphic to the product of n copies of $\mathbf{G}_m$).
(ii) *If $t = 2$, T is isomorphic to the product of n non-split tori of dimension 1 which are split by the quadratic extension $k(z_\ell)/k$ if $\ell \neq 2$ and by $k(i)/k$ if $\ell = 2$.*

Proof. We give the proof for $t = 2$ and $\ell > 2$: the case $t = 1$ is easier and the case $t = 2 = \ell$ requires similar, but more detailed, arguments.

Let $\gamma \in \Gamma_k$. We may write $\chi(\gamma)$ as $e_\gamma \cdot u_\gamma$, with $e_\gamma \in \{1, -1\}$ and $u_\gamma \in \{1 + \ell^m\mathbf{Z}_\ell\}$. There are three cases:

(a) $e_\gamma = -1$ and $v_\ell(u_\gamma - 1) = m$
(b) $e_\gamma = -1$ and $v_\ell(u_\gamma - 1) > m$
(c) $e_\gamma = 1$.

In case (a), the "complement" above shows that $\rho(\gamma)$ has -1 for eigenvalue with multiplicity n, hence $\rho(\gamma) = -1$ in $\mathrm{Aut}(T) \simeq \mathbf{GL}_n(\mathbf{Z})$.

In case (b), choose $g \in \Gamma_k$ of type (a); this is possible by the very definition of t and m. The element $g^2\gamma$ is of type (a) (this uses the fact that ℓ is odd); hence we have $\rho(g^2\gamma) = -1$ and since $\rho(g) = -1$ this shows that $\rho(\gamma) = -1$.

If γ is of type (c), then $g\gamma$ is of type (a) or (b) and we have $\rho(g\gamma) = -1$ hence $\rho(\gamma) = 1$.

In all cases, we have $\rho(\gamma) \in \{1, -1\}$, and more precisely $\rho(\gamma) = e_\gamma$. The corollary follows. □

It would be interesting to have a similar classification for $t > 2$.

5.3. **Proof of theorem 4′: construction of tori with large A's.** To prove th.4′ it is enough to construct a k-torus T, of dimension $n = \varphi(t)$, such that $T(k)$ contains a cyclic subgroup of order ℓ^m. Here is the construction:

Let K be the field $k(z_\ell)$ if $\ell \neq 2$ and the field $k(i)$ if $\ell = 2$. It is a cyclic extension of k of degree t with Galois group C_t. Let $T_1 = R_{K/k}\mathbf{G}_m$ be the torus: "multiplicative group of K"; we have $T_1(k) = K^*$, and $T_1(k)$ contains the group $\langle z_{\ell^m} \rangle$, cf. §4. If σ is a generator of C_t, σ acts on T_1, and we have $\sigma^t - 1 = 0$ in the ring $\mathrm{End}(T_1)$. Let us write the polynomial $X^t - 1$ as $\Phi_t(X) \cdot \Psi(X)$, where Φ_t is the t-th cyclotomic polynomial. We have $\Phi_t(\sigma)\Psi(\sigma) = 0$ in $\mathrm{End}(T_1)$. Let $T = \mathrm{Im}\,\Psi(\sigma)$ be the image of

$$\Psi(\sigma) : T_1 \to T_1.$$

One checks that

(a) $\dim T = \varphi(t)$

(b) $T(k)$ contains z_{ℓ^m}.

(For $\ell \neq 2$, (b) follows from the fact that the restriction of $\Psi(\sigma)$ to $\langle z_{\ell^m} \rangle$ is an automorphism. For $\ell = 2$, use the fact that T is the kernel of $\Phi_t(\sigma)$.)

Hence T has the required properties. □

Alternate description of T. It is enough to describe its character group $T^* = \mathrm{Hom}_{k_s}(T, \mathbf{G}_m)$, together with the action of Γ_k on T^*:

- $T^* = \mathbf{Z}[X]/\Phi_t(X) =$ algebraic integers of the cyclotomic field $\mathbf{Q}(\mu_t)$
- Γ_k acts on T^* by $\Gamma_k \to \mathrm{Im}\,\chi_{\ell^\infty} \to C_t \xrightarrow{\sim} \mathrm{Aut}\big(\mathbf{Q}(\mu_t)\big)$.

(It does not matter which isomorphism of C_t onto $\mathrm{Aut}\big(\mathbf{Q}(\mu_t)\big)$ one chooses; they all give isomorphic tori.)

5.4. **The S-bound for reductive groups.** Recall, cf. §3.3, that "reductive" ⇒ "connected".

Theorem 5. *Let G be a reductive group over k, of rank r, with Weyl group W. If A is a finite subgroup of $G(k)$, one has*

$$v_\ell(A) \leq m \left[\frac{r}{\varphi(t)} \right] + v_\ell(W).$$

Proof. As usual, we may assume that A is an ℓ-group. In that case it is nilpotent, and by the corollary to th.3″ of §3.3 there exists a maximal k-torus T of G whose normalizer $N = N_G(T)$ contains A. Put $W_T = N/T$; this is a finite k-group such that $W_T(k_s) \simeq W$. If A_T denotes the intersection of A with $T(k)$, we have an exact sequence

$$1 \to A_T \to A \to W_T(k)\,.$$

Hence $v_\ell(A) \leq v_\ell(A_T) + v_\ell\big(W_T(k)\big)$. By th.4, we have $v_\ell(A_T) \leq m \cdot [r/\varphi(t)]$; on the other hand $W_T(k)$ is isomorphic to a subgroup of W, hence $v_\ell\big(W_T(k)\big)$ $\leq v_\ell(W)$. The theorem follows. □

Corollary. *If $r < \varphi(t)$, then $G(k)$ is ℓ-torsion free* (i.e. does not contain any elements of order ℓ).

Proof. We have $\left[\frac{r}{\varphi(t)}\right] = 0$. Hence by th.5 it is enough to show that $v_\ell(W) = 0$, but this follows from th.1 of §1.1 since W is isomorphic to a subgroup of $\mathbf{GL}_r(\mathbf{Z})$ and $r < \varphi(t) \leq t \leq \ell - 1$. □

Remark. The "S-bound" given by th.5 looks *a priori* rather coarse:

(a) The torus T is not an arbitrary torus of dimension r; the fact that it is a subtorus of G puts non-trivial conditions on it; for instance the action of Γ_k on $T^* = \mathrm{Hom}_{k_s}(T, \mathbf{G}_m)$ stabilizes the set of roots.

(b) The group $W_T(k)$ is in general smaller than W itself, and the image of $N(k) \to W_T(k)$ may be even smaller.

It is therefore surprising how often the S-bound is close to being optimal. As an example, take $k = \mathbf{Q}$ and G of type E_8. We have $m = 1$ and $t = \ell - 1$ (except when $\ell = 2$ in which case $m = t = 2$), $r = 8$, $|W| = 2^{14}3^5 5^2 7$. The S-bound tells us that, if A is a finite subgroup of $G(\mathbf{Q})$, its order divides the number

$$M_S = 2^{30}\cdot 3^{13}\cdot 5^6\cdot 7^5\cdot 13^2\cdot 17\cdot 19\cdot 31\,.$$

We shall see later (cf. §6.3.2 and §7) that the best bound is $M = M_S/5\cdot 7\cdot 17$. In particular, the ℓ-factors of M_S are optimal for all ℓ's except $\ell = 5, 7$ and 17.

§6. The M-bound

6.1. **Notation.** From now on, G is a semisimple [1] group over k. We denote by R its root system (over k_s), by W its Weyl group, and by r its rank. The group W has a natural linear representation of degree r. The invariants of W acting on $\mathbf{Q}[x_1, \ldots, x_r]$ make up a graded polynomial algebra of the form $\mathbf{Q}[P_1, \ldots, P_r]$, where the P_i are homogeneous of degrees d_i, with $d_1 \leq d_2 \leq \cdots \leq d_r$, (Shephard-Todd theorem, cf. e.g. [LIE V], §5, th.4 or [Se 00], p.95). The d_i's are called the *invariant degrees* of W (or of G). One has

$$\prod d_i = |W| \quad \text{and} \quad \sum (2d_i - 1) = \dim G\,.$$

When G is quasi-simple (i.e. when R is irreducible) d_r is equal to the Coxeter number $h = (\dim G)/r - 1$, and one has the symmetry formula

$$d_i + d_{r+1-i} = h + 2\,.$$

Moreover, if $j < h$ is prime to h, then $j+1$ is one of the d_i's. These properties make $d_1, \ldots, d_r$ very easy to compute (see e.g. the tables of [LIE VI]).

For instance, for G of type E_8, the d_i's are: 2, 8, 12, 14, 18, 20, 24, 30.

Let $\mathrm{Dyn}(R)$ be the Dynkin diagram of R. There is a natural action of the Galois group Γ_k on $\mathrm{Dyn}(R)$: this follows from the fact that $\mathrm{Dyn}(R)$ can be defined intrinsically from $G_{/k_s}$ (cf. [LIE VIII], §4, no 4, Scholie, or [SGA 3],

[1] We could also accept inner forms of reductive groups, for instance $\mathbf{GL}_n$ or more generally $\mathbf{GL}_D$, where D is a central simple k-algebra with $[D : k] = n^2$. In that case, one has $r = n$, the d_i's are the integers $1, 2, \ldots, n$ and th.6 is valid, with the same proof.

exposé XXIV, §3, p.344). In what follows (with the only exception of §6.7) we make the assumption that *the action of* Γ_k *on* $\mathrm{Dyn}(R)$ *is trivial*: one then says that G is *of inner type* (it can be obtained from a split group G_0 by a Galois twist coming from the adjoint group of G_0).

Examples of groups of inner type :
- $\mathbf{SL}_n$, or more generally, $\mathbf{SL}_D$, where D is a central simple algebra over k.
- Any group G whose root system has no non-trivial automorphism, e.g. any group of type $A_1, B_r, C_r, G_2, F_4, E_7, E_8$.

6.2. **Statement of the theorem.** We fix ℓ, k, and the root system R of G. Recall that $\mathrm{Im}\,\chi_{\ell^\infty}$ is a closed subgroup of $\mathbf{Z}_\ell^*$. Define:

$$M(\ell,k,R) = \inf_{x\,\in\,\mathrm{Im}\,\chi_{\ell^\infty}} \sum v_\ell(x^{d_i}-1) = \inf_{g\,\in\,\Gamma_k} \sum v_\ell(\chi_{\ell^\infty}(g)^{d_i}-1)\,.$$

This is either an integer ≥ 0 or ∞ (it is ∞ if and only if the invariants m, t of k defined in §4 are such that $m = \infty$ and t divides one of the d_i's, see prop.4 below).

Theorem 6. *Let A be a finite subgroup of $G(k)$. Then $v_\ell(A) \leq M(\ell,k,R)$.* (Recall that G is semisimple of inner type, cf. §6.1.)

This is what we call the "M-bound" for $v_\ell(A)$. It will be proved in §6.5 below by a method similar to Minkowski's. We shall see in Lect. III that it is "optimal" except possibly in the case $\ell = 2$, type (c) of §4.2.

For computations, it is useful to write $M(\ell,k,R)$ explicitly in terms of the invariants t and m of §4:

Proposition 4. (1) *If $\ell \neq 2$ or if $\ell = 2, t = 1$* (case (a)), *one has*

$$M(\ell,k,R) = \sum_{\substack{i \\ d_i \equiv\, 0 \pmod t}} \big(m + v_\ell(d_i)\big)$$

(2) *If $\ell = 2$ and $t = 2$* $\big($cases (b) and (c)$\big)$, *one has*

$$M(2,k,R) = r_1 + mr_0 + v_2(W)\,,$$

where r_0 (resp. *r_1*) *is the number of indices i such that d_i is even* (resp. *d_i is odd*).

Proof. Let us begin with the case $\ell \neq 2$. One shows first that, if $t|d$, one has $v_\ell(x^d-1) \geq m + v_\ell(d)$ for every $x \in \mathrm{Im}\,\chi_{\ell^\infty}$. (This is easy, since x can be written as zu with $z^t = 1$ and $v_\ell(u-1) \geq m$, hence $x^d - 1 = u^d - 1$.)

