

Algebra & Number Theory

Volume 2

2008

No. 8



mathematical sciences publishers

Algebra & Number Theory

www.jant.org

EDITORS

MANAGING EDITOR

Bjorn Poonen
University of California
Berkeley, USA

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Georgia Benkart	University of Wisconsin, Madison, USA	Susan Montgomery	University of Southern California, USA
Dave Benson	University of Aberdeen, Scotland	Shigefumi Mori	RIMS, Kyoto University, Japan
Richard E. Borcherds	University of California, Berkeley, USA	Andrei Okounkov	Princeton University, USA
John H. Coates	University of Cambridge, UK	Raman Parimala	Emory University, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Victor Reiner	University of Minnesota, USA
Brian D. Conrad	University of Michigan, USA	Karl Rubin	University of California, Irvine, USA
Hélène Esnault	Universität Duisburg-Essen, Germany	Peter Sarnak	Princeton University, USA
Hubert Flenner	Ruhr-Universität, Germany	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Ronald Solomon	Ohio State University, USA
Andrew Granville	Université de Montréal, Canada	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Joseph Gubeladze	San Francisco State University, USA	J. Toby Stafford	University of Michigan, USA
Ehud Hrushovski	Hebrew University, Israel	Bernd Sturmfels	University of California, Berkeley, USA
Craig Huneke	University of Kansas, USA	Richard Taylor	Harvard University, USA
Mikhail Kapranov	Yale University, USA	Ravi Vakil	Stanford University, USA
Yujiro Kawamata	University of Tokyo, Japan	Michel van den Bergh	Hasselt University, Belgium
János Kollár	Princeton University, USA	Marie-France Vignéras	Université Paris VII, France
Hendrik W. Lenstra	Universiteit Leiden, The Netherlands	Kei-Ichi Watanabe	Nihon University, Japan
Yuri Manin	Northwestern University, USA	Andrei Zelevinsky	Northeastern University, USA
Barry Mazur	Harvard University, USA	Efim Zelmanov	University of California, San Diego, USA

PRODUCTION

ant@mathscipub.org

Paulo Ney de Souza, Production Manager

Silvio Levy, Senior Production Editor


See inside back cover or www.jant.org for submission instructions.

Regular subscription rate for 2008: \$180.00 a year (\$120.00 electronic only).

Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Algebra & Number Theory, ISSN 1937-0652, at Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

PUBLISHED BY

 **mathematical sciences publishers**
<http://www.mathscipub.org>

A NON-PROFIT CORPORATION

Typeset in L^AT_EX

Copyright ©2010 by Mathematical Sciences Publishers

Integral points on hyperelliptic curves

Yann Bugeaud, Maurice Mignotte, Samir Siksek,
Michael Stoll and Szabolcs Tengely

Let $C : Y^2 = a_n X^n + \cdots + a_0$ be a hyperelliptic curve with the a_i rational integers, $n \geq 5$, and the polynomial on the right-hand side irreducible. Let J be its Jacobian. We give a completely explicit upper bound for the integral points on the model C , provided we know at least one rational point on C and a Mordell–Weil basis for $J(\mathbb{Q})$. We also explain a powerful refinement of the Mordell–Weil sieve which, combined with the upper bound, is capable of determining all the integral points. Our method is illustrated by determining the integral points on the genus 2 hyperelliptic models $Y^2 - Y = X^5 - X$ and $\binom{Y}{2} = \binom{X}{5}$.

1. Introduction

Consider the hyperelliptic curve with affine model

$$C : Y^2 = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_0, \quad (1-1)$$

with a_0, \dots, a_n rational integers, $a_n \neq 0$, $n \geq 5$, and the polynomial on the right-hand side irreducible. Let $H = \max\{|a_0|, \dots, |a_n|\}$. In one of the earliest applications of his theory of lower bounds for linear forms in logarithms, Baker [1969] showed that any integral point (X, Y) on this affine model satisfies

$$\max(|X|, |Y|) \leq \exp \exp \exp((n^{10n} H)^{n^2}).$$

Such bounds have been improved considerably by many authors, including Sprindžuk [1977], Brindza [1984], Schmidt [1992], Poulakis [1991], Bilu [1995], Bugeaud [1997] and Voutier [1995]. Despite the improvements, the bounds remain astronomical and often involve inexplicit constants.

In this paper we explain a new method for explicitly computing the integral points on affine models of hyperelliptic curves (1-1). The method falls into two distinct steps:

MSC2000: primary 11G30; secondary 11J86.

Keywords: curve, integral point, Jacobian, height, Mordell–Weil group, Baker’s bound, Mordell–Weil sieve.

- (i) We give a completely explicit upper bound for the size of integral solutions of (1-1). This upper bound combines many refinements found in the papers of Voutier, Bugeaud, and others, together with Matveev's bounds [2000] for linear forms in logarithms, and a method for bounding the regulators based on a theorem of Landau [1918].
- (ii) The bounds obtained in (i), whilst substantially better than bounds given by earlier authors, are still astronomical. We explain a powerful variant of the Mordell–Weil sieve which, combined with the bound obtained in (i), is capable of showing that the known solutions to (1-1) are the only ones.

Step (i) requires two assumptions:

- (a) We assume that we know at least one rational point P_0 on C .
- (b) Let J be the Jacobian of C . We assume that a Mordell–Weil basis for $J(\mathbb{Q})$ is known.

For step (ii) we need assumptions (a), (b) and also:

- (c) We assume that the canonical height $\hat{h} : J(\mathbb{Q}) \rightarrow \mathbb{R}$ is explicitly computable and that we have explicit bounds for the difference

$$\mu_1 \leq h(D) - \hat{h}(D) \leq \mu'_1 \tag{1-2}$$

where h is an appropriately normalized logarithmic height on J that allows us to enumerate points P in $J(\mathbb{Q})$ with $h(P) \leq B$ for a given bound B .

Assumptions (a)–(c) deserve a comment or two. For many families of curves of higher genus, practical descent strategies are available for estimating the rank of the Mordell–Weil group; see for example [Cassels and Flynn 1996; Poonen and Schaefer 1997; Schaefer 1995; Stoll 2001]. To provably determine the Mordell–Weil group one however needs bounds for the difference between the logarithmic and canonical heights. For Jacobians of curves of genus 2 such bounds have been determined by Stoll [1999; 2002], building on previous work of Flynn and Smart [1997]. At present, no such bounds have been determined for Jacobians of curves of genus ≥ 3 , though work on this is in progress. The assumption about the knowledge of a rational point is a common sense assumption that brings some simplifications to our method, though the method can be modified to cope with the situation where no rational point is known. However, if a search on a curve of genus ≥ 2 reveals no rational points, it is probable that there are none, and the methods of Bruin and Stoll [2008a; 2008b; ≥ 2008] are likely to succeed in proving this.

We illustrate the practicality of our approach by proving:

Theorem 1.1. *The only integral solutions to the equation*

$$Y^2 - Y = X^5 - X \tag{1-3}$$

are

$$(X, Y) = (-1, 0), (-1, 1), (0, 0), (0, 1), (1, 0), (1, 1), (2, -5), \\ (2, 6), (3, -15), (3, 16), (30, -4929), (30, 4930).$$

Theorem 1.2. *The only integral solutions to the equation*

$$\begin{pmatrix} Y \\ 2 \end{pmatrix} = \begin{pmatrix} X \\ 5 \end{pmatrix} \tag{1-4}$$

are

$$(X, Y) = (0, 0), (0, 1), (1, 0), (1, 1), (2, 0), (2, 1), (3, 0), \\ (3, 1), (4, 0), (4, 1), (5, -1), (5, 2), (6, -3), (6, 4), \\ (7, -6), (7, 7), (15, -77), (15, 78), (19, -152), (19, 153).$$

Equations (1-3) and (1-4) are of historical interest and Section 2 gives a brief outline of their history. For now we merely mention that these two equations are the first two problems on a list of 22 unsolved Diophantine problems, compiled by Evertse and Tijdeman [2007] following a recent workshop on Diophantine equations at Leiden.

To appreciate why the innocent-looking (1-3) and (1-4) have resisted previous attempts, let us briefly survey the available methods which apply to hyperelliptic curves and then briefly explain why they fail in these cases. To determine the integral points on the affine model C given by (1-1) there are four available methods:

- (I) The first is Chabauty's elegant method which in fact determines all rational points on C in many cases, provided the rank of the Mordell–Weil group of its Jacobian is strictly less than the genus g ; see for example [Flynn 1997; Wetherell 1997]. Chabauty's method fails if the rank of the Mordell–Weil group exceeds the genus.
- (II) The second method is to use coverings, often combined with a version of Chabauty called *Elliptic Curve Chabauty*. See [Bruin 1999; 2003; Flynn and Wetherell 1999; 2001]. This approach often requires computations of Mordell–Weil groups over number fields (and does fail if the rank of the Mordell–Weil groups is too large).
- (III) The third method is to combine Baker's approach through S -units with the LLL algorithm to obtain all the solutions provided that certain relevant unit groups and class groups can be computed; for a modern treatment, see [Bilu and Hanrot 1998] or [Smart 1998, Section XIV.4]. This strategy often fails in practice as the number fields involved have very high degree.

- (IV) The fourth approach is to apply Skolem's method to the S -unit equations (see [Smart 1998, Section III.2]). This needs the same expensive information as method (III).

The Jacobians of the curves given by (1-3) and (1-4) respectively have ranks 3 and 6 and so Chabauty's method fails. To employ Elliptic Curve Chabauty would require the computation of Mordell–Weil groups of elliptic curves without rational 2-torsion over number fields of degree 5 (which does not seem practical at present). To apply the S -unit approach (with either LLL or Skolem) requires the computations of the unit groups and class groups of several number fields of degree 40—a computation that seems completely impractical at present.