This already shows that $M(\ell,k,R) \geq \sum_{t|d_i} \big(m + v_\ell(d_i)\big)$. To prove the opposite inequality, one chooses $x \in \mathrm{Im}\,\chi_{\ell^\infty}$ of the form zu with z of order t and $v_\ell(u-1) = m$. One gets (1).

The same argument works if $\ell = 2$ and $t = 1$. If $\ell = 2$ and $t = 2$, one has

$$\begin{aligned} v_2(x^d-1) &\geq m + v_2(d) \quad \text{if } d \text{ is even} \\ v_2(x^d-1) &\geq 1 \quad \text{if } d \text{ is odd}\,, \end{aligned}$$

for every $x \in \mathrm{Im}\,\chi_{2^\infty}$. This gives:

$$M(2,k,R) \geq \sum_{d_i \text{ odd}} 1 + \sum_{d_i \text{ even}} \big(m + v_2(d_i)\big) = r_1 + mr_0 + v_2(W)\,.$$

To get the opposite inequality, observe that $x = -1 + 2^m$ belongs to $\operatorname{Im} \chi_{2^\infty}$ and check that $\sum v_2(x^{d_i} - 1)$ is equal to $r_1 + mr_0 + v_2(W)$. □

Corollary. *Let $a(t)$ be the number of indices i such that $d_i \equiv 0 \pmod t$. If $a(t) = 0$, then $G(k)$ is ℓ-torsion free.*

Indeed, if $a(t) = 0$, the sum occurring in prop.4 is an empty sum, hence $M(\ell, k, R) = 0$ and one applies th.6. □

6.3. **Two examples: A_1 and E_8.** We take $k = \mathbf{Q}$, so that $t = \ell - 1$ and $m = 1$ if $\ell > 2$ and $t = m = 2$ if $\ell = 2$.

6.3.1. *Type A_1.* There is only one d_i, namely $d_1 = 2$, and prop.4 gives:

$$M\big(\ell, \mathbf{Q}, A_1\big) = \begin{cases} 3 & \text{if} \quad \ell = 2 \\ 1 & \text{if} \quad \ell = 3 \\ 0 & \text{if} \quad \ell > 3\,. \end{cases}$$

In other words, every finite subgroup of $G(\mathbf{Q})$ has an order which divides $2^3 \cdot 3$. This bound is optimal in the following sense:

(a) The split adjoint group $\mathbf{PGL}_2(\mathbf{Q})$ contains both a subgroup of order 3 and a dihedral subgroup of order 8 (but no subgroup of order 24).

(b) The simply connected group $\mathbf{SL}_{\mathbf{H}}(\mathbf{Q})$, where $\mathbf{H}$ is the standard quaternion division algebra, contains a subgroup of order 24 which is isomorphic to $\mathbf{SL}_2(\mathbf{F}_3)$. However the split group $\mathbf{SL}_2(\mathbf{Q})$ does not contain any subgroup of order 8 (but it does contain cyclic subgroups of order 3 and 4).

6.3.2. *Type E_8.* If we define $M(\mathbf{Q}, E_8)$ as $\prod_\ell \ell^{M(\ell, \mathbf{Q}, E_8)}$, prop.4 gives:

$$M(\mathbf{Q}, E_8) = 2^{30} \cdot 3^{13} \cdot 5^5 \cdot 7^4 \cdot 11^2 \cdot 13^2 \cdot 19 \cdot 31, \text{ see e.g. [Se 79], §3.3.}$$

By th.6, the order of every finite subgroup of $G(\mathbf{Q})$ divides $M(\mathbf{Q}, E_8)$. As we shall see in the next lecture, this multiplicative bound is optimal.

6.4. **A Chebotarev-style result.** We need such a result in order to generalize Minkowski's method of §1.

Let L be a normal domain which is finitely generated over $\mathbf{Z}$ as a ring, and let k be its field of fractions. If $d = \dim(L)$ denotes the Krull dimension of L ([AC VIII], §1), one has (*loc.cit.*, §2):

$$\begin{array}{lcllcl} d & = & 1 + \operatorname{tr.deg}(k/\mathbf{Q}) & \text{if} & \operatorname{char}(k) & = & 0 \\ d & = & \operatorname{tr.deg}(k/\mathbf{F}_p) & \text{if} & \operatorname{char}(k) & = & p > 0\,. \end{array}$$

Let $\operatorname{Specmax}(L)$ be the set of the maximal ideals of L (= set of closed points of $\operatorname{Spec}(L)$). If $x \in \operatorname{Specmax}(L)$, the residue field $\kappa(x) = L/x$ is finite (see e.g. [AC V], p. 68, cor. 1). We put $Nx = |\kappa(x)|$; it is the *norm* of x.

When $d = 0$, L is a finite field, and $\operatorname{Specmax}(L)$ has only one element. If $d > 0$ (e.g. when $\operatorname{char}(k) = 0$), then $\operatorname{Specmax}(L)$ is infinite. More precisely, the Dirichlet series $z(s) = \sum_x 1/(Nx)^s$ converges for $\operatorname{Re}(s) > d$, and one has

$$z(s) \sim \log\big(1/(s-d)\big) \quad \text{when} \quad s \to d \quad (\text{with } s > d)\,. \tag{6.4.1}$$

See [Se 65], §2.7, which only contains a sketch of proof; complete details (for a slightly weaker statement) can be found in [Pi 97], App. B [2]; see also [FW 84], pp.206-207.

Let now n be an integer ≥ 1 which is invertible in L (and hence in k). Let $\chi_n : \Gamma_k \to (\mathbf{Z}/n\mathbf{Z})^*$ denote the n-th cyclotomic character of k, cf. §4.0. As in §4, we shall be interested in $\operatorname{Im}\chi_n \subset (\mathbf{Z}/n\mathbf{Z})^*$.

Theorem 7. *Let c be an element of $(\mathbf{Z}/n\mathbf{Z})^*$, and let X_c be the set of all $x \in \operatorname{Specmax}(L)$ such that $Nx \equiv c \pmod{n}$. Then :*

a) *If $c \notin \operatorname{Im}\chi_n$, then $X_c = \varnothing$.*

b) *If $c \in \operatorname{Im}\chi_n$ and $d > 0$, then X_c is Zariski-dense in* $\operatorname{Specmax}(L)$ (or in $\operatorname{Spec}(L)$, this amounts to the same). *In particular, X_c is infinite.*

A more concrete formulation of b) is that, for every non-zero $f \in L$, there exists an x with $f \notin x$ and $Nx \equiv c \pmod{n}$.

Example. Take $L = \mathbf{Z}[1/n]$. Then $\operatorname{Specmax}(L)$ is the set of all prime numbers which do not divide n, and th.7 translates into Dirichlet's theorem on the existence of primes in arithmetic progressions.

Proof of theorem 7. The group $C = \operatorname{Im}\chi_n$ is the Galois group of the cyclotomic extension $k(z_n)/k$. Let L_n be the integral closure of L in $k(z_n)$. One checks by standard arguments that the ring extension L_n/L is finite and étale. In geometric terms, $\operatorname{Spec}(L_n) \to \operatorname{Spec}(L)$ is a finite étale covering. The group C acts freely on $\operatorname{Spec}(L_n)$, with quotient $\operatorname{Spec}(L)$. For every closed point x of $\operatorname{Spec}(L)$, the Frobenius element σ_x of x is a well-defined conjugacy class of C (hence an element of C since C is commutative). Moreover, if we view C as a subgroup of $(\mathbf{Z}/n\mathbf{Z})^*$, σ_x is the image of Nx in $\mathbf{Z}/n\mathbf{Z}$. This proves a).

Suppose now that $d > 0$ and that c belongs to $C = \operatorname{Im}\chi_n$. Let $z_c(s)$ be the Dirichlet series $\sum 1/(Nx)^s$, where the sum is over the elements x of X_c. The general Chebotarev density theorem ([Se 65], [Pi 97]) gives:

$$z_c(s) \sim \frac{1}{|C|}\log(1/(s-d)) \quad \text{when } s \to d \quad \text{with} \quad s > d\,. \tag{6.4.2}$$

In particular, we have $z_c(d) = +\infty$. If the Zariski closure $\overline{X}_c$ of X_c were of dimension $< d-1$, we would have $z_c(d) < \infty$, as one sees by splitting $\overline{X}_c$ into irreducible components, and applying (6.4.1). Hence b). □

6.5. **Proof of theorem 6.** Let $A \subset G(k)$ be as in th.6. We want to prove that

$$v_\ell(A) \leq M(\ell, k, R)\,.$$

We do it in three steps:

[2] When $\operatorname{char}(k) = 0$ one can give a stronger statement, in the spirit of the Prime Number Theorem:

For every $X \geq 2$, call $\pi_L(X)$ the number of $x \in \operatorname{Specmax}(L)$ such that $Nx \leq X$. Then:

$$\pi_L(X) = (1/d)\,X^d/\log X + O(X^d/\log^2 X) \quad \text{when} \quad X \to \infty\,.$$

The general Chebotarev density theorem can also be stated (and proved) in terms of such "natural" density (standard method: use Weil-Deligne estimates to reduce everything to the known case $d = 1$).

6.5.1. **The case where k is finite.** Put $q = |k|$. It is well-known that

$$|G(k)| = q^N \prod (q^{d_i} - 1), \qquad \text{where} \quad N = |R|/2 = \sum (d_i - 1).$$

This shows that $v_\ell(A) \leq \sum v_\ell(q^{d_i} - 1)$. The integer q, viewed as an element of $\mathbf{Z}_\ell^*$, is a topological generator of $\operatorname{Im} \chi_{\ell^\infty}$. Hence every element u of $\operatorname{Im} \chi_{\ell^\infty}$ is an ℓ-adic limit of powers of q and this implies that $v_\ell(u^d - 1) \geq v_\ell(q^d - 1)$ for every $d \geq 1$. Hence the lower bound which defines $M(\ell, k, R)$ is equal to $\sum v_\ell(q^{d_i} - 1)$; this proves th.6 in the case where k is finite.

6.5.2. **The case where k is finitely generated over its prime subfield.** By 6.5.1, we may assume that k is infinite. We need a subring L of k, with field of fractions k, which has the following properties:

(a) L is normal, finitely generated over $\mathbf{Z}$ and contains $1/\ell$.

(b) G comes by base change from a semisimple group scheme $\underline{G}$ over L, in the sense of [SGA 3], XIX. 2.7.

(c) A is contained in the group $\underline{G}(L)$ of the L-points of $\underline{G}$.

Lemma 7. *There exists such an L.*

This is standard, see e.g. [EGA IV], §8.1 □

Let us now choose $(L, \underline{G})$ with properties (a), (b) and (c). For every $x \in \operatorname{Specmax}(L)$, the fiber $\underline{G}_x$ of $\underline{G}$ at x is a semisimple group over $\kappa(x)$, of type R. Moreover, the Dynkin diagram of $\underline{G}$ is finite étale over $\operatorname{Spec}(L)$, cf. [SGA 3], XXIV.3.2; since it is "constant" for the generic fiber (i.e. over k) it is constant everywhere; this shows that the $\underline{G}_x$ are of inner type. The inclusion map $i : A \to \underline{G}(L)$ gives for every x a homomorphism $i_x : A \to \underline{G}\big(\kappa(x)\big)$. Since i is injective, there is an open dense subset X_0 of $\operatorname{Specmax}(L)$ such that i_x is injective for all $x \in X_0$. We thus get:

$$v_\ell(A) \;\leq\; v_\ell\big(\underline{G}(\kappa(x))\big) = \sum v_\ell\big((Nx)^{d_i} - 1\big) \quad \text{for all} \quad x \in X_0,$$

cf. 6.5.1. Let u be any element of $\operatorname{Im} \chi_{\ell^\infty}$. By applying th.7 to the image of u in $(\mathbf{Z}/\ell^j\mathbf{Z})^*$ with $j = 1, 2, \ldots$, we find a sequence of points x_j of X_0 such that $\lim Nx_j = u$ in $\mathbf{Z}_\ell^*$. We have:

$$v_\ell(u^{d_i} - 1) = \lim_{j \to \infty} \sum v_\ell\big((Nx_j)^{d_i} - 1\big)\,,$$

and applying the formula above to each of the x_j's we obtain

$$v_\ell(A) \leq \sum v_\ell(u^{d_i} - 1) \qquad \text{for every} \quad u \in \operatorname{Im} \chi_{\ell^\infty}\,.$$

This proves th.6 in the case 6.5.2.