Our paper is arranged as follows. Section 2 gives a brief history of (1-3) and (1-4). In Section 3 we show, after appropriate scaling, that an integral point (x, y) satisfies $x - \alpha = \kappa \xi^2$ where α is some fixed algebraic integer, $\xi \in \mathbb{Q}(\alpha)$, and κ is an algebraic integer belonging to a finite computable set. In Section 9 we give bounds for the size of solutions $x \in \mathbb{Z}$ to an equation of the form $x - \alpha = \kappa \xi^2$ where α and κ are fixed algebraic integers. Thus, in effect, we obtain bounds for the size of solutions of the integral points on our affine model (1-1). Sections 4–8 are preparation for Section 9: in particular Section 4 is concerned with heights; Section 5 explains how a theorem of Landau can be used to bound the regulators of number fields; Section 6 collects and refines various results on appropriate choices of systems of fundamental units; Section 7 is devoted to Matveev's bounds for linear forms in logarithms; in Section 8 we use Matveev's bounds and the results of previous sections to prove a bound on the size of solutions of unit equations; in Section 9 we deduce the bounds for x alluded to above from the bounds for solutions of unit equations. Despite our best efforts, the bounds obtained for x are still so large that no naive search up to those bounds is conceivable. Over Sections 10, 11 and 12 we explain how to sieve effectively up to these bounds using the Mordell–Weil group of the Jacobian. In particular, Section 11 gives a powerful refinement of the Mordell–Weil sieve (see [Bruin and Stoll 2008a; Bruin and Stoll \geq 2008]) which we expect to have applications elsewhere. Finally, in Section 13 we apply the method of this paper to prove Theorems 1.1 and 1.2.

2. History of (1-3) and (1-4)

Equation (1-3) is a special case of the family of Diophantine equations

$$Y^p - Y = X^q - X, \quad 2 \leq p < q. \quad (2-1)$$

This family has previously been studied by Fielder and Alford [1998] and by Mignotte and Pethő [1999]. The (genus 1) case $p = 2, q = 3$ was solved by

Mordell [1963] who showed that the only solutions in this case are

$$(X, Y) = (0, 0), (0, 1), (\pm 1, 0), (\pm 1, 1), (2, 3), (2, -2), (6, 15), (6, -14).$$

Felder and Alford presented the following list of solutions with $X, Y > 1$:

$$(p, q, X, Y) = (2, 3, 2, 3), (2, 3, 6, 15), (2, 5, 2, 6), (2, 5, 3, 16), \\ (2, 5, 30, 4930), (2, 7, 5, 280), (2, 13, 2, 91), (3, 7, 3, 13).$$

Mignotte and Pethő proved that for given p and q with $2 \leq p < q$, the Diophantine equation (2-1) has only a finite number of integral solutions. Assuming the *abc*-conjecture, they showed that (2-1) has only finitely many solutions with $X, Y > 1$.

If $p = 2$, $q > 2$ and y is a prime power, then Mignotte and Pethő found all solutions of the equation and these are all in Felder and Alford's list.

Equation (1-4) is a special case of the Diophantine equation

$$\binom{n}{k} = \binom{m}{l}, \quad (2-2)$$

in unknowns k, l, m, n . This is usually considered with the restrictions $2 \leq k \leq \frac{n}{2}$, and $2 \leq l \leq \frac{m}{2}$. The only known solutions (with these restrictions) are

$$\binom{16}{2} = \binom{10}{3}, \quad \binom{56}{2} = \binom{22}{3}, \quad \binom{120}{2} = \binom{36}{3}, \quad \binom{21}{2} = \binom{10}{4}, \\ \binom{153}{2} = \binom{19}{5}, \quad \binom{78}{2} = \binom{15}{5} = \binom{14}{6}, \quad \binom{221}{2} = \binom{17}{8}, \\ \binom{F_{2i+2}F_{2i+3}}{F_{2i}F_{2i+3}} = \binom{F_{2i+2}F_{2i+3} - 1}{F_{2i}F_{2i+3} + 1} \quad \text{for } i = 1, 2, \dots,$$

where F_n is the n -th Fibonacci number. It is known that there are no other nontrivial solutions with $\binom{n}{k} \leq 10^{30}$ or $n \leq 1000$; see [de Weger 1997]. The infinite family of solutions was found by Lind [1968] and Singmaster [1975].

Equation (2-2) has been completely solved for pairs

$$(k, l) = (2, 3), (2, 4), (2, 6), (2, 8), (3, 4), (3, 6), (4, 6).$$

These are the cases when one can easily reduce the equation to the determination of solutions of a number of Thue equations or elliptic Diophantine equations. Avanesov [1966/1967] found all solutions of (2-2) with $(k, l) = (2, 3)$. De Weger [1996] and independently Pintér [1995] solved the equation with $(k, l) = (2, 4)$. The case $(k, l) = (3, 4)$ reduces to the equation

$$Y(Y + 1) = X(X + 1)(X + 2)$$

which was solved by Mordell [1963]. The remaining pairs

$$(2, 6), (2, 8), (3, 6), (4, 6)$$

were treated by Stroeker and de Weger [1999], using linear forms in elliptic logarithms.

There are also some general finiteness results related to (2-2). Kiss [1988] proved that if $k = 2$ and l is a given odd prime, then the equation has only finitely many positive integral solutions. Using Baker's method, Brindza [1991] showed that (2-2) with $k = 2$ and $l \geq 3$ has only finitely many positive integral solutions.

3. Descent

Consider the integral points on the affine model of the hyperelliptic curve (1-1). If the polynomial on the right-hand side is reducible then the obvious factorisation argument reduces the problem of determining the integral points for (1-1) to determining those on simpler hyperelliptic curves, or on genus 1 curves. The integral points on a genus 1 curve can be determined by highly successful algorithms (see for example [Smart 1998; Stroeker and Tzanakis 2003]) based on LLL and David's bound for linear forms in elliptic logarithms.

We therefore suppose henceforth that the polynomial on the right-hand side of (1-1) is irreducible; this is certainly the most difficult case. By appropriate scaling, one transforms the problem of integral points on (1-1) to integral points on a model of the form

$$ay^2 = x^n + b_{n-1}x^{n-1} + \cdots + b_0, \quad (3-1)$$

where a and b_i are integers, with $a \neq 0$. We shall work henceforth with this model of the hyperelliptic curve. Denote the polynomial on the right-hand side by f and let α be a root of f . Then a standard argument shows that

$$x - \alpha = \kappa \zeta^2$$

where $\kappa, \zeta \in K = \mathbb{Q}(\alpha)$ and κ is an *algebraic integer that comes from a finite computable set*. In this section we suppose that the Mordell–Weil group $J(\mathbb{Q})$ of the curve C is known, and we show how to compute such a set of κ using our knowledge of the Mordell–Weil group $J(\mathbb{Q})$. The method for doing this depends on whether the degree n is odd or even.

3A. The odd degree case. Each coset of $J(\mathbb{Q})/2J(\mathbb{Q})$ has a coset representative of the form $\sum_{i=1}^m (P_i - \infty)$ where the set $\{P_1, \dots, P_m\}$ is stable under the action of Galois, and where all $y(P_i)$ are nonzero. Now write $x(P_i) = \gamma_i/d_i^2$ where γ_i is an algebraic integer and $d_i \in \mathbb{Z}_{\geq 1}$; moreover if P_i, P_j are conjugate then we may suppose that $d_i = d_j$ and so γ_i, γ_j are conjugate. To such a coset representative of

$J(\mathbb{Q})/2J(\mathbb{Q})$ we associate

$$\kappa = a^{(m \bmod 2)} \prod_{i=1}^m (\gamma_i - \alpha d_i^2).$$

Lemma 3.1. *Let \mathcal{K} be a set of κ associated as above to a complete set of coset representatives of $J(\mathbb{Q})/2J(\mathbb{Q})$. Then \mathcal{K} is a finite subset of \mathbb{O}_K and if (x, y) is an integral point on the model (3-1) then $x - \alpha = \kappa \zeta^2$ for some $\kappa \in \mathcal{K}$ and $\zeta \in K$.*

Proof. This follows trivially from the standard homomorphism

$$\theta : J(\mathbb{Q})/2J(\mathbb{Q}) \rightarrow K^*/K^{*2}$$

that is given by

$$\theta \left(\sum_{i=1}^m (P_i - \infty) \right) = a^m \prod_{i=1}^m (x(P_i) - \alpha) \pmod{K^{*2}}$$

for coset representatives $\sum (P_i - \infty)$ with $y(P_i) \neq 0$; see [Stoll 2001, Section 4]. □

3B. The even degree case. As mentioned in the introduction, we shall assume the existence of at least one rational point P_0 . If P_0 is one of the two points at infinity, let $\epsilon_0 = 1$. Otherwise, as f is irreducible, $y(P_0) \neq 0$; write $x(P_0) = \gamma_0/d_0^2$ with $\gamma_0 \in \mathbb{Z}$ and $d_0 \in \mathbb{Z}_{\geq 1}$ and let $\epsilon_0 = \gamma_0 - \alpha d_0^2$.

Each coset of $J(\mathbb{Q})/2J(\mathbb{Q})$ has a coset representative of the form $\sum_{i=1}^m (P_i - P_0)$ where the set $\{P_1, \dots, P_m\}$ is stable under the action of Galois, and where all $y(P_i)$ are nonzero for $i = 1, \dots, m$. Write $x(P_i) = \gamma_i/d_i^2$ where γ_i is an algebraic integer and $d_i \in \mathbb{Z}_{\geq 1}$; moreover if P_i, P_j are conjugate then we may suppose that $d_i = d_j$ and so γ_i, γ_j are conjugate. To such a coset representative of $J(\mathbb{Q})/2J(\mathbb{Q})$ we associate

$$\epsilon = \epsilon_0^{(m \bmod 2)} \prod_{i=1}^m (\gamma_i - \alpha d_i^2).$$

Lemma 3.2. *Let \mathcal{E} be a set of ϵ associated as above to a complete set of coset representatives of $J(\mathbb{Q})/2J(\mathbb{Q})$. Let Δ be the discriminant of the polynomial f . For each $\epsilon \in \mathcal{E}$, let \mathcal{B}_ϵ be the set of square-free rational integers supported only by primes dividing $a \Delta \text{Norm}_{K/\mathbb{Q}}(\epsilon)$. Let $\mathcal{K} = \{\epsilon b : \epsilon \in \mathcal{E}, b \in \mathcal{B}_\epsilon\}$. Then \mathcal{K} is a finite subset of \mathbb{O}_K and if (x, y) is an integral point on the model (3-1) then $x - \alpha = \kappa \zeta^2$ for some $\kappa \in \mathcal{K}$ and $\zeta \in K$.*

Proof. In our even degree case, the homomorphism θ takes values in K^*/\mathbb{Q}^*K^{*2} . Thus if (x, y) is an integral point on the model (3-1), we have that $(x - \alpha) = \epsilon b \zeta^2$

for some $\epsilon \in \mathcal{E}$ and b a square-free rational integer. A standard argument shows that $2 \mid \text{ord}_{\wp}(x - \alpha)$ for all prime ideals $\wp \nmid a\Delta$. Hence, $2 \mid \text{ord}_{\wp}(b)$ for all $\wp \nmid a\Delta\epsilon$. Let $\wp \mid p$ where p is a rational prime not dividing $a\Delta \text{Norm}_{K/\mathbb{Q}}(\epsilon)$. Then p is unramified in K/\mathbb{Q} and so $\text{ord}_p(b) = \text{ord}_{\wp}(b) \equiv 0 \pmod{2}$. This shows that $b \in \mathcal{B}_\epsilon$ and proves the lemma. \square

3C. Remarks. The following remarks are applicable to both odd and even degree cases.