[Variant: One reduces the general case to the case where $\dim(L) = 1$ by using Hilbert's irreducibility theorem, as explained in [Se 81], p.2; in the case $\dim(L) = 1$, one can apply the standard Chebotarev theorem instead of the general one.]

6.5.3. **The general case.** The same argument as for lemma 7 shows that G comes by base change from a semisimple group G' over a subfield k' of k which is finitely generated over the prime subfield of k (i.e. $\mathbf{F}_p$ or $\mathbf{Q}$). Moreover, one may assume (after enlarging k' if necessary) that A is contained in $G'(k')$. The Galois group $\Gamma_{k'}$ acts on the Dynkin diagram $\operatorname{Dyn}(R)$ of G' (which is the same as the one of G). Let k'' be the Galois

extension of k' corresponding to the kernel of $\Gamma_{k'} \to \mathrm{Aut}\ \mathrm{Dyn}(R)$. Since G is of inner type over k, the field k'' is contained in k. By base change to k'', G' gives a semisimple group G'' which is of inner type and we may apply 6.5.2 to (G'', A). We get $v_\ell(A) \leq M(\ell, k'', R)$. Since k'' is contained in k, we have $M(\ell, k'', R) \leq M(\ell, k, R)$: the group $\mathrm{Im}\,\chi_{\ell^\infty}$ can only decrease by field extensions. Hence $v_\ell(A) \leq M(\ell, k, R)$. □

6.5.4. **Remark.** Surprisingly, the proof above does not really use the hypothesis that A is a subgroup of $G(k)$. It uses only that A *acts freely on* G, viewed merely as a k-variety (and not as a homogeneous space); this is indeed enough to ensure that $v_\ell(A) \leq v_\ell(G(k))$ when k is finite. Here is an example: take $G = \mathbf{SL}_2$, $\ell = 2$, $k = \mathbf{Q}$; the M-bound is 3, which means that any finite 2-subgroup of $\mathbf{SL}_2(\mathbf{Q})$ has order ≤ 8. As was said in §6.3.1, there is in fact no subgroup of order 8 in $\mathbf{SL}_2(\mathbf{Q})$. But one can make a cyclic group of order 8 act freely on the variety $\mathbf{SL}_2$: take for instance the group generated by the automorphism:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} d-c & -c-d \\ \frac{(a-b)}{2} & \frac{(a+b)}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ \frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}.$$

Hence, even in this bad-looking case, the M-bound can claim to be "optimal".

6.6. **An analogue of Sylow's theorem.**

Theorem 8. *Let A and A' be two finite ℓ-subgroups of $G(k)$. Assume that $v_\ell(A)$ is equal to the* M*-bound $M(\ell, k, R)$. Then there exists $y \in G(\bar{k})$ such that $yA'y^{-1} \subset A$.*

Corollary. *If both A and A' attain the* M*-bound, then they are geometrically conjugate* (i.e. conjugate in $G(\bar{k})$). *In particular, they are isomorphic.*

Proof. We may assume that k is finitely generated over its prime subfield. If it is finite, th.8 is just a special case of Sylow's theorem. Let us assume that k is infinite, and choose $L, \underline{G}$ as in §6.5.2 with $A, A' \subset \underline{G}(L)$. Let Y be the subscheme of $\underline{G}$ made up of the points y with $yA'y^{-1} \subset A$. Let X be the set of all $x \in \mathrm{Specmax}(L)$ such that Nx, viewed as an element of $\mathbf{Z}_\ell^*$, is of the form $z_t u$ with z_t of order t and $v_\ell(u-1) = m$ (note that m is finite, cf. §4.3). It follows from th.7, applied to $n = \ell^{m+1}$, that X is Zariski-dense in $\mathrm{Spec}(L)$. If $x \in \mathrm{Specmax}(L)$, the groups A and A' inject into $\underline{G}(\kappa(x))$ (this is an easy consequence of the hypothesis that ℓ is invertible in L). If moreover x belongs to X, then the same computation as in §5.2 shows that $v_\ell\big(\underline{G}(\kappa(x))\big)$ is equal to the M-bound, hence A is an ℓ-Sylow of $\underline{G}(\kappa(x))$. By Sylow's theorem, this shows that A' is conjugate in $\underline{G}(\kappa(x))$ to a subgroup of A. In particular, the fiber at x of $Y \to \mathrm{Spec}(L)$ is non-empty. Since X is Zariski-dense, this implies that the generic fiber $Y_{/k}$ of $Y \to \mathrm{Spec}(L)$ is non-empty, i.e. that $Y(\bar{k})$ is non-empty. □

Remark. One can show that Y is smooth over L, and hence that $Y(k_s) \neq \varnothing$ which is slightly more precise than $Y(\bar{k}) \neq \varnothing$.

Exercise. Show that a family of polynomial equations with coefficients in $\mathbf{Z}$ has a solution in $\mathbf{C}$ if and only if it has a solution in $\mathbf{Z}/p\mathbf{Z}$ for infinitely many p's.

6.7. **Arbitrary semisimple algebraic groups.** In the previous sections, we have assumed that G is of inner type, i.e. that the natural homomorphism

$$\varepsilon : \Gamma_k \to \operatorname{Aut} \operatorname{Dyn}(R)$$

is trivial. Let us now look briefly at the general case, where no hypotheses on ε are made. In order to state the result which replaces th.6 we need to introduce the linear representations ε_d of Γ_k defined as follows:

Let $S = \mathbf{Q}[P_1, \ldots, P_r]$ be the $\mathbf{Q}$-algebra of W-invariant polynomials, cf. §6.1. Let $I = (P_1, \ldots, P_r)$ be the augmentation ideal of S; put $V = I/I^2$. The vector space V is of dimension r, and is graded; the dimension of its d-th component V_d is equal to the number of indices i with $d_i = d$. The group $\operatorname{Aut} \operatorname{Dyn}(R)$ acts on S, V and the V_d's; by composing this action with ε, we get for each $d > 0$ a linear representation

$$\varepsilon_d : \Gamma_k \to \operatorname{Aut}(V_d)\,.$$

Theorem 6′. *Let A be a finite subgroup of $G(k)$. Then*:

$$v_\ell(A) \leq \inf_{g \in \Gamma_k} \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(g)^d - \varepsilon_d(g))\big)$$

(The determinant is relative to the vector space $V_d \otimes \mathbf{Q}_\ell$.)

Proof (sketch). The method is the same as the one used for th.6. There are three steps:

(1) Reduction to the case where k is finitely generated over its prime subfield; this is easy.

(2) Reduction to the case where k is finite, via the general Chebotarev density theorem instead of th.7.

(3) The case where k is finite. In that case, if $q = |k|$, and if σ is the Frobenius generator of Γ_k, one has (cf. e.g. [St 68] th. 11.16)

$$v_\ell\big(G(k)\big) = \sum_d v_\ell\big(\det(q^d - \varepsilon_d(\sigma))\big) = \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(\sigma)^d - \varepsilon_d(\sigma))\big)$$

hence the desired formula:

$$(*) \qquad v_\ell(A) \leq \sum_d v_\ell\big(\det(\chi_{\ell^\infty}(g)^d - \varepsilon_d(g))\big)$$

in the special case $g = \sigma$. By applying this to the finite extensions of k, one sees that the inequality $(*)$ is valid for all $\sigma^n, n = 1, 2, \ldots$, and hence for all $g \in \Gamma_k$, since the σ^n are dense in Γ_k. □

Remark. One may also prove th.6′ using ℓ-adic cohomology, cf. §6.8.

Example. Take R of type A_2, so that $\operatorname{Aut} \operatorname{Dyn}(R) = \{1, -1\}$ and ε may be viewed as a quadratic character of Γ_k. The V_d's are of dimension 1 for $d = 2, 3$ and are 0 otherwise. The action of $\operatorname{Aut} \operatorname{Dyn}(R)$ on V_d is trivial for all d, except $d = 3$. Hence $\varepsilon_2 = 1$, $\varepsilon_3 = \varepsilon$, and th.6′ can be rewritten as:

$$v_\ell(A) \leq \inf_{g \in \Gamma_k} \big\{ v_\ell(\chi_{\ell^\infty}(g)^2 - 1) + v_\ell\big(\chi_{\ell^\infty}(g)^3 - \varepsilon(g)\big) \big\}\,.$$

A similar result holds for the types A_r ($r > 2$), D_r (r odd) and E_6, with 2 (resp. 3) replaced by the even d_i's (resp. the odd d_i's).

6.8. **The cohomological method.** Let us consider first the general situation suggested in §6.5.4 where a finite group A acts freely on a quasi-projective k-variety X. As explained in [Il 05], §7, one can then give an upper bound for $v_\ell(A)$ in terms of the action of Γ_k on the étale cohomology of X. More precisely, let $H^i_c(X)$ denote the i-th étale cohomology group of $X_{/k_s}$, with proper support and coefficients $\mathbf{Q}_\ell$; it is a finite dimensional $\mathbf{Q}_\ell$-vector space which is 0 for $i > 2\cdot\dim(X)$. There is a natural action of Γ_k on $H^i_c(X)$, and, for each $g \in \Gamma_k$, one can define the "Lefschetz number" $\Lambda_X(g)$ by the usual formula:

$$\Lambda_X(g) = \sum_i (-1)^i \mathrm{Tr}\big(g|H^i_c(X)\big)\,.$$

One has $\Lambda_X(g) \in \mathbf{Z}_\ell$. Moreover:

Theorem 6″. $v_\ell(A) \leq \inf_{g\in\Gamma_k} v_\ell\big(\Lambda_X(g)\big)$.

Proof. See [Il 05], §7, especially cor.7.5. The proof follows the same pattern as the other proofs of the present §: one uses Chebotarev to reduce to the case where k is finite, in which case the result follows from the fact, due to Grothendieck, that, if σ is the (geometric) Frobenius generator of Γ_k, then $\Lambda_X(\sigma)$ is equal to $|X(k)|$, hence is divisible by $|A|$ since the action of A is free. (As in the proof of th.6′, one applies this, not only to σ but also to its powers σ^n, $n > 0$, and one uses the fact that the σ^n are dense in Γ_k.) □

If one applies th.6″ to $A \subset G(k)$, with A acting by left translations on $X = G$, one recovers th.6 and th.6′, thanks to the known structure of the cohomology of G, cf. e.g. [SGA $4\frac{1}{2}$], p. 230.

6.9. **The Cremona group: open problems.** Recall that the *Cremona group* $\mathbf{Cr}_r(k)$ is the group of k-automorphisms of the field $k(X_1, \ldots, X_r)$, i.e. the group of birational automorphisms (or "pseudo-automorphisms", cf. [De 70]) of the projective r-space over k. For $r = 1$, one has $\mathbf{Cr}_1(k) = \mathbf{PGL}_2(k)$. Let us assume that $r \geq 2$. As explained in [De 70], $\mathbf{Cr}_r$ is not an algebraic group, but looks like a kind of very large semisimple group of rank r (very large indeed: its "Weyl group" is the infinite group $\mathbf{GL}_r(\mathbf{Z})$). Not much is known about the finite subgroups of $\mathbf{Cr}_r(k)$ beyond the classical case $r = 2$ and k algebraically closed. Here is a question suggested by §5.1:

- Is it true that $\mathbf{Cr}_r(k)$ has no ℓ-torsion if $\varphi(t) > r$?

In the special case $k = \mathbf{Q}$, $r = 2$ or 3, this amounts to:

- Is it true that the fields $\mathbf{Q}(X_1, X_2)$ and $\mathbf{Q}(X_1, X_2, X_3)$ have no automorphism of prime order ≥ 11? (Automorphisms of order 2, 3, 5 and 7 do exist.)

It would be very interesting to attack these questions using cohomology, but I do not see how to do this. It is not even clear how to define cohomological invariants of $\mathbf{Cr}_r(\mathbf{C})$, since there is no natural topology on that group. Still, one would like to give a meaning to a sentence such as

"$\mathbf{Cr}_r(\mathbf{C})$ is connected for $r \geq 1$ and simply-connected for $r \geq 2$."

III. Construction of large subgroups

§7. Statements

We keep the notation of Lecture II: k, ℓ, χ_{ℓ^∞}, t, m, ... We consider only semisimple groups over k with a root system R which is *irreducible.* The M-bound of §6.2 will be denoted by $M(\ell, k, R)$; it only depends on the pair (ℓ, k) via the invariants t and m, and on R via the degrees $d_1, \ldots, d_r$ of W. We limit ourselves to the case $m < \infty$; see §14 for the case $m = \infty$.

A pair (G, A), where G is of inner type with root system R, and $A \subset G(k)$ is a finite group, will be called *optimal* if $v_\ell(A)$ is equal to the M-bound $M(\ell, k, R)$. (We could assume that A is an ℓ-group, but this would not be convenient for the constructions which follow.) Our goal is to prove:

Theorem 9. *If $\ell \neq 2$, an optimal pair (G, A) exists (for any k, R).*

Theorem 10. *If $\ell = 2$, an optimal pair (G, A) exists if $\operatorname{Im} \chi_{2^\infty}$ is of type* (a) *or* (b) *in the sense of* §4.2 (i.e. if $\operatorname{Im} \chi_{2^\infty}$ can be topologically generated by one element).

Theorem 11. *In the case $\ell = 2$ and type* (c)*, there exists (G, A) with*

$$v_2(A) = r_0 m + v_2(W)$$

where r_0 is the number of indices i such that d_i is even.

Note that here the M-bound is $M(2, k, R) = r_1 + r_0 m + v_2(W)$ with $r_1 = r - r_0$, cf. §6.2, prop.4. Hence $v_2(A)$ differs from $M(2, k, R)$ only by r_1. In particular, A is optimal if $r_1 = 0$. Hence:

Corollary. *If all the d_i's are even* (i.e. if $-1 \in W$)*, then an optimal pair (G, A) exists for $\ell = 2$* (and hence for all ℓ's, thanks to th.9).

This applies in particular to the exceptional types G_2, F_4, E_7 and E_8.

Remarks. (1) The simplest case where the M-bound is not attained is $k = \mathbf{Q}$, $\ell = 2$, R of type A_2, where $m = 2$, $r_0 = 1$, $r = 2$, the M-bound is 4, and it follows from [Sch 05] that $v_2(A) \leq 3$ for every finite subgroup A of $G(\mathbf{Q})$.

(2) In Theorems 9, 10 and 11, no claim is made on the structure of G except that it is of inner type and that its root system is of type R. However, if one looks closely at the proofs given in the next sections, one sees that G can be chosen to have the following properties:
- it is simply connected;
- it splits over the cyclotomic field $k(z_\ell)$ if $\ell > 2$, and over $k(i)$ if $\ell = 2$.

Simple examples (such as $k = \mathbf{Q}$, $\ell = 3$, G of type G_2) show that it is not always possible to have G split over k.

(3) If G is not chosen carefully, the group $G(k)$ may not contain any large ℓ-subgroup, even if k contains all the roots of unity. For instance, when R is of type A_1 (resp. of type E_8) it is easy (resp. it is possible) to construct a pair (G, k) such that the only torsion elements of $G(k)$ have order 1 or 2 (resp. $G(k)$ is torsion free).

(4) The three theorems above are almost obvious if the characteristic is $p \neq 0$ (especially Theorem 11 since type (c) never occurs!): one takes a

finite field k_0 contained in k which has the same invariants t and m (this is easily seen to be possible – if k is finitely generated over $\mathbf{F}_p$, one chooses the maximal finite subfield of k), and one takes for G the group deduced by base change from a split group G_0 over k_0 with root system R. If we choose for A the finite group $G_0(k_0)$, it is clear from the way we got the M-bound that $v_\ell(A) = M(\ell, k_0, R) = M(\ell, k, R)$, so that (G, A) is optimal.

In what follows, we shall assume that $\mathrm{char}(k) = 0$. Note also that we could replace k by any subfield having the same invariants t and m, for instance the intersection of k with the field of ℓ^∞-roots of unity. We could thus assume that *k is a cyclotomic number field*, if needed.

The proof of Theorem 9 will be given first for classical groups (§9), by explicit elementary constructions similar to those of Schur. The more interesting case of exceptional groups (§12) will use different methods, based on Galois twists (§10), Tits groups and braid groups (§11). The case of $\ell = 2$ will be given in §13. The last section (§14) is about $m = \infty$.

§8. Arithmetic methods ($k = \mathbf{Q}$)

These methods are not strong enough to prove the statements of §7, but they give very interesting special cases.

8.10. **Euler characteristics.** Here, the ground field is $\mathbf{Q}$. One starts from a split simply connected group scheme G over $\mathbf{Z}$ (this makes sense, cf. [SGA 3]). One may thus speak of the group $\Gamma = G(\mathbf{Z})$ of the *integral points* of G. It is a discrete subgroup of $G(\mathbf{R})$. Its Euler characteristic $\chi(\Gamma)$ ("caractéristique d'Euler-Poincaré" in French) is well-defined (see [Se 71] and [Se 79]); it is a rational number. Moreover it is proved in [Ha 71] that

$$\chi(\Gamma) = c \prod_{i=1}^{r} \frac{1}{2}\zeta(1 - d_i) = c \prod_{i=1}^{r} \frac{b_{d_i}}{2d_i}, \tag{8.10.1}$$

where b_d is the d-th Bernoulli number, ζ is the zeta function and $c = |W|/|W_K|$ where W_K is the Weyl group of a maximal compact subgroup K of $G(\mathbf{R})$. Assume that all d_i's are *even* (if not, all the terms in (8.10.1) are zero). Using standard properties of Bernoulli numbers, one can check that *the* M*-bound relative to ℓ is* $M = \sum_i v_\ell\big(\mathrm{den}\,\big(\frac{1}{2}\zeta(1-d_i)\big)\big)$, where "den" means denominator. Hence, if ℓ does not divide c, and does not divide the numerator of any $\frac{1}{2}\zeta(1 - d_i)$ (which is the case if ℓ is a so-called regular prime), one sees that *the denominator of* $\mathrm{EP}(\Gamma)$ *is divisible by* ℓ^M. But a theorem of K. Brown [Br 74] shows that this is only possible if Γ contains a finite subgroup of order ℓ^M. Hence we get an optimal pair (provided $(c, \ell) = 1$, and ℓ is regular, say).

Example. Take G of type E_8; here $c = 3^3 \cdot 5$, and the numerators of the $\frac{1}{2}\zeta(1 - d_i)$ do not cancel any denominator. Hence one obtains that a split E_8 contains an optimal A for all $\ell \neq 3, 5$, with the extra information that A can be found inside the group $\Gamma = G(\mathbf{Z})$ – but no information on what it looks like!

8.11. **Mass formulae.** In [Gr 96], B. Gross considers **Q**-forms of G such that $G(\mathbf{R})$ is *compact*; he also requires another condition which guarantees that G has a *smooth model over* $\mathbf{Z}$. This condition is fulfilled for types B, D, G_2, F_4 and E_8. He then proves a *mass formula* à la Minkowski ([Gr 96], prop.2.2):

$$\sum \frac{1}{|A_\sigma|} = \prod_{i=1}^{r} \frac{1}{2}\zeta(1-d_i)$$

where the A_σ are the $\mathbf{Z}$-points of the smooth models of G over $\mathbf{Z}$ (taken up to conjugation). Each A_σ is finite. It is then clear that, if ℓ^N is the ℓ-th part of the denominator of $\prod_{i=1}^r \frac{1}{2}\zeta(1-d_i)$, the ℓ-Sylow subgroup of one of the A_σ has order $\geq \ell^N$. If N is equal to the Minkowski bound M (which happens if ℓ does not divide the numerator of any of the $\frac{1}{2}\zeta(1-d_i)$), then such a Sylow has order ℓ^M, and we get an optimal pair. Note that there is no extra factor "c" as in (8.10.1). This works very well for G_2, F_4, E_8 (and some classical groups too, cf. [Gr 96]):

G_2 - Here the mass is $\frac{1}{4}\zeta(-1)\zeta(-5) = \dfrac{1}{2^6 3^3 7}$, and it is obtained with just one A_σ, which turns out to be isomorphic to $G_2(\mathbf{F}_2)$.

F_4 - There are two A_σ's and the mass formula is

$$\begin{aligned} \frac{1}{2^{15}\cdot 3^6\cdot 5^2\cdot 7} + \frac{1}{2^{12}\cdot 3^5\cdot 7^2\cdot 13} &= \tfrac{1}{16}\zeta(-1)\zeta(-5)\zeta(-7)\zeta(-11) \\ &= \frac{691}{2^{15}\cdot 3^6\cdot 5^2\cdot 7^2\cdot 13}. \end{aligned}$$

E_8 - Here the numerator is very large, but the denominator is exactly what is needed for the M-bound, namely:

$$2^{30}\cdot 3^{13}\cdot 5^5\cdot 7^4\cdot 11^2\cdot 13^2\cdot 19\cdot 31.$$

§9. Proof of theorem 9 for classical groups

Here $\ell \neq 2$. Recall that $\operatorname{Im}\chi_{\ell^\infty} = C_t \times \{1+\ell^m\mathbf{Z}_\ell\}$, where $m \geq 1$ and t divides $\ell - 1$. The M-bound is

$$M = \sum_{\substack{i \\ d_i \equiv 0 (\mathrm{mod}\, t)}} \big(m + v_\ell(d_i)\big).$$

We denote by K the field $k(z_\ell)$ generated by a root of unity of order ℓ. It is a cyclic extension of k, of degree t, with Galois group C_t. It contains z_{ℓ^m} but not $z_{\ell^{m+1}}$, cf. §4.1.

9.1. **The groups A_N and A_N^1.** If N is an integer ≥ 1, we denote by A_N the subgroup of $\mathbf{GL}_N(K)$ (where $K = k(z_\ell)$ as above) generated by the symmetric group S_N and the diagonal matrices whose entries are ℓ^m-th roots of unity (wreath product of S_N with a cyclic group of order ℓ^m). We have

$$v_\ell(A_N) = mN + v_\ell(N!). \tag{9.1.1}$$

The image of $\det_K : A_N \to K^*$ is $\{\pm 1\} \times \langle z_{\ell^m} \rangle$. Hence the kernel A_N^1 is such that

$$v_\ell(A_N^1) = m(N-1) + v_\ell(N!). \tag{9.1.2}$$

We are going to use A_N, and sometimes A_N^1, in order to construct optimal subgroups for the classical groups $\mathbf{SL}_n$, $\mathbf{SO}_n$ and $\mathbf{Sp}_n$; this is what Schur did in [Sch 05], §6, for the case of $\mathbf{GL}_n$.

9.2. **The case of $\mathbf{SL}_n$.** The d_i's are 2, 3,..., n. If we put $N = \left[\frac{n}{t}\right]$, we have

$$M = mN + v_\ell(N!) \qquad \text{if } t \geq 2, \tag{9.2.1}$$

$$M = m(N-1) + v_\ell(N!) \qquad \text{if } t = 1, \text{ in which case } N = n. \tag{9.2.2}$$

In the case $t \geq 2$, we take $A_N \subset \mathbf{GL}_N(K) \subset \mathbf{GL}_{Nt}(k)$, and observe that $\det_k(A_N)$ is equal to ± 1 (indeed, if $g \in A_N$, then $\det_k(g) = N_{K/k}\big(\det_K(g)\big)$ and one checks that $N_{K/k}\big(z_{\ell^m}\big) = 1$). This shows that an ℓ-Sylow of A_N is contained in $\mathbf{SL}_{Nt}(k)$ and hence in $\mathbf{SL}_n(k)$. By (9.2.1) we get an optimal pair.