- (i) We point out that we can still obtain a suitable (though larger) set of κ that satisfies the conclusions of Lemmas 3.1 and 3.2, even if we do not know coset representatives for $J(\mathbb{Q})/2J(\mathbb{Q})$, provided we are able to compute the class group and unit group of the number field K ; for this see for example [Bruin 1999, Section 2.2].
- (ii) We can use local information at small and bad primes to restrict the set \mathcal{H} further, compare [Bruin and Stoll 2008a; 2008b], where this is applied to rational points. In our case, we can restrict the local computations to $x \in \mathbb{Z}_p$ instead of \mathbb{Q}_p .

4. Heights

We fix once and for all the following notation.

K	a number field,
\mathbb{O}_K	the ring of integers of K ,
M_K	the set of all places of K ,
M_K^0	the set of non-Archimedean places of K ,
M_K^∞	the set of Archimedean places of K ,
v	a place of K ,
K_v	the completion of K at v ,
d_v	the local degree $[K_v : \mathbb{Q}_v]$.

For $v \in M_K$, we let $|\cdot|_v$ be the usual normalized valuation corresponding to v ; in particular if v is non-Archimedean and p is the rational prime below v then $|p|_v = p^{-1}$. Thus if L/K is a field extension, and ω a place of L above v then $|\alpha|_\omega = |\alpha|_v$, for all $\alpha \in K$.

Define

$$\|\alpha\|_v = |\alpha|_v^{d_v}.$$

Hence for $\alpha \in K^*$, the product formula states that

$$\prod_{v \in M_K} \|\alpha\|_v = 1.$$

In particular, if v is Archimedean, corresponding to a real or complex embedding σ of K , then

$$|\alpha|_v = |\sigma(\alpha)| \quad \text{and} \quad \|\alpha\|_v = \begin{cases} |\sigma(\alpha)| & \text{if } \sigma \text{ is real,} \\ |\sigma(\alpha)|^2 & \text{if } \sigma \text{ is complex.} \end{cases}$$

For $\alpha \in K$, the (absolute) logarithmic height $h(\alpha)$ is given by

$$h(\alpha) = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} d_v \log \max\{1, |\alpha|_v\} = \frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} \log \max\{1, \|\alpha\|_v\}. \tag{4-1}$$

The absolute logarithmic height of α is independent of the field K containing α .

We shall need the following elementary properties of heights.

Lemma 4.1. *For any nonzero algebraic number α , we have $h(\alpha^{-1}) = h(\alpha)$. For algebraic numbers $\alpha_1, \dots, \alpha_n$, we have*

$$\begin{aligned} h(\alpha_1 \alpha_2 \cdots \alpha_n) &\leq h(\alpha_1) + \cdots + h(\alpha_n), \\ h(\alpha_1 + \cdots + \alpha_n) &\leq \log n + h(\alpha_1) + \cdots + h(\alpha_n). \end{aligned}$$

Proof. The lemma is [Silverman 1986, Exercise 8.8]. We do not know of a reference for the proof and so we will indicate briefly the proof of the second (more difficult) inequality. For $v \in M_K$, choose i_v in $\{1, \dots, n\}$ to satisfy

$$\max\{|\alpha_1|_v, \dots, |\alpha_n|_v\} = |\alpha_{i_v}|_v.$$

Note that

$$|\alpha_1 + \cdots + \alpha_n|_v \leq \epsilon_v |\alpha_{i_v}|_v,$$

where $\epsilon_v = n$ if v is Archimedean or $\epsilon_v = 1$ otherwise. Thus

$$\begin{aligned} \log \max\{1, |\alpha_1 + \cdots + \alpha_n|_v\} &\leq \log \epsilon_v + \log \max\{1, |\alpha_{i_v}|_v\} \\ &\leq \log \epsilon_v + \sum_{i=1}^n \log \max\{1, |\alpha_i|_v\}. \end{aligned}$$

Observe that

$$\frac{1}{[K : \mathbb{Q}]} \sum_{v \in M_K} d_v \log \epsilon_v = \frac{\log n}{[K : \mathbb{Q}]} \sum_{v \in M_K^\infty} d_v = \log n.$$

The desired inequality follows from the definition of logarithmic height (4-1). \square

4A. Height lower bound. We need the following result of Voutier [1996] concerning Lehmer’s problem.

Lemma 4.2. *Let K be a number field of degree d . Let*

$$\partial_K = \begin{cases} \frac{\log 2}{d} & \text{if } d = 1, 2, \\ \frac{1}{4} \left(\frac{\log \log d}{\log d} \right)^3 & \text{if } d \geq 3. \end{cases}$$

Then, for every nonzero algebraic number α in K , which is not a root of unity,

$$\deg(\alpha) h(\alpha) \geq \partial_K.$$

Throughout, by the logarithm of a complex number, we mean the principal determination of the logarithm. In other words, if $x \in \mathbb{C}^*$ we express $x = re^{i\theta}$ where $r > 0$ and $-\pi < \theta \leq \pi$; we then let $\log x = \log r + i\theta$.

Lemma 4.3. *Let K be a number field and let*

$$\partial'_K = \left(1 + \frac{\pi^2}{\partial_K^2} \right)^{1/2}.$$

For any nonzero $\alpha \in K$ and any place $v \in M_K$,

$$\log |\alpha|_v \leq \deg(\alpha) h(\alpha), \quad \log \|\alpha\|_v \leq [K : \mathbb{Q}] h(\alpha).$$

Moreover, if α is not a root of unity and σ is a real or complex embedding of K then

$$|\log \sigma(\alpha)| \leq \partial'_K \deg(\alpha) h(\alpha).$$

Proof. The first two inequalities are an immediate consequence of the definition of absolute logarithmic height. For the last, write $\sigma(\alpha) = e^{a+ib}$, with $a = \log |\sigma(\alpha)|$ and $|b| \leq \pi$, and let $d = \deg(\alpha)$. Then we have

$$|\log \sigma(\alpha)| = (a^2 + b^2)^{1/2} \leq (\log^2 |\sigma(\alpha)| + \pi^2)^{1/2} \leq ((d h(\alpha))^2 + \pi^2)^{1/2}.$$

By Lemma 4.2 we have $d h(\alpha) \geq \partial_K$, so

$$|\log \sigma(\alpha)| \leq d h(\alpha) \left(1 + \frac{\pi^2}{\partial_K^2} \right)^{1/2},$$

as required. □

5. Bounds for regulators

Later on we need to give upper bounds for the regulators of complicated number fields of high degree. The following lemma, based on bounds of Landau [1918], is an easy way to obtain reasonable bounds.

Lemma 5.1. *Let K be a number field with degree $d = u + 2v$ where u and v are respectively the numbers of real and complex embeddings. Denote the absolute*

discriminant by D_K and the regulator by R_K , and the number of roots of unity in K by w . Suppose, moreover, that L is a real number such that $D_K \leq L$. Let

$$a = 2^{-v} \pi^{-d/2} \sqrt{L}.$$

Define the function $f_K(L, s)$ by

$$f_K(L, s) = 2^{-u} w a^s \left(\Gamma\left(\frac{s}{2}\right)\right)^u (\Gamma(s))^v s^{d+1} (s-1)^{1-d},$$

and let $B_K(L) = \min\{f_K(L, 2-t/1000) : t = 0, 1, \dots, 999\}$. Then $R_K < B_K(L)$.

Proof. Landau [1918, proof of Hilfssatz 1] established the inequality

$$R_K < f_K(D_K, s)$$

for all $s > 1$. It is thus clear that $R_K < B_K(L)$. □

Remark 5.2. For a complicated number field of high degree it is difficult to calculate the discriminant D_K exactly, though it is easy to give an upper bound L for its size. It is also difficult to minimise the function $f_K(L, s)$ analytically, but we have found that the above gives an accurate enough result, which is easy to calculate on a computer.

6. Fundamental units

For the number fields we are concerned with, we shall need to work with a certain system of fundamental units.

Lemma 6.1 [Bugeaud and Györy 1996, Lemma 1]. *Let K be a number field of degree d and let $r = r_K$ be its unit rank and R_K its regulator. Define the constants*

$$c_1 = c_1(K) = \frac{(r!)^2}{2^{r-1} d^r}, \quad c_2 = c_2(K) = c_1 \left(\frac{d}{\partial_K}\right)^{r-1}, \quad c_3 = c_3(K) = c_1 \frac{d^r}{\partial_K}.$$

Then K admits a system $\{\varepsilon_1, \dots, \varepsilon_r\}$ of fundamental units such that:

- (i) $\prod_{i=1}^r h(\varepsilon_i) \leq c_1 R_K$.
- (ii) $h(\varepsilon_i) \leq c_2 R_K, 1 \leq i \leq r$.
- (iii) *Write \mathcal{M} for the $r \times r$ -matrix $(\log \|\varepsilon_i\|_v)$, where v runs over r of the Archimedean places of K and $1 \leq i \leq r$. Then the absolute values of the entries of \mathcal{M}^{-1} are bounded above by c_3 .*

Lemma 6.2. *Let K be a number field of degree d , and let $\{\varepsilon_1, \dots, \varepsilon_r\}$ be a system of fundamental units as in Lemma 6.1. Define the constant $c_4 = c_4(K) = r d c_3$. Suppose $\varepsilon = \zeta \varepsilon_1^{b_1} \dots \varepsilon_r^{b_r}$, where ζ is a root of unity in K . Then*

$$\max\{|b_1|, \dots, |b_r|\} \leq c_4 h(\varepsilon).$$

Proof. Note that for any Archimedean place v of K ,

$$\log \|\varepsilon\|_v = \sum b_i \log \|\varepsilon_i\|_v.$$

The lemma now follows from part (iii) of Lemma 6.1, plus the fact that $\log \|\varepsilon\|_v \leq d h(\varepsilon)$ for all v given by Lemma 4.3. □

The following result is a special case of [Bugeaud and Győry 1996, Lemma 2].