In the case $t = 1$, we use the same construction with A_N^1 instead of A_N. The comparison of (9.1.2) and (9.2.2) shows that we get an optimal pair.

9.3. **The case of the orthogonal and symplectic groups, t odd.** Let us consider the case of $\mathbf{Sp}_{2n}$. The d_i's are equal to $2, 4, \ldots, 2n$. Hence, if we put $N = \left[\frac{n}{t}\right]$, the M-bound is $mN + v_\ell(N!)$. There is a natural embedding:

$$\mathbf{GL}_N \to \mathbf{Sp}_{2N} \to \mathbf{Sp}_{2n}$$

defined by $x \mapsto \begin{pmatrix} x & 0 \\ 0 & {}^t x^{-1} \end{pmatrix}$. The image of A_N by that embedding is optimal.

The same construction works for $\mathbf{SO}_{2n}$ and $\mathbf{SO}_{2n+1}$. (Note that, in all these cases, we get the *split* forms of the groups of type B_n, C_n, D_n. This is no longer true in the case t is even – nor in the cases of §12.)

9.4. **The case of the orthogonal and symplectic groups, t even.** Since t is even, the group $C_t = \mathrm{Gal}(K/k)$ contains an element σ of order 2; its image in $\mathbf{Z}_\ell^*$ is -1. Let K_0 be the subfield of K fixed by σ; we have $[K : K_0] = 2$, $[K_0 : k] = t_0$ with $t_0 = t/2$. Moreover $\sigma(z_{\ell^m})$ is equal to $(z_{\ell^m})^{-1}$; i.e. σ acts on z_{ℓ^m} just as complex conjugation does. Let us define an *hermitian form* h on K^N (where N is a given integer ≥ 1) by the standard formula

$$h(x,y) = \sum_{i=1}^{N} x_i \cdot \sigma(y_i), \qquad \text{if } x = (x_1, \ldots, x_N),\ y = (y_1, \ldots, y_N).$$

If $\mathbf{U}_N$ denotes the *unitary group* associated with h, it is clear that *the group A_N defined in §9.1 is contained in $\mathbf{U}_N(K)$.* [We use here the traditional notation $\mathbf{U}_N(K)$ for the unitary group; this is a bit misleading, since $\mathbf{U}_N$ is an algebraic group over K_0, and we are taking its K_0-points.]

Let $\delta \in K^*$ be such that $\sigma(\delta) = -\delta$, e.g. $\delta = z_\ell - z_\ell^{-1}$. We have $K = K_0 \oplus \delta \cdot K_0$, and $h(x,y)$ can be decomposed as

$$h(x,y) = q_0(x,y) + \delta \cdot b_0(x,y), \text{ with } q_0(x,y) \in K_0, \quad b_0(x,y) \in K_0.$$

Then q_0 (resp. b_0) is a non-degenerate symmetric (resp. alternating) K_0-bilinear form of rank $2N$.
Its trace $q = \mathrm{Tr}_{K_0/k}\, q_0(x,y)$ (resp. $b = \mathrm{Tr}_{K_0/k}\, b_0(x,y)$) is of rank $2Nt_0 = Nt$ over k. We thus get embeddings:

$$A_N \to \mathbf{U}_N(K) \to \mathbf{SO}_{2N}(K_0) \to \mathbf{SO}_{Nt}(k) \tag{9.4.1}$$

$$A_N \to \mathbf{U}_N(K) \to \mathbf{Sp}_{2N}(K_0) \to \mathbf{Sp}_{Nt}(k). \tag{9.4.2}$$

Now, for a given n, let us define N by $N = \left[\frac{2n}{t}\right] = \left[\frac{n}{t_0}\right]$. By (9.4.2), we get an embedding

$$A_N \to \mathbf{Sp}_{Nt}(k) \to \mathbf{Sp}_{2n}(k),$$

and one checks that it is optimal.

The same method gives an embedding of A_N into $\mathbf{SO}_{Nt}(k)$, hence into $\mathbf{SO}_{2n+1}(k)$, and this embedding is also optimal. As for $\mathbf{SO}_{2n}(k)$, one has to be more careful. The method does give an embedding of A_N into the $\mathbf{SO}_{2n}$ group relative to some quadratic form Q, but we have to ensure that such an $\mathbf{SO}_{2n}$ group is *of inner type* i.e. that $\mathrm{disc}(Q) = (-1)^n$ in k^*/k^{*2}. There are three cases:

a) If $2n > Nt$ (i.e. if t does not divide $2n$), we choose $Q = q \oplus q_1$, where q_1 has rank $2n - Nt$, and is such that $\mathrm{disc}(q) \cdot \mathrm{disc}(q_1) = (-1)^n$. We then have $A_N \subset \mathbf{SO}_{2n,Q}(k)$ and this is optimal.

b) If $2n = Nt$ and N is even, we have $\mathrm{disc}(q) = d^N$, where $d = \mathrm{disc}(K_0/k)$, hence $\mathrm{disc}(q) = 1$ in k^*/k^{*2}, which is the same as $(-1)^n$ since n is even.

c) If $2n = Nt$ and N is odd, we use an optimal subgroup A of $\mathbf{SO}_{2n-1}(k)$ relative to a quadratic form q_0 of rank $2n-1$. By adding to q_0 a suitable quadratic form of rank 1, we get a quadratic form of rank $2n$ and discriminant $(-1)^n$, as wanted. The corresponding embedding

$$A \to \mathbf{SO}_{2n-1}(k) \to \mathbf{SO}_{2n}(k)$$

is optimal. (Note that the d_i's for type D_n are $2, 4, \ldots, 2n-2$, and n. Hence, if $t \nmid n$, the M-bound for D_n is the same as the M-bound for B_{n-1}.)

§10. Galois twists

To handle exceptional groups, we have to use *twisted* inner forms instead of split ones. We shall only need the most elementary case of twisting, namely the one coming from a homomorphism $\varphi : \Gamma_k \to \mathrm{Aut}(G)$. Let us recall what this means (cf. for example [Se 64], chapter III):

Let K/k be a finite Galois extension. Let X be an algebraic variety over k, assumed to be quasi-projective (the case where X is affine would be enough). Choose a homomorphism

$$\varphi : \mathrm{Gal}(K/k) \to \mathrm{Aut}_k\, X.$$

The *twist* X_φ of X by φ is a variety over k which can be characterized as follows:

There is a K-isomorphism $\theta : X_{/K} \to X_{\varphi/K}$ such that $\gamma(\theta) = \theta \circ \varphi(\gamma)$ for every $\gamma \in \mathrm{Gal}(K/k)$.

(Here $X_{/K}$ denotes the K-variety deduced from X by the base change $k \to K$, and $\varphi(\gamma) \in \mathrm{Aut}_k\, X$ is viewed as belonging to $\mathrm{Aut}_K\, X_{/K}$.)

One shows (as a special case of Galois descent) that such a pair (X_φ, θ) exists, and is unique, up to isomorphism.

It is sometimes convenient to identify the K-points of X and X_φ via the isomorphism θ. But one should note that this is not compatible with the natural action of $\mathrm{Gal}(K/k)$ on $X(K)$ and $X_\varphi(K)$; one has

$$\gamma\big(\theta(x)\big) = \varphi(\gamma)\big(\gamma(x)\big) \qquad \text{if } \gamma \in \mathrm{Gal}(K/k), x \in X(K).$$

In other words, if we identify $X_\varphi(K)$ with $X(K)$, an element γ of $\mathrm{Gal}(K/k)$ acts on $X_\varphi(K)$ by the *twisted action* :

$$x \mapsto \varphi(\gamma)\big(\gamma(x)\big)$$

In particular, the *k-rational points of X_φ* correspond (via θ^{-1}) to the points $x \in X(K)$ such that $\gamma(x) = \varphi(\gamma^{-1})x$ for every $\gamma \in \mathrm{Gal}(K/k)$.

In what follows we apply the φ-twist to $X =$ split form of G, with $\varphi(\gamma)$ being a k-automorphism of G for every $\gamma \in \mathrm{Gal}(K/k)$. In that case, G_φ is a k-form of G; this form is inner if all $\varphi(\gamma)$ belong to $G^{\mathrm{ad}}(K)$ where G^{ad} is the adjoint group of G. The effect of the twist is to make k-rational some elements of G which were not. In order to define φ, we shall have to use the k-automorphisms of G provided by the Tits group W^*, see next section.

§11. A general construction

Here, G is a split simply connected group over k, and T is a maximal split torus of G. We put $N = N_G(T)$ and $W = N/T$ is the Weyl group.

11.1. **The Tits group.** The exact sequence $1 \to T \to N \to W \to 1$ does not split in general. However Tits ([Ti 66a], [Ti 66b]) has shown how to construct a subgroup[3] W^* of $N(k)$ having the following properties:

(1) The map $W^* \to W$ is surjective.

(2) The group $W^* \cap T$ is equal to the subgroup T_2 of T made up of the points x of T with $x^2 = 1$.

We thus have a commutative diagram, where the vertical maps are inclusions:

$$\begin{array}{ccccccccc} 1 & \to & T_2 & \to & W^* & \to & W & \to & 1 \\ \downarrow & & \downarrow & & \downarrow & & \downarrow & & \\ 1 & \to & T & \to & N & \to & W & \to & 1 \end{array}$$

We refer to Tits (*loc. cit.*) and to Bourbaki[4] ([LIE X], pp. 115–116, exerc. 12, 13) for the construction and the properties of W^*. For instance:

If G comes from a split group scheme $\underline{G}$ over $\mathbf{Z}$, then W^* is equal to $\underline{N}(\mathbf{Z})$, the group of *integral points* of the group scheme $\underline{N}$.

[3]The construction of W^* depends on more than (G, T): one needs a *pinning* ("épinglage") of (G, T) in the sense of [SGA 3], XXIII.1.1.

[4]Bourbaki works in the context of compact real Lie groups; his results can easily be translated to the algebraic setting we use here.

In the case of $\mathbf{SL}_n$, this means that one can choose for W^* the group of monomial matrices with non-zero entries ± 1 and determinant 1. For $n = 2$, W^* is the cyclic group of order 4 generated by $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

Note also that W^* is a quotient of the *braid group* $\mathbf{B}_W$ associated to W. (For the definition of the braid group of a Coxeter group, see e.g. [BM 97].)

11.2. **Special elements of W.** We now go back to our general notation $\ell, m, t, \ldots$ of Lecture II. Recall that the M-bound $M = M(\ell, k, R)$ is given by

$$M = \sum_{t|d_i} (m + v_\ell(d_i)), \qquad \text{cf. §6.2.} \tag{11.2.1}$$

Let $a(t)$ be the number of indices i such that $d_i \equiv 0 \pmod{t}$. We may rewrite (11.2.1) as

$$M = ma(t) + \sum_{t|d_i} v_\ell(d_i). \tag{11.2.2}$$

Note that, if no d_i is divisible by t, we have $M = 0$ and the trivial group $A = 1$ is optimal. Hence *we shall assume in what follows that* $a(t) \geq 1$.

Let now w be an element of W. We shall say that w is *special* (with respect to t and ℓ) if it has the following four properties:
(1) w has order t in W.
(2) w is the image of an element w^* of W^* such that $(w^*)^t \in T_2 \cap C(G)$, where $C(G)$ is the center of G.
(3) The characteristic polynomial of w (in the natural r-dimensional representation of W) is divisible by $(\Phi_t)^{a(t)}$, where Φ_t is the t-th cyclotomic polynomial.
(Equivalently: if z_t denotes a primitive t-th root of unity, then z_t is an eigenvalue of w of multiplicity at least $a(t)$.)
(4) Let $C_W(w)$ be the centralizer of w in W. Then:

$$v_\ell(C_W(w)) \geq \sum_{t|d_i} v_\ell(d_i).$$

Remark. The reader may wonder whether special elements exist for a given pair (t, ℓ) (with $a(t) > 0$ and $\ell \equiv 1 \pmod{t}$, of course). The answer is "no" in general: if R is of type C_3 and $t = 4$, no element of W^* has both properties (1) and (2). Fortunately, the answer is "yes" for the exceptional types $G_2, \ldots, E_8$, cf. §12.

Example : the regular case. Suppose that $w \in W$ is *regular of order* t in the sense of Springer[5] ([Sp 74], bottom of p. 170 - see also [BM 97], §3). This means that w has an eigenvector v, with eigenvalue z_t, such that v does not belong to any reflecting hyperplane. *Then w is special* (for any ℓ with $\ell \equiv 1 \pmod{t}$). Indeed:

[5]With a slight difference: Springer requires $t > 1$ and we don't; it is convenient to view $w = 1$ as a regular element of W.

Note that, if t is given, there is a very simple criterion ensuring the existence of a regular element of W of order t: the number of indices i such that $d_i \equiv 0 \pmod{t}$ should be equal to the number of i's such that $d_i \equiv 2 \pmod{t}$, cf. Lehrer-Springer [LS 99], cor.5.5.

(1) is obvious.

(2) follows from the fact, proved in [BM 97], §3, that w has a lifting $\mathbf{w}$ in the braid group $\mathbf{B}_W$ with $\mathbf{w}^t = \boldsymbol{\pi}$, where $\boldsymbol{\pi}$ has an image π in W which belongs to $T_2 \cap C(G)$. In Bourbaki's notation ([LIE X], p.116) π is the canonical element z_G of the center of G.

(3) is proved in [Sp 74], th. 4.2.

(4) is proved in [Sp 74], th. 4.2, in the stronger form $|C_W(w)| = \prod_{t|d_i} d_i$.

Special cases

$t = 1$. Here $w = 1$ and $w^* = \pi$ (one could also take $w^* = 1$).

$t = 2$. Here $w = w_0 =$ longest element of W. When -1 belongs to W, one has $w_0 = -1$ and w_0^* is *central* in W^* (because $\mathbf{w}_0$ is central in $\mathbf{B}_W$, cf. [BM 97], 1.2 and 3.4). In that case the inner automorphism of G defined by w_0^* is a "Weyl-Chevalley involution": it acts on T by $t \mapsto t^{-1}$.

11.3. **An auxiliary result.**

Lemma 7. *Suppose $w \in W$ is special of order t. Then it is possible to choose a lifting w^* of w in W^* which satisfies*

(2*) $(w^*)^t \in T_2 \cap C(G)$

and

(4*) $v_\ell\,(C_{W^*}(w^*)) \geq \sum_{t|d_i} v_\ell(d_i)$.

Proof. Let P be an ℓ-Sylow of $C_W(w)$; the groups P and $\langle w\rangle$ commute, and $P \cap \langle w\rangle = 1$ since w has order t and ℓ is prime to t (since $\ell \equiv 1 \bmod t$). Hence the group P_w generated by w and P is the direct product $P \times \langle w\rangle$. Since $\ell \neq 2$, its 2-Sylow subgroup is contained in $\langle w\rangle$. Put $C_2 = T_2 \cap C(G)$. We have an exact sequence:

$$1 \to T_2/C_2 \to W^*/C_2 \to W \to 1.$$

By property (2) of w, this exact sequence splits over $\langle w\rangle$, hence over the 2-Sylow of P_w; since the order of T_2/C_2 is a power of 2, this implies that it splits over P_w. We thus get an element w' of W^*/C_2, of order t, which lifts w, and centralizes a subgroup P' of W^*/C_2 isomorphic to P. We then choose for w^* a representative of w' in W^*; it has property (2*), moreover its centralizer contains the inverse image of P', which is canonically isomorphic to $C_2 \times P'$. By property (4) we have

$$v_\ell(P') = v_\ell(P) \geq \sum_{t|d_i} v_\ell(d_i).$$

This shows that w^* has property (4*). □

Remark. In the case where w is regular, one can do without lemma 7. Indeed the braid group construction of [BM 97] gives a lifting w^* of w having property (2*) and such that the map $C_{W^*}(w^*) \to C_W(w)$ is surjective.

11.4. **The main result.**

Proposition 5. *Suppose W contains an element w which is special with respect to t and ℓ. Then there exist an inner twist G_φ of G* (cf. §10) *and a finite ℓ-subgroup A of $G_\varphi(k)$ such that the pair (G_φ, A) is optimal in the sense of* §7.

(In particular, th.9 is true for (k, ℓ, R).)

Proof. As in §9, we put $K = k(z_\ell)$, where z_ℓ is a root of unity of order ℓ. Let $C_t = \mathrm{Gal}(K/k)$; it is a cyclic group of order t.

Choose $w^* \in W^*$ with the properties of lemma 7 and let σ be the inner automorphism of G defined by w^*. Since σ has order t, there exists an injective homomorphism:

$$\varphi : C_t \to G^{\mathrm{ad}}(k) \subset \mathrm{Aut}_k(G)$$

which maps C_t onto the subgroup $\langle\sigma\rangle$ of $\mathrm{Aut}_k(G)$ generated by σ. As explained in §10, we may then define the *φ-twist* G_φ of G, relatively to the Galois extension K/k. The group G_φ is an inner form of G; it has the same root system R. It remains to construct a finite ℓ-subgroup A of $G_\varphi(k)$ such that (G_φ, A) is optimal, i.e. $v_\ell(A) = ma(t) + \sum_{t|d_i} v_\ell(d_i)$, cf. (11.2.2).

We take for A the semi-direct product $E_m \cdot P$, with $E_m \subset T_\varphi(k)$ and $P \subset N_\varphi(k)$, where E_m and P are defined as follows:

(1) P is an ℓ-Sylow of $C_{W^*}(w^*)$. By lemma 7 we have $v_\ell(P) \geq \sum_{t|d_i} v_\ell(d_i)$. Note that the points of P are fixed by σ. Hence these points are rational over k not only in the group G but also in the group G_φ.

(2) E_m is the subgroup of $T_\varphi(k)$ made up of the elements x such that $x^{\ell^m} = 1$.

It is clear that P normalizes E_m, and that $P \cap E_m = 1$.

Lemma 8. *The group E_m contains a product of $a(t)$ copies of the group $\mathbf{Z}/\ell^m\mathbf{Z}$.*

This implies that $v_\ell(E_m) \geq ma(t)$ and hence

$$v_\ell(A) = v_\ell(E_m) + v_\ell(P) \geq ma(t) + \sum_{t|d_i} v_\ell(d_i).$$

We thus get $v_\ell(A) \geq M$ and since M is an upper bound for $v_\ell(A)$ we have $v_\ell(A) = M$. □

Proof of lemma 8. Consider first the subgroup T_{ℓ^m} of $T(k_s)$ made up of the elements x with $x^{\ell^m} = 1$. Since T is k-split, and $K = k(z_\ell) = k(z_{\ell^m})$ (cf. §4 and §9), the points of T_{ℓ^m} are rational over K. If we write $T_{\ell^m}(K)$ additively, it becomes a free $\mathbf{Z}/\ell^m\mathbf{Z}$-module of rank r and the action of a generator s of C_t is by $x \mapsto sx$, where s is identified with an element of order t in $\mathbf{Z}_\ell^*$ (i.e. $s =$"z_t" with our usual notation for roots of unity). As for the action of w^* (i.e. of w) on $T_{\ell^m}(K)$, it can be put in diagonal form since w is of order t and t divides $\ell - 1$; its diagonal elements are r elements $y_1, \ldots, y_r$ of $\mathbf{Z}/\ell^m\mathbf{Z}$, with $y_i^t = 1$. Let c be the largest integer such that $(\Phi_t)^c$ divides the characteristic polynomial of w. By property (3) of 11.2, we have $c \geq a(t)$ (in fact, $c = a(t)$, by [Sp 74], th. 3.4). This implies that the family of the y_i's contains c times each primitive t-th root of unity (viewed as element of $(\mathbf{Z}/\ell^m\mathbf{Z})^*$). In particular, there is a $\mathbf{Z}/\ell^m\mathbf{Z}$-submodule X of $T_{\ell^m}(K)$ which is free of rank c and on which w acts by $x \mapsto z_t^{-1}x$. If we twist G, T, T_{ℓ^m} by φ, the new action of $C_t = \mathrm{Gal}(K/k)$ on X is trivial (cf. end of §10). This means that X is contained in $T_\varphi(k)$, hence in E_m, which proves the lemma. □

Note the following consequence of proposition 4:

Corollary. *If W contains a t-regular element in the sense of* [Sp 74]*, then theorem 9 is true for k, ℓ, R.*

In the case $t = 1$, no twist is necessary (one takes $w = 1$, $w^* = 1$, cf. §11.2).

§12. Proof of theorem 9 for exceptional groups

In each case we will show that the Weyl group contains an element w which is special with respect to t and ℓ, so that we may apply prop.5.

12.1. **The case of G_2.** The degrees d_i are $d_1 = 2, d_2 = 6$. Since t divides one of them, t is a divisor of 6, hence is regular ([Sp 74], no. 5.4). We may then apply prop.5. □

Explicit description of w, w^*: if c is a Coxeter element of W, c is of order 6, and every lifting c^* of c in W^* has order 6. Hence, for any divisor t of 6, we may take $w = c^{6/t}$ and $w^* = (c^*)^{6/t}$.

12.2. **The case of F_4.** The d_i's are: 2, 6, 8, 12. All their divisors are regular (Springer, *loc. cit.*). One concludes as for G_2. □

12.3. **The case of E_6.** The d_i's are: 2, 5, 6, 8, 9, 12. All their divisors are regular, except $t = 5$. In that case, choose any element $w \in W$ of order 5. Since the kernel of $W^* \to W$ is a 2-group, w can be lifted to an element w^* of W^* of order 5. Conditions (1) and (2) of §11.2 are obviously satisfied. The same is true for condition (3), since $a(5) = 1$ (only one of the d_i's is divisible by 5), and w has at least one eigenvalue of order 5. As for condition (4), it is trivial, since $\ell \equiv 1 \pmod 5$ implies $\ell \geq 11$, and ℓ does not divide any of the d_i's, so that $\sum_{t|d_i} v_\ell(d_i)$ is 0. Hence w is special with respect to $(5, \ell)$. □

12.4. **The case of E_7.** The d_i's are: 2, 6, 8, 10, 12, 14, 18. By [Sp 74], *loc.cit.* all their divisors are regular except 4, 5, 8, 10, 12. If $t = 4, 5, 8$ or 12, t already occurs for E_6, with the same values of $a(t)$, namely 2, 1, 1 and 1. Hence, we have E_6-special elements w_4, w_5, w_8 and w_{12} in $W(E_6)$. One then takes their images in $W(E_7)$ by the injective map $W(E_6) \to W(E_7)$, and one checks that they are E_7-special (here again condition (4) is trivial since $v_\ell(d_i) = 0$ for all the ℓ's with $\ell \equiv 1 \pmod t$).

As for $t = 10$, one takes $w = -w_5$, which makes sense since $-1 \in W$. The element -1 (usually denoted by w_0) can be lifted to a central element ε of W^* with $\varepsilon^2 \in T_2 \cap C(G)$; this is a general property of the case $-1 \in W$ (which reflects the fact that -1 is 2-regular, see end of §11.2). Hence, if w_5^* is a lifting of w_5 of order 5, εw_5^* is a lifting of w of order 10, and this shows that w is special with respect to 10 and ℓ. □

12.5. **The case of E_8.** The d_i's are: 2, 8, 12, 14, 18, 20, 24, 30. By [Sp 74], *loc.cit.*, all their divisors are regular except 7, 9, 14, 18.

If $t = 7$ (resp. 9), one chooses $w_7 \in W$ of order 7 (resp. $w_9 \in W$ of order 9). Since 7 and 9 are odd, condition (2) of §11.2 is satisfied. The same is true for condition (3) because $a(t) = 1$, and for condition (4) because $v_\ell(d_i) = 0$ for all i.