Lemma 6.3. *Let K be a number field of unit rank r and regulator K . Let α be a nonzero algebraic integer belonging to K . Then there exists a unit ε of K such that*

$$h(\alpha\varepsilon) \leq c_5 R_K + \frac{\log |\text{Norm}_{K/\mathbb{Q}}(\alpha)|}{[K : \mathbb{Q}]}$$

where

$$c_5 = c_5(K) = \frac{r^{r+1}}{2\delta_K^{r-1}}.$$

Lemma 6.4. *Let K be a number field, $\beta, \varepsilon \in K^*$ with ε being a unit. Let σ be the real or complex embedding that makes $|\sigma(\beta\varepsilon)|$ minimal. Then*

$$h(\beta\varepsilon) \leq h(\beta) - \log |\sigma(\beta\varepsilon)|.$$

Proof. As usual, write $d = [K : \mathbb{Q}]$ and $d_v = [K_v : \mathbb{Q}_v]$. Then

$$\begin{aligned} h(\beta\varepsilon) &= h\left(\frac{1}{\beta\varepsilon}\right) \\ &= \frac{1}{d} \sum_{v \in M_K^\infty} d_v \max\{0, \log(|\beta\varepsilon|_v^{-1})\} + \frac{1}{d} \sum_{v \in M_K^0} d_v \max\{0, \log(|\beta\varepsilon|_v^{-1})\} \\ &\leq \log(|\sigma(\beta\varepsilon)|^{-1}) + \frac{1}{d} \sum_{v \in M_K^0} d_v \max\{0, \log(|\beta|_v^{-1})\} \\ &\leq -\log |\sigma(\beta\varepsilon)| + \frac{1}{d} \sum_{v \in M_K} d_v \max\{0, \log(|\beta|_v^{-1})\} \\ &\leq -\log |\sigma(\beta\varepsilon)| + h(\beta). \end{aligned} \quad \square$$

7. Matveev’s lower bound for linear forms in logarithms

Let L be a number field and let σ be a real or complex embedding. For $\alpha \in L^*$ we define the *modified logarithmic height of α with respect to σ* to be

$$h_{L,\sigma}(\alpha) := \max\{[L : \mathbb{Q}] h(\alpha), |\log \sigma(\alpha)|, 0.16\}.$$

The modified height is clearly dependent on the number field; we shall need the following Lemma which gives a relation between the modified and absolute height.

Lemma 7.1. *Let $K \subseteq L$ be number fields and write*

$$\partial_{L/K} = \max \left\{ [L : \mathbb{Q}], [K : \mathbb{Q}] \partial'_K, \frac{0.16[K : \mathbb{Q}]}{\partial_K} \right\}.$$

Then for any $\alpha \in K$ which is neither zero nor a root of unity, and any real or complex embedding σ of L ,

$$h_{L,\sigma}(\alpha) \leq \partial_{L/K} h(\alpha).$$

Proof. By Lemma 4.3 we have

$$[K : \mathbb{Q}] \partial'_K h(\alpha) \geq \partial'_K \deg(\alpha) h(\alpha) \geq |\log \sigma(\alpha)|.$$

Moreover, by Lemma 4.2,

$$\frac{0.16[K : \mathbb{Q}] h(\alpha)}{\partial_K} \geq \frac{0.16 \deg(\alpha) h(\alpha)}{\partial_K} \geq 0.16.$$

The lemma follows. □

We shall apply lower bounds on linear forms, more precisely a version of Matveev’s estimates [2000]. We recall that \log denotes the principal determination of the logarithm.

Lemma 7.2. *Let L be a number field of degree d , with $\alpha_1, \dots, \alpha_n \in L^*$. Define a constant*

$$C(L, n) := 3 \cdot 30^{n+4} \cdot (n + 1)^{5.5} d^2 (1 + \log d).$$

Consider the “linear form”

$$\Lambda := \alpha_1^{b_1} \cdots \alpha_n^{b_n} - 1,$$

where b_1, \dots, b_n are rational integers and let $B := \max\{|b_1|, \dots, |b_n|\}$. If $\Lambda \neq 0$, and σ is any real or complex embedding of L then

$$\log |\sigma(\Lambda)| > -C(L, n)(1 + \log(nB)) \prod_{j=1}^n h_{L,\sigma}(\alpha_j).$$

Proof. This straightforward corollary of Matveev’s estimates is [Bugeaud et al. 2006, Theorem 9.4]. □

8. Bounds for unit equations

Now we are ready to prove an explicit version of [Bugeaud 1997, Lemma 4]. The proposition below allows us to replace in the final estimate the regulator of the larger field by the product of the regulators of two of its subfields. This often results in a significant improvement of the upper bound for the height. This idea is due to Voutier [1995].

Proposition 8.1. *Let L be a number field of degree d , which contains K_1 and K_2 as subfields. Let R_{K_i} (respectively r_i) be the regulator (respectively the unit rank) of K_i . Suppose further that v_1, v_2 and v_3 are nonzero elements of L with height $\leq H$ (with $H \geq 1$) and consider the unit equation*

$$v_1\varepsilon_1 + v_2\varepsilon_2 + v_3\varepsilon_3 = 0 \tag{8-1}$$

where ε_1 is a unit of K_1 , ε_2 a unit of K_2 and ε_3 a unit of L . Then, for $i = 1$ and 2 ,

$$h\left(\frac{v_i\varepsilon_i}{v_3\varepsilon_3}\right) \leq A_2 + A_1 \log(H + \max\{h(v_1\varepsilon_1), h(v_2\varepsilon_2)\}),$$

where

$$A_1 = 2H \cdot C(L, r_1 + r_2 + 1) \cdot c_1(K_1)c_1(K_2)\partial_{L/L} \cdot (\partial_{L/K_1})^{r_1} \cdot (\partial_{L/K_2})^{r_2} \cdot R_{K_1}R_{K_2},$$

$$A_2 = 2H + A_1 + A_1 \log((r_1 + r_2 + 1) \cdot \max\{c_4(K_1), c_4(K_2), 1\}).$$

Proof. Let $\{\mu_1, \dots, \mu_{r_1}\}$ and $\{\rho_1, \dots, \rho_{r_2}\}$ be respectively systems of fundamental units for K_1 and K_2 as in Lemma 6.1; in particular we know that

$$\prod_{j=1}^{r_1} h(\mu_j) \leq c_1(K_1)R_{K_1}, \quad \prod_{j=1}^{r_2} h(\rho_j) \leq c_1(K_2)R_{K_2}. \tag{8-2}$$

We can write

$$\varepsilon_1 = \zeta_1 \mu_1^{b_1} \cdots \mu_{r_1}^{b_{r_1}}, \quad \varepsilon_2 = \zeta_2 \rho_1^{f_1} \cdots \rho_{r_2}^{f_{r_2}},$$

where ζ_1 and ζ_2 are roots of unity and b_1, \dots, b_{r_1} , and f_1, \dots, f_{r_2} are rational integers. Set

$$B_1 = \max\{|b_1|, \dots, |b_{r_1}|\}, \quad B_2 = \max\{|f_1|, \dots, |f_{r_2}|\}, \quad B = \max\{B_1, B_2, 1\}.$$

Set $\alpha_0 = -\frac{\zeta_2 v_2}{\zeta_1 v_1}$ and $b_0 = 1$. By (8-1),

$$\frac{v_3\varepsilon_3}{v_1\varepsilon_1} = \alpha_0^{b_0} \mu_1^{-b_1} \cdots \mu_{r_1}^{-b_{r_1}} \rho_1^{f_1} \cdots \rho_{r_2}^{f_{r_2}} - 1.$$

Now choose the real or complex embedding σ of L such that $|\sigma(\frac{v_3\varepsilon_3}{v_1\varepsilon_1})|$ is minimal.

We apply Matveev’s estimate (Lemma 7.2) to this “linear form”, obtaining

$$\log\left|\sigma\left(\frac{v_3\varepsilon_3}{v_1\varepsilon_1}\right)\right| > -C(L, n)(1 + \log(nB)) h_{L,\sigma}(\alpha_0) \prod_{j=1}^{r_1} h_{L,\sigma}(\mu_j) \prod_{j=1}^{r_2} h_{L,\sigma}(\rho_j),$$

where $n = r_1 + r_2 + 1$. Using Lemma 7.1 and (8-2) we obtain

$$\prod_{j=1}^{r_1} h_{L,\sigma}(\mu_j) \leq (\partial_{L/K_1})^{r_1} \prod_{j=1}^{r_1} h(\mu_j) \leq c_1(K_1)(\partial_{L/K_1})^{r_1} R_{K_1},$$

and a similar estimate for $\prod_{j=1}^{r_2} h_{L,\sigma}(\rho_j)$. Moreover, again by Lemma 7.1 and Lemma 4.1, $h_{L,\sigma}(\alpha_0) \leq 2H\delta_{L/L}$. Thus

$$\log \left| \sigma \left(\frac{v_3 \varepsilon_3}{v_1 \varepsilon_1} \right) \right| > -A_1(1 + \log(nB)).$$

Now applying Lemma 6.4, we obtain that

$$h \left(\frac{v_3 \varepsilon_3}{v_1 \varepsilon_1} \right) \leq h \left(\frac{v_3}{v_1} \right) + A_1(1 + \log(nB)) \leq 2H + A_1(1 + \log(nB)).$$

The proof is complete on observing, from Lemma 6.2, that

$$B \leq \max\{c_4(K_1), c_4(K_2), 1\} \cdot \max\{h(\varepsilon_1), h(\varepsilon_2), 1\},$$

and from Lemma 4.1,

$$h(v_i \varepsilon_i) \leq h(\varepsilon_i) + h(v_i) \leq h(\varepsilon) + H. \quad \square$$

9. Upper bounds for the size of integral points on hyperelliptic curves

We shall need the following standard sort of lemma.