If $t = 14$ (resp. 18), one takes $w = -w_7$ (resp. $w = -w_9$), as we did for E_7. □

§13. Proof of theorems 10 and 11

Here $\ell = 2$. There are three cases (cf. §4.2):

(a) $\operatorname{Im}\chi_{2^\infty} = 1 + 2^m\mathbf{Z}_2$ with $m \geq 2$. In that case the M-bound is $rm + v_2(W)$, and th.10 asserts that an optimal pair (G, A) exists for every type R.

(b) $\operatorname{Im}\chi_{2^\infty} = \langle -1 + 2^m \rangle$, with $m \geq 2$. The M-bound is $r_0 m + r_1 + v_2(W)$, where r_0 (resp. r_1) is the number of i's such that d_i is odd (resp. even). Here, too, th.10 asserts that an optimal pair exists.

(c) $\operatorname{Im}\chi_{2^\infty} = \langle -1, 1 + 2^m \rangle$, with $m \geq 2$.

The M-bound is the same as in case (b), but th.11 does not claim that it can be met (i.e. that an optimal pair exists); it merely says that there is a pair (G, A) with $v_2(A) = r_0 m + v_2(W)$; such a pair is optimal only when $r_1 = 0$, i.e. when -1 belongs to the Weyl group.

13.1. **Proof of theorem 10 in case (a).** We take G split and simply connected, and we choose a maximal split torus T. We use the notation (N, W, W^*) of §11. Let E be the 2-torsion subgroup of $T(k)$. Since T is isomorphic to the product of r copies of $\mathbf{G}_m$, E is isomorphic to a product of m copies of $\mathbf{Z}/2^m\mathbf{Z}$, cf. §4.2. Hence $v_2(E) = rm$. The group E is normalized by the Tits group W^*; we define A as $A = E \cdot W^*$. The exact sequence

$$1 \to E \to A \to W \to 1$$

shows that $v_2(A) = rm + v_2(W)$. Hence (G, A) is optimal.

13.2. **Cases (b) and (c).** As in §11.4, we start with a split G, with a split maximal torus T. We define N, W, W^* as usual. After choosing an order on the root system R, we may view W as a Coxeter group; let w_0 be its longest element. It has order 2, and it is regular in the sense of Springer [Sp 74]. As explained in §11.2, this implies that there is a lifting w_0^* of w_0 in W^* which has the following two properties:

(i) its square belongs to the center of G;

(ii) the natural map $C_{W^*}(w_0^*) \to C_W(w_0)$ is surjective.

Let σ be the inner automorphism of G defined by w_0^*. By (i), we have $\sigma^2 = 1$. Let $K = k(i)$ and let φ be the homomorphism of $\operatorname{Gal}(K/k)$ into $\operatorname{Aut}_k(G)$ whose image is $\{1, \sigma\}$. Let us define G_φ as the *φ-twist* of G, in the sense defined in §10. Denote by T_φ, N_φ and W_φ^* the φ-twists of T, N and W^*. We have an exact sequence

$$1 \to T_\varphi \to N_\varphi \to W_\varphi \to 1,$$

where W_φ is the φ-twist of W. Note that $W_\varphi(k)$ is equal to the centralizer $C_W(w_0)$ of w_0 in W, and similarly $W_\varphi^*(k)$ is equal to $C_{W^*}(w_0^*)$.

As in §13.1, let E be the 2-torsion subgroup of $T_\varphi(k)$. It is normalized by $C_{W^*}(w_0^*)$. Define $A \subset G_\varphi(k)$ to be the group $A = E \cdot C_{W^*}(w_0^*)$. By (ii), we have an exact sequence:

$$1 \to E \to A \to C_W(w_0) \to 1\,,$$

which shows that $v_2(A) = v_2(E) + v_2\big(C_W(w_0)\big)$. The fact that w_0 is regular of order 2 implies that

$$|C_W(w_0)| = \prod_{2|d_i} d_i,$$

hence $v_2\big(C_W(w_0)\big) = \sum v_2(d_i) = v_2(W)$. This gives:

$$v_2(A) = v_2(E) + v_2(W). \tag{13.2.1}$$

Proposition 6. *We have* :

$v_2(E) = r_1 + r_0 m$ *in case* (b)
$v_2(E) = r_0 m$ *in case* (c).

In case (b), this shows that (G_φ, A) is optimal, which proves th.10. Similarly, the fact that $v_2(A) = r_0 m + v_2(W)$ proves th.11 in case (c).

13.3. **Proof of proposition 6.** We need to describe explicitly the torus T_φ. To do so, let us first define the following two tori:

$\mathbf{G}_m^\sigma$ = 1-dimensional torus deduced from $\mathbf{G}_m$ by Galois twist relatively to K/k. Its group of k-points is $K_1^* = \operatorname{Ker} N_{K/k} : K^* \to k^*$.

$R_{K/k}\mathbf{G}_m$ = 2-dimensional torus deduced from $\mathbf{G}_m$ by Weil's restriction of scalars relatively to K/k. Its group of k-points is K^*.

Lemma 9. *The torus T_φ is isomorphic to the product of r_1 copies of $R_{K/k}\mathbf{G}_m$ and $r_0 - r_1$ copies of $\mathbf{G}_m^\sigma$.*

Proof. The character group $X = \operatorname{Hom}(T, \mathbf{G}_m)$ is free of rank r, with basis the fundamental weights $\omega_1, \ldots, \omega_r$. This gives a decomposition of T as

$$T = T_1 \times T_2 \times \cdots \times T_r\,,$$

where each T_i is canonically isomorphic to $\mathbf{G}_m$. Let $\tau = -w_0$ be the opposition involution of the root system R; it permutes $\omega_1, \ldots, \omega_r$ with r_1 orbits of order 2, and $r_0 - r_1$ orbits of order 1. (This follows from the fact that -1 is an eigenvalue of w_0 of multiplicity r_0.) The involution τ permutes the tori T_j. If an index j is fixed by τ, then w_0 acts on T_j by $t \mapsto t^{-1}$ and the twisted torus $(T_j)_\varphi$ is isomorphic to $\mathbf{G}_m^\sigma$; similarly, if τ permutes j and j', the torus $(T_j \times T_{j'})_\varphi$ is isomorphic to $R_{K/k}\mathbf{G}_m$. This proves lemma 9. □

End of the proof of prop.6. The 2-torsion subgroup of $\mathbf{G}_m^\sigma(k) = K_1^*$ is cyclic of order 2^m; the 2-torsion subgroup of $R_{K/k}\mathbf{G}_m(k) = K^*$ is cyclic of order 2^{m+1} in case (b) and of order 2^m in case (c). We get what we wanted, namely:

case (b): $v_2(E) = r_1(m+1) + (r_0 - r_1)m = r_0 m + r_1$
case (c): $v_2(E) = r_1 m + (r_0 - r_1)m = r_0 m$.

This completes the proof of prop.6, and hence of th.10 and th.11. □

13.4. **Remarks on the non simply connected case.** The proof above could have been given without assuming that the split group G is simply connected. The main difference is in lemma 9: in the general case, the torus T_φ is a product of three factors (instead of two):

$$T_\varphi = (\mathbf{G}_m)^\alpha \times (\mathbf{G}_m^\sigma)^\beta \times (R_{K/k}\mathbf{G}_m)^\gamma\,,$$

where α, β, γ are integers, with $\beta + \gamma = r_0$ and $\alpha + \gamma = r_1$. This gives the following formulae for $v_2(E)$:

case (b) : $v_2(E) = \alpha + \beta m + \gamma(m+1) = r_1 + r_0 m$

case (c) : $v_2(E) = \alpha + \beta m + \gamma m = \alpha + r_0 m$.

In case (b) one finds the same value for $v_2(A)$, namely the M-bound. In case (c) one finds a result which is intermediate between the M-bound $r_1 + r_0 m + v_2(W)$ and the value $r_0 m + v_2(W)$ given by th.11.

Examples (assuming we are in case (c)).

- *Type A_r, r even.* One finds that α is always 0, so that one does not gain anything by choosing non simply connected groups. Indeed, in that case, it is possible to prove, by a variant of Schur's method, that the value of $v_2(A)$ given by th.11 is best possible.

- *Type A_r, r odd ≥ 3.* Here $r_1 = (r-1)/2$. One finds that $\alpha = 0$ if $r \equiv 1 \pmod 4$, but that α can be equal to 1 if $r \equiv 3 \pmod 4$. When $r = 3$, we thus get $\alpha = r_1$; this shows that the M-bound is best possible for type A_3.

- *Type D_r, r odd.* Here $r_1 = 1$, and if one chooses G neither simply connected nor adjoint, one has $\alpha = 1$. This means that the orthogonal group $\mathbf{SO}_{2r}$ has an inner k-form which contains an optimal A. (Note the case $r = 3$, where $D_3 = A_3$.)

- *Type E_6.* Here $r_1 = 2$, and one has $\alpha = 0$ both for the simply connected group and for the adjoint group (indeed, α is 0 for every adjoint group). I do not know whether the bound of th.11 is best possible in this case.

§14. The case $m = \infty$

14.1. **Statements.** We keep the notation $(G, R, W, d_i, \ell, t, m)$ of §4 and §6; as before, we assume that G is of inner type.

We consider the case $m = \infty$, i.e. the case where *the image of χ_{ℓ^∞} is finite*; that image is then cyclic of order t, cf. §4.

Let $a(t)$ be the number of i's such that $d_i \equiv 0 \pmod t$. If $a(t) = 0$, then $G(k)$ is ℓ-torsion free, cf. §6.2, cor.to prop. 4. In what follows, we shall thus assume that $a(t) \geq 1$. In that case, $G(k)$ may contain infinite ℓ-subgroups (we say that a group is an *ℓ-group* if every element of that group has order a power of ℓ). The following two theorems show that $a(t)$ controls the size of such a subgroup:

Theorem 12. *Let A be an ℓ-subgroup of $G(k)$. Then A contains a subgroup of finite index isomorphic to the ℓ-group $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^a$, with $a \leq a(t)$.*

(Note that $\mathbf{Q}_\ell/\mathbf{Z}_\ell$ is the union of an increasing sequence of cyclic groups of order $\ell, \ell^2, \ldots$; it is the analogue of $\mathbf{Z}/\ell^m\mathbf{Z}$ for $m = \infty$.)

The bound $a \leq a(t)$ of th.12. is optimal. More precisely:

Theorem 13. *There exist a semisimple group G of inner type, with root system R, and an ℓ-subgroup A of $G(k)$, such that A is isomorphic to the product of $a(t)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.*

14.2. **Proof of theorem 12.** We need a few lemmas:

Lemma 10. *Any finitely generated ℓ-subgroup of $G(k)$ is finite.*

Proof. Let B be a finitely generated ℓ-subgroup of $G(k)$. We may embed B in $\mathbf{GL}_n(k)$ for n large enough. By a known result (see §1.2) there exists a subgroup B' of B, of finite index, which is torsion-free if $\mathrm{char}(k) = 0$, and has only p-torsion if $\mathrm{char}(k) = p$. Since B' is an ℓ-group, this means that $B' = 1$, hence B is finite. □

Lemma 11. *There exists a maximal k-torus of G which is normalized by A.* (Recall that A is an ℓ-subgroup of $G(k)$.)

Proof. Let F be the set of all finite subgroups of A, ordered by inclusion. Lemma 10 implies that, if B_1 and B_2 belong to F, so does $\langle B_1, B_2\rangle$. Let X be the k-variety parametrizing the maximal tori of G; it is a homogeneous space of G. If $B \in F$, let X^B be the subvariety of X fixed by B; a point of X^B corresponds to a maximal torus of G normalized by B. By the noetherian property of the scheme X, one may choose $B_0 \in F$ such that X^{B_0} is minimal among the X^B 's. If $B \in F$, then $X^{\langle B_0, B\rangle}$ is contained in X^{B_0}, hence equal to X^{B_0}. This shows that X^{B_0} is contained in all the X^B 's, i.e. that every maximal torus which is normalized by B_0 is normalized by all the B's, hence by A. By the corollary to th.3″ of §3.3 (applied to the finite ℓ-group B_0) there exists such a torus which is defined over k. □

Lemma 12. *Let $u \in \mathbf{M}_r(\mathbf{Z}_\ell)$ be an $r \times r$ matrix with coefficients in $\mathbf{Z}_\ell$, which we view as an endomorphism of $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^r$. Then* $\mathrm{Ker}(u)$ *has a subgroup of finite index isomorphic to the product of $r - \mathrm{rank}(u)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.*

In other words, the "corank" of $\mathrm{Ker}(u)$ is equal to $r - \mathrm{rank}(u)$.