Lemma 9.1. *Let a, b, c, y be positive numbers and suppose that*

$$y \leq a + b \log(c + y).$$

Then

$$y \leq 2b \log b + 2a + c.$$

Proof. Let $z = c + y$, so that

$$z \leq (a + c) + b \log z.$$

Now we apply the case $h = 1$ of [Pethö and de Weger 1986, Lemma 2.2]; this gives

$$z \leq 2(b \log b + a + c). \quad \square$$

Theorem 9.2. *Let α be an algebraic integer of degree at least 3, and let κ be an integer belonging to K . Let $\alpha_1, \alpha_2, \alpha_3$ be distinct conjugates of α and $\kappa_1, \kappa_2, \kappa_3$ be the corresponding conjugates of κ . Let*

$$K_1 = \mathbb{Q}(\alpha_1, \alpha_2, \sqrt{\kappa_1 \kappa_2}), \quad K_2 = \mathbb{Q}(\alpha_1, \alpha_3, \sqrt{\kappa_1 \kappa_3}), \quad K_3 = \mathbb{Q}(\alpha_2, \alpha_3, \sqrt{\kappa_2 \kappa_3}),$$

and

$$L = \mathbb{Q}(\alpha_1, \alpha_2, \alpha_3, \sqrt{\kappa_1 \kappa_2}, \sqrt{\kappa_1 \kappa_3}).$$

Let R be an upper bound for the regulators of K_1, K_2 and K_3 . Let r be the maximum of the unit ranks of K_1, K_2, K_3 . Let

$$\begin{aligned}
 c_j^* &= \max_{1 \leq i \leq 3} c_j(K_i), \\
 N &= \max_{1 \leq i, j \leq 3} \left| \text{Norm}_{\mathbb{Q}(\alpha_i, \alpha_j)/\mathbb{Q}}(\kappa_i(\alpha_i - \alpha_j)) \right|^2, \\
 H^* &= c_5^* R + \frac{\log N}{\min_{1 \leq i \leq 3} [K_i : \mathbb{Q}]} + h(\kappa), \\
 A_1^* &= 2H^* \cdot C(L, 2r + 1) \cdot (c_1^*)^2 \partial_{L/L} \cdot \left(\max_{1 \leq i \leq 3} \partial_{L/K_i} \right)^{2r} \cdot R^2, \\
 A_2^* &= 2H^* + A_1^* + A_1^* \log((2r + 1) \cdot \max\{c_4^*, 1\}).
 \end{aligned}$$

If $x \in \mathbb{Z} \setminus \{0\}$ satisfies $x - \alpha = \kappa \zeta^2$ for some $\zeta \in K$ then

$$\log |x| \leq 8A_1^* \log(4A_1^*) + 8A_2^* + H^* + 20 \log 2 + 13 h(\kappa) + 19 h(\alpha).$$

Proof. Conjugating the relation $x - \alpha = \kappa \zeta^2$ appropriately and taking differences we obtain

$$\alpha_1 - \alpha_2 = \kappa_2 \zeta_2^2 - \kappa_1 \zeta_1^2, \quad \alpha_3 - \alpha_1 = \kappa_1 \zeta_1^2 - \kappa_3 \zeta_3^2, \quad \alpha_2 - \alpha_3 = \kappa_3 \zeta_3^2 - \kappa_2 \zeta_2^2.$$

Let

$$\tau_1 = \kappa_1 \zeta_1, \quad \tau_2 = \sqrt{\kappa_1 \kappa_2} \zeta_2, \quad \tau_3 = \sqrt{\kappa_1 \kappa_3} \zeta_3.$$

Observe that

$$\kappa_1(\alpha_1 - \alpha_2) = \tau_2^2 - \tau_1^2, \quad \kappa_1(\alpha_3 - \alpha_1) = \tau_1^2 - \tau_3^2, \quad \kappa_1(\alpha_2 - \alpha_3) = \tau_3^2 - \tau_2^2,$$

and

$$\tau_2 \pm \tau_1 \in K_1, \quad \tau_1 \pm \tau_3 \in K_2, \quad \tau_3 \pm \tau_2 \in \sqrt{\frac{\kappa_1}{\kappa_2}} K_3.$$

We claim that each $\tau_i \pm \tau_j$ can be written in the form $v\varepsilon$ where ε is a unit in one of the K_i and $v \in L$ is an integer satisfying $h(v) \leq H^*$. Let us show this for $\tau_2 - \tau_3$; the other cases are either similar or easier. Note that $\tau_2 - \tau_3 = \sqrt{\frac{\kappa_1}{\kappa_2}} v''$ where v'' is an integer belonging to K_3 . Moreover, v'' divides

$$\sqrt{\frac{\kappa_2}{\kappa_1}}(\tau_3 - \tau_2) \cdot \sqrt{\frac{\kappa_2}{\kappa_1}}(\tau_3 + \tau_2) = \kappa_2(\alpha_2 - \alpha_3).$$

Hence $|\text{Norm}_{K_3/\mathbb{Q}}(v'')| \leq N$. By Lemma 6.3, we can write $v'' = v'\varepsilon$ where $\varepsilon \in K_3$ and

$$h(v') \leq c_5(K_3)R + \frac{\log N}{[K_3 : \mathbb{Q}]}.$$

Now let $v = \sqrt{\frac{\kappa_1}{\kappa_2}} v'$. Thus $\tau_2 - \tau_3 = v\varepsilon$ where $h(v) \leq h(v') + h(\kappa) \leq H^*$ proving our claim.

We apply Proposition 8.1 to the unit equation

$$(\tau_1 - \tau_2) + (\tau_3 - \tau_1) + (\tau_2 - \tau_3) = 0,$$

which is indeed of the form $\nu_1 \varepsilon_1 + \nu_2 \varepsilon_2 + \nu_3 \varepsilon_3 = 0$ where the ν_i and ε_i satisfy the conditions of that proposition with H replaced by H^* . We obtain

$$h\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \leq A_2^* + A_1^* \log(H^* + \max\{h(\tau_2 - \tau_3), h(\tau_1 - \tau_2)\}).$$

Observe that

$$\begin{aligned} h(\tau_i \pm \tau_j) &\leq \log 2 + h(\tau_i) + h(\tau_j) \leq \log 2 + 2h(\kappa) + 2h(\zeta) \\ &\leq \log 2 + 3h(\kappa) + h(x - \alpha) \leq 2\log 2 + 3h(\kappa) + h(\alpha) + \log |x|, \end{aligned}$$

where we have made repeated use of Lemma 4.1. Thus

$$h\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \leq A_2^* + A_1^* \log(A_3^* + \log |x|),$$

where $A_3^* = H^* + 2\log 2 + 3h(\kappa) + h(\alpha)$.

We also apply Proposition 8.1 to the unit equation

$$(\tau_1 + \tau_2) + (\tau_3 - \tau_1) - (\tau_2 + \tau_3) = 0,$$

to obtain precisely the same bound for $h\left(\frac{\tau_1 + \tau_2}{\tau_1 - \tau_3}\right)$. Using the identity

$$\left(\frac{\tau_1 - \tau_2}{\tau_1 - \tau_3}\right) \cdot \left(\frac{\tau_1 + \tau_2}{\tau_1 - \tau_3}\right) = \frac{\kappa_1(\alpha_2 - \alpha_1)}{(\tau_1 - \tau_3)^2},$$

we obtain that

$$h(\tau_1 - \tau_3) \leq \frac{\log 2 + h(\kappa)}{2} + h(\alpha) + A_2^* + A_1^* \log(A_3^* + \log |x|).$$

Now

$$\begin{aligned} \log |x| &\leq \log 2 + h(\alpha) + h(x - \alpha_1) \\ &\leq \log 2 + h(\alpha) + h(\kappa) + 2h(\tau_1) \quad (\text{using } x - \alpha_1 = \frac{\tau_1^2}{\kappa_1}) \\ &\leq 5\log 2 + h(\alpha) + h(\kappa) + 2h(\tau_1 + \tau_3) + 2h(\tau_1 - \tau_3) \\ &\leq 5\log 2 + h(\alpha) + h(\kappa) + 2h\left(\frac{\kappa_1(\alpha_3 - \alpha_1)}{\tau_1 - \tau_3}\right) + 2h(\tau_1 - \tau_3) \\ &\leq 7\log 2 + 5h(\alpha) + 3h(\kappa) + 4h(\tau_1 - \tau_3) \\ &\leq 9\log 2 + 9h(\alpha) + 5h(\kappa) + 4A_2^* + 4A_1^* \log(A_3^* + \log |x|). \end{aligned}$$

The theorem follows from Lemma 9.1. □

10. The Mordell–Weil sieve I

The Mordell–Weil sieve is a technique that can be used to show the nonexistence of rational points on a curve (for example [Bruin and Stoll 2008a; ≥ 2008]), or to help determine the set of rational points in conjunction with the method of Chabauty (for example [Bruin and Elkies 2002]); for connections to the Brauer–Manin obstruction see, for example, [Flynn 2004; Poonen 2006; Stoll 2007]. In this section and the next we explain how the Mordell–Weil sieve can be used to show that any rational point on a curve of genus ≥ 2 is either a known rational point or a very large rational point.

In this section we let C/\mathbb{Q} be a smooth projective curve (not necessarily hyperelliptic) of genus $g \geq 2$ and we let J be its Jacobian. As indicated in the introduction, we assume the knowledge of some rational points on C ; henceforth let D be a fixed rational point on C (or even a fixed rational divisor of degree 1) and let J be the corresponding Abel–Jacobi map:

$$J : C \rightarrow J, \quad P \mapsto [P - D].$$

Let W be the image in J of the known rational points on C . The Mordell–Weil sieve is a strategy for obtaining a very large and “smooth” positive integer B such that

$$J(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q}).$$

Recall that a positive integer B is called *A-smooth* if all its prime factors are $\leq A$. By saying that B is smooth, we loosely mean that it is *A-smooth* with A much smaller than B .

Let S be a finite set of primes, which for now we assume to be primes of good reduction for the curve C . The basic idea is to consider the following commutative diagram:

$$\begin{array}{ccc} C(\mathbb{Q}) & \xrightarrow{J} & J(\mathbb{Q})/BJ(\mathbb{Q}) \\ \downarrow & & \downarrow \alpha \\ \prod_{p \in S} C(\mathbb{F}_p) & \xrightarrow{J} & \prod_{p \in S} J(\mathbb{F}_p)/BJ(\mathbb{F}_p). \end{array}$$

The image of $C(\mathbb{Q})$ in $J(\mathbb{Q})/BJ(\mathbb{Q})$ must then be contained in the subset of $J(\mathbb{Q})/BJ(\mathbb{Q})$ of elements that map under α into the image of the lower horizontal map. If we find that this subset equals the image of W in $J(\mathbb{Q})/BJ(\mathbb{Q})$, then we have shown that

$$J(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q})$$

as desired. Note that, at least in principle, the required computation is finite: each set $C(\mathbb{F}_p)$ is finite and can be enumerated, hence $J(C(\mathbb{F}_p))$ can be determined, and

we assume that we know explicit generators of $J(\mathbb{Q})$, which allows us to construct the finite set $J(\mathbb{Q})/BJ(\mathbb{Q})$. In practice, and in particular for the application we have in mind here, we will need a very large value of B , so this naive approach is much too inefficient. In [Bruin and Stoll 2008a; \geq 2008], the authors describe how one can perform this computation in a more efficient way.

One obvious improvement is to replace the lower horizontal map in the diagram above by a product of maps

$$C(\mathbb{Q}_p) \xrightarrow{J} G_p/BG_p$$

with suitable finite quotients G_p of $J(\mathbb{Q}_p)$. We have used this to incorporate information modulo higher powers of p for small primes p . This kind of information is often called “deep” information, as opposed to the “flat” information obtained from reduction modulo good primes.

We can always force B to be divisible by any given (not too big) number. In our application we will want B to kill the rational torsion subgroup of J .

11. The Mordell–Weil sieve II

We continue with the notation of Section 10. Let W be the image in $J(\mathbb{Q})$ of all the known rational points on C . We assume that the strategy of Section 10 is successful in yielding a large “smooth” integer B such that any point $P \in C(\mathbb{Q})$ satisfies $J(P) - w \in BJ(\mathbb{Q})$ for some $w \in W$, and moreover, that B kills all the torsion of $J(\mathbb{Q})$.

Let D_1, \dots, D_r be a basis of the free part of $J(\mathbb{Q})$ and let

$$\phi : \mathbb{Z}^r \rightarrow J(\mathbb{Q}), \quad \phi(a_1, \dots, a_r) = \sum a_i D_i,$$

so that the image of ϕ is simply the free part of $J(\mathbb{Q})$. Our assumption now is that

$$J(C(\mathbb{Q})) \subset W + \phi(B\mathbb{Z}^r).$$

Set $L_0 = B\mathbb{Z}^r$. We explain a method of obtaining a (very long) decreasing sequence of lattices in \mathbb{Z}^r :

$$B\mathbb{Z}^r = L_0 \supseteq L_1 \supseteq L_2 \supseteq \dots \supseteq L_k \tag{11-1}$$

such that

$$J(C(\mathbb{Q})) \subset W + \phi(L_j)$$

for $j = 1, \dots, k$.

If q is a prime of good reduction for J we denote by

$$\phi_q : \mathbb{Z}^r \rightarrow J(\mathbb{F}_q), \quad \phi_q(a_1, \dots, a_r) = \sum a_i \tilde{D}_i,$$

and so $\phi_q(\mathbf{1}) = \widetilde{\phi(\mathbf{1})}$.

Lemma 11.1. *Let W be a finite subset of $J(\mathbb{Q})$, and let L be a subgroup of \mathbb{Z}^r . Suppose that $J(C(\mathbb{Q})) \subset W + \phi(L)$. Let q be a prime of good reduction for C and J . Let L' be the kernel of the restriction $\phi_q|_L$. Let $\mathbf{l}_1, \dots, \mathbf{l}_m$ be representatives of the nonzero cosets of L/L' and suppose that $\tilde{w} + \phi_q(\mathbf{l}_i) \notin J(C(\mathbb{F}_q))$ for all $w \in W$ and $i = 1, \dots, m$. Then $J(C(\mathbb{Q})) \subset W + \phi(L')$.*

Proof. Suppose $P \in C(\mathbb{Q})$. Since $J(C(\mathbb{Q})) \subset W + \phi(L)$, we may write $J(P) = w + \phi(\mathbf{I})$ for some $\mathbf{I} \in L$. Now let $\mathbf{l}_0 = \mathbf{0}$, so that $\mathbf{l}_0, \dots, \mathbf{l}_m$ represent all cosets of L/L' . Then $\mathbf{I} = \mathbf{l}_i + \mathbf{I}'$ for some $\mathbf{I}' \in L'$ and $i = 0, \dots, m$. However, $\phi_q(\mathbf{I}') = 0$, or in other words, $\phi(\widetilde{\mathbf{I}'}) = 0$. Hence

$$J(\tilde{P}) = J(\widetilde{P}) = \tilde{w} + \phi_q(\mathbf{I}) = \tilde{w} + \phi_q(\mathbf{l}_i) + \phi_q(\mathbf{I}') = \tilde{w} + \phi_q(\mathbf{l}_i).$$

By hypothesis, $\tilde{w} + \phi_q(\mathbf{l}_i) \notin J(C(\mathbb{F}_q))$ for $i = 1, \dots, m$, so $i = 0$ and so $\mathbf{l}_i = 0$. Hence $J(P) = w + \mathbf{I}' \in W + L'$ as required. \square

We obtain a very long strictly decreasing sequence of lattices as in (11-1) by repeated application of Lemma 11.1. However, the conditions of Lemma 11.1 are unlikely to be satisfied for a prime q chosen at random. Here we give criteria that we have employed in practice to choose the primes q :

- (I) $\gcd(B, \#J(\mathbb{F}_q)) > (\#J(\mathbb{F}_q))^{0.6}$.
- (II) $L' \neq L$.
- (III) $\#W \cdot (\#L/L' - 1) < 2q$.
- (IV) $\tilde{w} + \phi_q(\mathbf{l}_i) \notin J(C(\mathbb{F}_q))$ for all $w \in W$ and $i = 1, \dots, m$.

The criteria (I)–(IV) are listed in the order in which we check them in practice. Criterion (IV) is just the criterion of the lemma. Criterion (II) ensures that L' is strictly smaller than L , otherwise we gain no new information. Although we would like L' to be strictly smaller than L , we do not want the index L/L' to be too large and this is reflected in Criteria (I) and (III). Note that the number of checks required by Criterion (IV) (or the lemma) is $\#W \cdot (\#L/L' - 1)$. If this number is large then Criterion (IV) is likely to fail. Let us look at this in probabilistic terms. Assume that the genus of C is 2. Then the probability that a random element of $J(\mathbb{F}_q)$ lies in the image of $C(\mathbb{F}_q)$ is about $\frac{1}{q}$. If $N = \#W \cdot (\#L/L' - 1)$ then the probability that Criterion (IV) is satisfied is about $(1 - q^{-1})^N$. Since $(1 - q^{-1})^q \sim e^{-1}$, we do not want N to be too large in comparison to q , and this explains the choice of $2q$ in Criterion (III).

We still have not justified Criterion (I). The computation involved in obtaining L' is a little expensive. Since we need to do this with many primes, we would like a way of picking only primes where this computation is not wasted, and in particular $\#L/L'$ is not too large. Now at every stage of our computations, L will be some element of our decreasing sequence (11-1) and so contained in $B\mathbb{Z}^r$. Criterion (I)

ensures that a “large chunk” of L will be in the kernel of $\phi_q : \mathbb{Z}^r \rightarrow J(\mathbb{F}_q)$ and so that $\#L/L'$ is not too large. The exponent 0.6 in Criterion (I) is chosen on the basis of computational experience.

12. Lower bounds for the size of rational points

In this section, we suppose that the strategy of Sections 10 and 11 succeeded in showing that $J(C(\mathbb{Q})) \subset W + \phi(L)$ for some lattice L of huge index in \mathbb{Z}^r , where W is the image in J of the set of known rational points in C . In this section we provide a lower bound for the size of rational points not belonging to the set of known rational points.

Lemma 12.1. *Let W be a finite subset of $J(\mathbb{Q})$, and let L be a sublattice of \mathbb{Z}^r . Suppose that $J(C(\mathbb{Q})) \subset W + \phi(L)$. Let μ_1 be a lower bound for $h - \hat{h}$ as in (1-2). Let*

$$\mu_2 = \max\left\{\sqrt{\hat{h}(w)} : w \in W\right\}.$$

Let M be the height-pairing matrix for the Mordell–Weil basis D_1, \dots, D_r and let $\lambda_1, \dots, \lambda_r$ be its eigenvalues. Let

$$\mu_3 = \min\{\sqrt{\lambda_j} : j = 1, \dots, r\}.$$

Let $m(L)$ be the Euclidean norm of the shortest nonzero vector of L , and suppose that $\mu_3 m(L) \geq \mu_2$. Then, for any $P \in C(\mathbb{Q})$, either $J(P) \in W$ or

$$h(J(P)) \geq (\mu_3 m(L) - \mu_2)^2 + \mu_1.$$

Note that $m(L)$ is called the *minimum of L* and can be computed using an algorithm of Fincke and Pohst [1985].

Proof. Suppose that $J(P) \notin W$. Then $J(P) = w + \phi(\mathbf{I})$ for some nonzero element $\mathbf{I} \in L$. In particular, if $\|\cdot\|$ denotes Euclidean norm then $\|\mathbf{I}\| \geq m(L)$.

We can write $M = N\Lambda N^t$ where N is orthogonal and Λ is the diagonal matrix with diagonal entries λ_i . Let $\mathbf{x} = \mathbf{I}N$ and write $\mathbf{x} = (x_1, \dots, x_r)$. Then

$$\hat{h}(\phi(\mathbf{I})) = \mathbf{I}M\mathbf{I}^t = \mathbf{x}\Lambda\mathbf{x}^t \geq \mu_3^2 \|\mathbf{x}\|^2 = \mu_3^2 \|\mathbf{I}\|^2 \geq \mu_3^2 m(L)^2.$$

Now recall that $D \mapsto \sqrt{\hat{h}(D)}$ defines a norm on $J(\mathbb{Q}) \otimes \mathbb{R}$ and so by the triangle inequality

$$\sqrt{\hat{h}(J(P))} \geq \sqrt{\hat{h}(\phi(\mathbf{I}))} - \sqrt{\hat{h}(w)} \geq \mu_3 m(L) - \mu_2.$$

The lemma now follows from (1-2). □

Remark 12.2. We can replace $\mu_3 m(L)$ with the minimum of L with respect to the height pairing matrix. This should lead to a very slight improvement. Since in practice our lattice L has very large index, computing the minimum of L with

Table 1

coset of $J(\mathbb{Q})/2J(\mathbb{Q})$	κ	unit rank of K_i	bound R for regulator of K_i	bound for $\log x$
0	1	12	1.8×10^{26}	1.0×10^{263}
D_1	-2α	21	6.2×10^{53}	7.6×10^{492}
D_2	$4 - 2\alpha$	25	1.3×10^{54}	2.3×10^{560}
D_3	$-4 - 2\alpha$	21	3.7×10^{55}	1.6×10^{498}
$D_1 + D_2$	$-2\alpha + \alpha^2$	21	1.0×10^{52}	3.2×10^{487}
$D_1 + D_3$	$2\alpha + \alpha^2$	25	7.9×10^{55}	5.1×10^{565}
$D_2 + D_3$	$-4 + \alpha^2$	21	3.7×10^{55}	1.6×10^{498}
$D_1 + D_2 + D_3$	$8\alpha - 2\alpha^3$	25	7.9×10^{55}	5.1×10^{565}

respect to the height pairing matrix may require the computation of the height pairing matrix to very great accuracy, and such a computation is inconvenient. We therefore prefer to work with the Euclidean norm on \mathbb{Z}^r .