Proof. Same as that of lemma 4 of §5.2: by reduction to the case where u is a diagonal matrix. □

Lemma 13. *Let z_t be a primitive t-th root of unity, and let w be an element of W. The multiplicity of z_t as an eigenvalue of w is $\leq a(t)$.*

Proof. See [Sp 74], th.3.4(i) where it is deduced from the fact that the polynomial $\det(t - w)$ divides $\prod_i (t^{d_i} - 1)$. □

Lemma 14. *Let T be a maximal k-torus of G, and let $T(k)_\ell$ be the ℓ-torsion subgroup of $T(k)$. We have* $\mathrm{corank}\ T(k)_\ell \leq a(t)$.

As above, the "corank" of a commutative ℓ-group is the largest n such that the group contains the product of n copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$.

Proof. As in §5.2, let $Y(T) = \mathrm{Hom}_{k_s}(\mathbf{G}_m, T)$ be the group of cocharacters of T. The action of the Galois group Γ_k on $Y(T)$ gives a homomorphism

$$\rho : \Gamma_k \to \mathrm{Aut}\, Y(T) \simeq \mathbf{GL}_r(\mathbf{Z})$$

and the image of ρ is contained in the Weyl group W (this is still another way of saying that G is of inner type). The group Γ_k acts on $T(k_s)_\ell \simeq (\mathbf{Q}_\ell/\mathbf{Z}_\ell)^r$ by $\rho \otimes \chi$, where $\chi = \chi_{\ell^\infty}$. Let us now choose $g \in \Gamma_k$ such that $\chi(g) = z_t^{-1}$, where z_t is an element of order t of $\mathbf{Z}_\ell^*$, and let $w = \rho(g)$. The element g acts on $T(k_s)_\ell$ by wz_t^{-1}. Let T_g be the kernel of $g - 1$ on $T(k_s)_\ell$. By lemma 12, we have $\mathrm{corank}\ (T_g) = r - \mathrm{rank}(g-1)$, which is equal to the multiplicity of z_t as an eigenvalue of w; using lemma 13, we get $\mathrm{corank}(T_g) \leq a(t)$, and since $T(k)_\ell$ is contained in T_g, we have $\mathrm{corank}(T(k)_\ell) \leq a(t)$. □

End of the proof of th.12. By lemma 11, there is a maximal k-torus T of G which is normalized by A. Let $A^\circ = A \cap T(k)$. Then A° is an abelian subgroup of A of finite index. Since A° is contained in $T(k)_\ell$, lemma 14 shows that A° is isomorphic to the product of a finite group with a product of at most $a(t)$ copies of $\mathbf{Q}_\ell/\mathbf{Z}_\ell$. □

14.3. **Proof of theorem 13.** We follow the same strategy as for theorem 9, 10 and 11. There are three cases:

14.3.1. **Classical groups** ($\ell \neq 2$). We change slightly the definitions of §9.1: we define A_N as the subgroup of $\mathbf{GL}_N(K)$, with $K = k(z_\ell)$, made up of the diagonal matrices of order a power of ℓ; it is isomorphic to $(\mathbf{Q}_\ell/\mathbf{Z}_\ell)^N$.

For any given $n \geq 2$, we put $N = [n/t]$ and we get embeddings

$$A_N \to \mathbf{GL}_N(K) \to \mathbf{GL}_{Nt}(k) \to \mathbf{GL}_n(k).$$

If $t > 1$, one checks that the k-determinant of every element of A_N is 1; we thus get an embedding $A_N \to \mathbf{SL}_n(k)$ which has the required properties since $N = a(t)$ in that case. When $t = 1$, we replace A_N by the subgroup of its elements of k-determinant 1, and we also get what we want. This solves the case of type A_r. Types B_r, C_r and D_r are then treated by the methods of §9.3 and §9.4.

14.3.2. **Exceptional groups** ($\ell \neq 2$). One replaces prop.5 of §11.4 by a statement giving the existence of $A \subset G_\varphi(k)$ with $A \simeq (\mathbf{Q}_\ell/\mathbf{Z}_\ell)^{a(t)}$. The proof is the same. One then proceeds as in §12.

14.3.3. **The case** $\ell = 2$. Same method as in §13. □

References

[A V] N. Bourbaki, *Algèbre, Chapitre* V, Masson, Paris, 1981.

[AC N] N. Bourbaki, *Algèbre Commutative, Chapitre* N, Hermann-Masson, Paris, 1961–1998.

[Bl 04] H. Blichfeldt, *On the order of linear homogeneous groups,* Trans. Amer. Math. Soc. **5** (1904), 310–325.

[BM 97] M. Broué and J. Michel, *Sur certains éléments réguliers des groupes de Weyl et les variétés de Deligne–Lusztig associées*, in Finite Reductive Groups: Related Structures and Representations, M. Cabanes (*edit.*), Progress in Math. **141**, Birkhäuser– Boston, 1997, 73–139.

[Bo 69] A. Borel, *Groupes arithmétiques*, Hermann, Paris 1969.

[Bo 91] A. Borel, *Linear Algebraic Groups*, second edition, Springer-Verlag, 1991.

[Br 74] K. Brown, *Euler characteristics of discrete groups and G-spaces,* Invent. math. **27** (1974), 229–264.

[Br 01] M. Broué, *Reflection groups, braid groups, Hecke algebras, finite reductive groups,* in Current Developments in Mathematics 2000, International Press, 2001, 1-107.

[BS 53] A. Borel and J-P. Serre, *Sur certains sous-groupes des groupes de Lie compacts*, Comm.Math.Helv. **27** (1953), 128–139 (= A. Borel, Coll. Works, vol.I, n°24).

[Bu 11] W. Burnside, *Theory of Groups of Finite Order,* second edition, Cambridge Univ.Press. 1911; reprinted by Dover Publ., 1955.

[De 70] M. Demazure, *Sous-groupes algébriques de rang maximum du groupe de Cremona,* Ann.scient.E.N.S. (4) **3** (1970), 507–588.

[EGA IV] A. Grothendieck, *Eléments de Géométrie Algébrique* (rédigés avec la collaboration de J. Dieudonné), Chap.IV, Etude Locale des Schémas et des Morphismes de Schémas (Troisième Partie), Publ.Math.I.H.E.S. **28** (1966).

[Fe 97] W. Feit, *Finite linear groups and theorems of Minkowski and Schur*, Proc. A.M.S. **125** (1997), 1259-1262.

[FW 84] G. Faltings, G. Wüstholz et al, *Rational Points*, Seminar Bonn-Wuppertal 1983/1984, Vieweg, Braunschweig, 1984.

[GL 06] R.M. Guralnick and M. Lorenz, *Orders of finite groups of matrices*, Contemp.Math., to appear.

[GMS 03] S. Garibaldi, A. Merkurjev and J.-P. Serre, *Cohomological Invariants in Galois Cohomology,* A.M.S. Lect. Series **28** (2003).

[Gr 96] B. H. Gross, *Groups over* $\mathbf{Z}$*,* Invent.math. **124** (1996), 263–279.

[Ha 71] G. Harder, *A Gauss-Bonnet formula for discrete arithmetically defined groups,* Ann.Sci. E.N.S. (4) **4** (1971), 409–455.

[Il 06] L. Illusie, *Miscellany on traces in ℓ-adic cohomology: a survey*, Jap. J.Math., (new series), **1** (2006), 107–136.

[LIE N] N. Bourbaki, *Groupes et Algèbres de Lie, Chapitre* N, Hermann-Masson, Paris 1972-1982.

[LS 99] G.I. Lehrer and T.A. Springer, *Reflection subquotients of unitary reflection groups*, Canadian J. Math. **51** (1999), 1175-1193.

[Mi 87] H. Minkowski, *Zur Theorie der positiven quadratischen Formen,* J.Crelle **101** (1887), 196–202 (= Ges.Abh., Band I, n°VI).

[Pi 97] R. Pink, *The Mumford-Tate conjecture for Drinfeld-modules,* Publ. Res. Inst. Math. Sci. **33** (1997), 393–425.

[Ro 58] P. Roquette, *Realisierung von Darstellungen endlicher nilpotenter Gruppen*, Arch.Math. **9** (1958), 241-250.

[Sch 05] I. Schur, *Über eine Klasse von endlichen Gruppen linearer Substitutionen*, Sitz.Preuss.Akad.Wiss. Berlin (1905), 77–91 (= Ges.Abh., Band I, n° 6).

[Se 64] J.-P. Serre, *Cohomologie Galoisienne*, Lect.Notes in Math. **5**, Springer- Verlag, 1964; fifth revised edition, 1994; English translation: *Galois Cohomology*, corrected second printing, Springer–Verlag, 2002.

[Se 65] J.-P. Serre, *Zeta and L functions*, in Arithmetical Algebraic Geometry (Proc. Conf. Purdue Univ. 1963), 82–92, Harper and Row, New York, 1965 (= Oe.64).

[Se 71] J.-P. Serre, *Cohomologie des groupes discrets,* Ann.of Math. Studies **70**, 77–169, Princeton, 1971 (= Oe.88).

[Se 79] J.-P. Serre, *Arithmetic groups,* in Homological Group Theory, C.T. Wall edit., LMS Lect.Notes Series **36**, Cambridge Univ.Press (1979), 105–136 (= Oe.120).

[Se 81] J-P. Serre, *Lettres à Ken Ribet du* 1/1/1981 *et du* 29/1/81, reproduced in Coll. Papers IV, 1–20 (= Oe.133).

[Se 93] J.-P. Serre, *Gèbres*, L'Ens.Math. (2) **39** (1993), 33–85 (= Oe.160).

[Se 00] J.-P. Serre, *Local Algebra,* Springer-Verlag, 2000.

[SGA 3] M. Demazure and A. Grothendieck, *Schémas en Groupes*, Lect.Notes in Math. **151-153**, Springer-Verlag, 1970.

[SGA $4\frac{1}{2}$] P. Deligne *et al*, *Cohomologie Étale*, Lect.Notes in Math. **569**, Springer-Verlag, 1977.

[Sp 74] T. A. Springer, *Regular elements of finite reflection groups*, Invent.math. **25** (1974), 159–198.

[SS 68] T. A. Springer and R. Steinberg, *Conjugacy Classes*, in Seminar on Algebraic Groups and Related Finite Groups, Lect.Notes in Math. **131**, Springer-Verlag, 1970 (= R. Steinberg, Coll.Papers, n°25).

[St 67] R. Steinberg, *Lectures on Chevalley Groups,* Yale University, 1967.

[St 68] R. Steinberg, *Endomorphisms of linear algebraic groups,* A.M.S.Memoirs, **80**, 1968 (= Coll.Papers, n°23).

[SZ 96] A. Silverberg and Yu.G. Zarhin, *Variations on a theme of Minkowski and Serre*, J.Pure Applied Algebra **111** (1996), 285-302.

[Th 60-64] J.G. Thompson, *Normal p-complements for finite groups*, Math.Zeit. **72** (1960), 332–354 and J. Algebra **1** (1964), 43–46.

[Ti 66a] J. Tits, *Normalisateurs de tores.* I. *Groupes de Coxeter étendus*, J. Algebra **1** (1966), 96–116.

[Ti 66b] J. Tits, *Sur les constantes de structure et le théorème d'existence des algèbres de Lie semi-simples*, Publ.Math. I.H.E.S. **31** (1966), 21–58.

J.-P. Serre
Collège de France
3, rue d'Ulm
F-75005 PARIS.