13. Proofs of Theorems 1.1 and 1.2

The equation $Y^2 - Y = X^5 - X$ is transformed into

$$C : \quad 2y^2 = x^5 - 16x + 8, \tag{13-1}$$

via the change of variables $y = 4Y - 2$ and $x = 2X$ which preserves integrality. We shall work with the model (13-1). Let C be the smooth projective genus 2 curve with affine model given by (13-1), and let J be its Jacobian. Using MAGMA [Bosma et al. 1997] we know that $J(\mathbb{Q})$ is free of rank 3 with Mordell–Weil basis given by

$$D_1 = (0, 2) - \infty, \quad D_2 = (2, 2) - \infty, \quad D_3 = (-2, 2) - \infty.$$

The MAGMA programs used for this step are based on Stoll’s papers [1999; 2001; 2002].

Let $f = x^5 - 16x + 8$. Let α be a root of f . We shall choose for coset representatives of $J(\mathbb{Q})/2J(\mathbb{Q})$ the linear combinations $\sum_{i=1}^3 n_i D_i$ with $n_i \in \{0, 1\}$. Then

$$x - \alpha = \kappa \zeta^2,$$

where $\kappa \in \mathcal{H}$ and \mathcal{H} is constructed as in Lemma 3.1. We tabulate the κ corresponding to the $\sum_{i=1}^3 n_i D_i$ in Table 1.

Next we compute the bounds for $\log x$ given by Theorem 9.2 for each value of κ . We implemented our bounds in MAGMA. Here the Galois group of f is S_5 which implies that the fields K_1, K_2, K_3 corresponding to a particular κ are isomorphic. The unit ranks of K_i , the bounds for their regulators as given by Lemma 5.1, and the corresponding bounds for $\log x$ are tabulated also in Table 1.

A quick search reveals 17 rational points on C :

$$\begin{aligned} \infty, (-2, \pm 2), (0, \pm 2), (2, \pm 2), (4, \pm 22), (6, \pm 62), \\ (1/2, \pm 1/8), (-15/8, \pm 697/256), (60, \pm 9859). \end{aligned}$$

Let W denote the image of this set in $J(\mathbb{Q})$. Applying the implementation of the Mordell–Weil sieve due to Bruin and Stoll which is explained in Section 10 we obtain that $J(C(\mathbb{Q})) \subseteq W + BJ(\mathbb{Q})$ where

$$\begin{aligned} B &= 4449329780614748206472972686179940 \\ &652515754483274306796568214048000 \\ &= 2^8 \cdot 3^4 \cdot 5^3 \cdot 7^3 \cdot 11^2 \cdot 13^2 \cdot 17^2 \cdot 19 \cdot 23 \cdot 29 \cdot 31^2 \cdot \prod_{\substack{37 \leq p \leq 149 \\ p \neq 107}} p. \end{aligned}$$

For this computation, we used “deep” information modulo $2^9, 3^6, 5^4, 7^3, 11^3, 13^2, 17^2, 19^2$, and “flat” information from all primes $p < 50000$ such that $\#J(\mathbb{F}_p)$ is 500-smooth (but keeping only information coming from the maximal 150-smooth quotient group of $J(\mathbb{F}_p)$). Recall that an integer is called A -smooth if all its prime divisors are $\leq A$. This computation took about 7 hours on a 2 GHz Intel Core 2 CPU.

We now apply the new extension of the Mordell–Weil sieve which is explained in Section 11. We start with $L_0 = B\mathbb{Z}^3$ where B is as above. We successively apply Lemma 11.1 using all primes $q < 10^6$ which are primes of good reduction and satisfy criteria (I)–(IV) of Section 11. There are 78,498 primes less than 10^6 . Of these, we discard 2, 139, 449 as they are primes of bad reduction for C . This leaves us with 78,495 primes. Of these, Criterion (I) fails for 77,073 of them, Criterion (II) fails for 220 of the remaining, Criterion (III) fails for 43 primes that survive Criteria (I) and (II), and Criterion (IV) fails for 237 primes that survive Criteria (I)–(III). Altogether, only 922 primes $q < 10^6$ satisfy Criteria (I)–(IV) and increase the index of L .

The index of the final L in \mathbb{Z}^3 is approximately 3.32×10^{3240} . This part of the computation lasted about 37 hours on a 2.8 GHZ Dual-Core AMD Opteron.

Let μ_1, μ_2, μ_3 be as in the notation of Lemma 12.1. Using MAGMA we find $\mu_1 = 2.677, \mu_2 = 2.612$ and $\mu_3 = 0.378$ (to 3 decimal places). The shortest vector of the final lattice L is of Euclidean length approximately 1.156×10^{1080} (it should be no surprise that this is roughly the cube root of the index of L in \mathbb{Z}^3). By Lemma 12.1 if $P \in C(\mathbb{Q})$ is not one of the 17 known rational points then

$$h(J(P)) \geq 1.9 \times 10^{2159}.$$

If P is an integral point, then $h(J(P)) = \log 2 + 2 \log x(P)$. Thus

$$\log x(P) \geq 0.95 \times 10^{2159}.$$

This contradicts the bounds for $\log x$ in Table 1 and shows that the integral point P must be one of the 17 known rational points. This completes the proof of Theorem 1.1. The proof of Theorem 1.2 is similar and we omit the details.

The reader can find the MAGMA programs for verifying the above computations at: <http://www.warwick.ac.uk/staff/S.Siksek/progs/intpoint/>.

Acknowledgments

We are grateful to the referee and editors for many useful comments, and to Mr. Homero Gallegos–Ruiz for spotting many misprints.

References

- [Avanesov 1966/1967] È. T. Avanesov, “Lösung eines Problems der figurierten Zahlen”, *Acta Arith.* **12** (1966/1967), 409–420. In Russian. MR 35 #6619 Zbl 0153.06403
- [Baker 1969] A. Baker, “Bounds for the solutions of the hyperelliptic equation”, *Proc. Cambridge Philos. Soc.* **65** (1969), 439–444. MR 38 #3226 Zbl 0174.33803
- [Bilu 1995] Y. Bilu, “Effective analysis of integral points on algebraic curves”, *Israel J. Math.* **90**:1-3 (1995), 235–252. MR 96e:11082 Zbl 0840.11028
- [Bilu and Hanrot 1998] Y. F. Bilu and G. Hanrot, “Solving superelliptic Diophantine equations by Baker’s method”, *Compositio Math.* **112**:3 (1998), 273–312. MR 99d:11028 Zbl 0915.11065
- [Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, “The Magma algebra system. I. The user language”, *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR 1484478 Zbl 0898.68039
- [Brindza 1984] B. Brindza, “On S -integral solutions of the equation $y^m = f(x)$ ”, *Acta Math. Hungar.* **44**:1-2 (1984), 133–139. MR 85m:11017 Zbl 0552.10009
- [Brindza 1991] B. Brindza, “On a special superelliptic equation”, *Publ. Math. Debrecen* **39**:1-2 (1991), 159–162. MR 92j:11029 Zbl 0749.11024
- [Bruin 1999] N. R. Bruin, *Chabauty methods and covering techniques applied to generalized Fermat equations*, PhD Thesis, University of Leiden, Leiden, 1999.
- [Bruin 2003] N. Bruin, “Chabauty methods using elliptic curves”, *J. Reine Angew. Math.* **562** (2003), 27–49. MR 2004j:11051 Zbl 1135.11320
- [Bruin and Elkies 2002] N. Bruin and N. D. Elkies, “Trinomials $ax^7 + bx + c$ and $ax^8 + bx + c$ with Galois groups of order 168 and $8 \cdot 168$ ”, pp. 172–188 in *Algorithmic number theory* (Sydney, 2002), edited by F. Claus and K. D. R., Lecture Notes in Comput. Sci. **2369**, Springer, Berlin, 2002. MR 2005d:11094 Zbl 1058.11044
- [Bruin and Stoll 2008a] N. Bruin and M. Stoll, “Deciding existence of rational points on curves: an experiment”, *Experiment. Math.* **17** (2008), 181–189. arXiv math/0604524
- [Bruin and Stoll 2008b] N. Bruin and M. Stoll, “Two-cover descent on hyperelliptic curves”, preprint, 2008. arXiv 0803.2052
- [Bruin and Stoll \geq 2008] N. Bruin and M. Stoll, “The Mordell–Weil sieve: proving the non-existence of rational points on curves”. In preparation.

- [Bugeaud 1997] Y. Bugeaud, “Bounds for the solutions of superelliptic equations”, *Compositio Math.* **107**:2 (1997), 187–219. MR 98c:11025 Zbl 0886.11016
- [Bugeaud and Győry 1996] Y. Bugeaud and K. Győry, “Bounds for the solutions of unit equations”, *Acta Arith.* **74**:1 (1996), 67–80. MR 97b:11045 Zbl 0861.11023
- [Bugeaud et al. 2006] Y. Bugeaud, M. Mignotte, and S. Siksek, “Classical and modular approaches to exponential Diophantine equations. I. Fibonacci and Lucas perfect powers”, *Ann. of Math.* (2) **163**:3 (2006), 969–1018. MR 2007f:11031 Zbl 1113.11021
- [Cassels and Flynn 1996] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus 2*, London Mathematical Society Lecture Note Series **230**, Cambridge University Press, Cambridge, 1996. MR 97i:11071 Zbl 0857.14018
- [Evertse and Tijdeman 2007] J.-H. Evertse and R. Tijdeman, “Some open problems about Diophantine equations”, 2007, available at <http://www.math.leidenuniv.nl/~evertse/07-workshop-problems.pdf>.
- [Fielder and Alford 1998] D. C. Fielder and C. O. Alford, “Observations from computer experiments on an integer equation”, pp. 93–103 in *Applications of Fibonacci numbers* (Graz, 1996), vol. 7, edited by G. E. Bergum et al., Kluwer Acad. Publ., Dordrecht, 1998. MR 1638435 Zbl 0913.11014
- [Fincke and Pohst 1985] U. Fincke and M. Pohst, “Improved methods for calculating vectors of short length in a lattice, including a complexity analysis”, *Math. Comp.* **44**:170 (1985), 463–471. MR 86e:11050 Zbl 0556.10022
- [Flynn 1997] E. V. Flynn, “A flexible method for applying Chabauty’s theorem”, *Compositio Math.* **105**:1 (1997), 79–94. MR 97m:11083 Zbl 0882.14009
- [Flynn 2004] E. V. Flynn, “The Hasse principle and the Brauer–Manin obstruction for curves”, *Manuscripta Math.* **115**:4 (2004), 437–466. MR 2005j:11047 Zbl 1069.11023
- [Flynn and Smart 1997] E. V. Flynn and N. P. Smart, “Canonical heights on the Jacobians of curves of genus 2 and the infinite descent”, *Acta Arith.* **79**:4 (1997), 333–352. MR 98f:11066 Zbl 0895.11026
- [Flynn and Wetherell 1999] E. V. Flynn and J. L. Wetherell, “Finding rational points on bielliptic genus 2 curves”, *Manuscripta Math.* **100**:4 (1999), 519–533. MR 2001g:11098 Zbl 1029.11024
- [Flynn and Wetherell 2001] E. V. Flynn and J. L. Wetherell, “Covering collections and a challenge problem of Serre”, *Acta Arith.* **98**:2 (2001), 197–205. MR 2002b:11088 Zbl 1049.11066
- [Kiss 1988] P. Kiss, “On the number of solutions of the Diophantine equation $\binom{x}{p} = \binom{y}{2}$ ”, *Fibonacci Quart.* **26**:2 (1988), 127–130. MR 89f:11050 Zbl 0641.10016
- [Landau 1918] E. Landau, “Verallgemeinerung eines Pólyaschen Satzes auf algebraische Zahlkörper.”, *Gött. Nachr.* (1918), 478–488. JFM 46.0267.01
- [Lind 1968] D. A. Lind, “The quadratic field $Q(\sqrt{5})$ and a certain Diophantine equation”, *Fibonacci Quart.* **6**:3 (1968), 86–93. MR 38 #112 Zbl 0174.33801
- [Matveev 2000] E. M. Matveev, “An explicit lower bound for a homogeneous rational linear form in logarithms of algebraic numbers. II”, *Izv. Ross. Akad. Nauk Ser. Mat.* **64**:6 (2000), 125–180. English translation in *Izv. Math.* **64** (2000), no. 6, 1217–1269. MR 2002e:11091 Zbl 1013.11043
- [Mignotte and Pethő 1999] M. Mignotte and A. Pethő, “On the Diophantine equation $x^p - x = y^q - y$ ”, *Publ. Mat.* **43**:1 (1999), 207–216. MR 2000d:11044 Zbl 0949.11022
- [Mordell 1963] L. J. Mordell, “On the integer solutions of $y(y+1) = x(x+1)(x+2)$ ”, *Pacific J. Math.* **13** (1963), 1347–1351. MR 27 #3590 Zbl 0124.27402
- [Pethő and de Weger 1986] A. Pethő and B. M. M. de Weger, “Products of prime powers in binary recurrence sequences. I. The hyperbolic case, with an application to the generalized Ramanujan–Nagell equation”, *Math. Comp.* **47**:176 (1986), 713–727. MR 87m:11027a Zbl 0623.10011

- [Pintér 1995] Á. Pintér, “A note on the Diophantine equation $\binom{x}{4} = \binom{y}{2}$ ”, *Publ. Math. Debrecen* **47**:3-4 (1995), 411–415. MR 96i:11027 Zbl 0856.11019
- [Poonen 2006] B. Poonen, “Heuristics for the Brauer–Manin obstruction for curves”, *Experiment. Math.* **15**:4 (2006), 415–420. MR 2008d:11062 Zbl 05196200
- [Poonen and Schaefer 1997] B. Poonen and E. F. Schaefer, “Explicit descent for Jacobians of cyclic covers of the projective line”, *J. Reine Angew. Math.* **488** (1997), 141–188. MR 98k:11087 Zbl 0888.11023
- [Poulakis 1991] D. Poulakis, “Solutions entières de l’équation $Y^m = f(X)$ ”, *Sém. Théor. Nombres Bordeaux* (2) **3**:1 (1991), 187–199. MR 93a:11025 Zbl 0733.11009
- [Schaefer 1995] E. F. Schaefer, “2-descent on the Jacobians of hyperelliptic curves”, *J. Number Theory* **51**:2 (1995), 219–232. MR 96c:11066 Zbl 0832.14016
- [Schmidt 1992] W. M. Schmidt, “Integer points on curves of genus 1”, *Compositio Math.* **81**:1 (1992), 33–59. MR 93e:11076 Zbl 0747.11026
- [Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1986. MR 95m:11054 Zbl 0585.14026
- [Singmaster 1975] D. Singmaster, “Repeated binomial coefficients and Fibonacci numbers”, *Fibonacci Quart.* **13**:4 (1975), 295–298. MR 54 #224 Zbl 0324.05007
- [Smart 1998] N. P. Smart, *The algorithmic resolution of Diophantine equations*, London Mathematical Society Student Texts **41**, Cambridge University Press, Cambridge, 1998. MR 2000c:11208 Zbl 0907.11001
- [Sprindžuk 1977] V. G. Sprindžuk, “The arithmetic structure of integer polynomials and class numbers”, *Trudy Mat. Inst. Steklov.* **143** (1977), 152–174, 210. MR 58 #589
- [Stoll 1999] M. Stoll, “On the height constant for curves of genus two”, *Acta Arith.* **90**:2 (1999), 183–201. MR 2000h:11069 Zbl 0932.11043
- [Stoll 2001] M. Stoll, “Implementing 2-descent for Jacobians of hyperelliptic curves”, *Acta Arith.* **98**:3 (2001), 245–277. MR 2002b:11089 Zbl 0972.11058
- [Stoll 2002] M. Stoll, “On the height constant for curves of genus two. II”, *Acta Arith.* **104**:2 (2002), 165–182. MR 2003f:11093 Zbl 1139.11318
- [Stoll 2007] M. Stoll, “Finite descent obstructions and rational points on curves”, *Algebra Number Theory* **1**:4 (2007), 349–391. MR 2008i:11086
- [Stroeker and de Weger 1999] R. J. Stroeker and B. M. M. de Weger, “Elliptic binomial Diophantine equations”, *Math. Comp.* **68**:227 (1999), 1257–1281. MR 99i:11122 Zbl 0920.11014
- [Stroeker and Tzanakis 2003] R. J. Stroeker and N. Tzanakis, “Computing all integer solutions of a genus 1 equation”, *Math. Comp.* **72**:244 (2003), 1917–1933. MR 2004b:11037 Zbl 1089.11019
- [Voutier 1995] P. M. Voutier, “An upper bound for the size of integral solutions to $Y^m = f(X)$ ”, *J. Number Theory* **53**:2 (1995), 247–271. MR 96f:11049 Zbl 0842.11008
- [Voutier 1996] P. Voutier, “An effective lower bound for the height of algebraic numbers”, *Acta Arith.* **74**:1 (1996), 81–95. MR 96j:11098 Zbl 0838.11065
- [de Weger 1996] B. M. M. de Weger, “A binomial Diophantine equation”, *Quart. J. Math. Oxford Ser. (2)* **47**:186 (1996), 221–231. MR 97c:11041 Zbl 0863.11022
- [de Weger 1997] B. M. M. de Weger, “Equal binomial coefficients: some elementary considerations”, *J. Number Theory* **63**:2 (1997), 373–386. MR 98b:11027 Zbl 0873.11023
- [Wetherell 1997] J. L. Wetherell, *Bounding the Number of Rational Points on Certain Curves of High Rank*, PhD Thesis, University of California, Berkeley, 1997.

Communicated by Bjorn Poonen

Received 2008-01-28

Revised 2008-09-02

Accepted 2008-09-12

- bugeaud@math.u-strasbg.fr *Université Louis Pasteur, U. F. R. de mathématiques,
7, rue René Descartes, 67084 Strasbourg Cedex, France
<http://www-irma.u-strasbg.fr/~bugeaud/>*
- mignotte@math.u-strasbg.fr *Université Louis Pasteur, U. F. R. de mathématiques,
7, rue René Descartes, 67084 Strasbourg Cedex, France*
- s.siksek@warwick.ac.uk *Institute of Mathematics, University of Warwick,
Coventry CV4 7AL, United Kingdom
<http://www.warwick.ac.uk/~maseap/>*
- Michael.Stoll@uni-bayreuth.de *Mathematisches Institut, Universität Bayreuth,
95440 Bayreuth, Germany
<http://www.mathe2.uni-bayreuth.de/stoll/>*
- tengely@math.klte.hu *Institute of Mathematics, University of Debrecen, Number
Theory Research Group, Hungarian Academy of Sciences,
P.O.Box 12, 4010 Debrecen, Hungary
<http://www.math.klte.hu/~tengely/>*

Smooth curves having a large automorphism p -group in characteristic $p > 0$

Michel Matignon and Magali Rocher

Let k be an algebraically closed field of characteristic $p > 0$ and C a connected nonsingular projective curve over k with genus $g \geq 2$. This paper continues our study of *big actions*, that is, pairs (C, G) where G is a p -subgroup of the k -automorphism group of C such that $|G|/g > 2p/(p-1)$. If G_2 denotes the second ramification group of G at the unique ramification point of the cover $C \rightarrow C/G$, we display necessary conditions on G_2 for (C, G) to be a big action, which allows us to pursue the classification of big actions.

Our main source of examples comes from the construction of curves with many rational points using ray class field theory for global function fields, as initiated by J.-P. Serre and continued by Lauter and by Auer. In particular, we obtain explicit examples of big actions with G_2 abelian of large exponent.

1. Introduction

This is the first of a set of three papers (together with [Rocher 2008a; 2008b]) whose main object is to study G -actions on connected nonsingular projective curves of genus $g \geq 2$ defined over an algebraically closed field of characteristic $p > 0$, when G is a p -group such that $|G| > 2g p/(p-1)$. One of our aims is to display some universal families and discuss the corresponding deformation space.

For more than a century, the study of finite groups G acting faithfully on smooth complete curves defined over an algebraically closed field k of characteristic $p \geq 0$ has produced a vast literature. Already back in the nineteenth century progress was made in the case of characteristic zero, with the works of Schwartz, Klein, Hurwitz, Wiman and others. The full automorphism group of a compact Riemann surface of genus $g \geq 2$ was proved by Hurwitz [1892] to be finite and of order at most $84(g-1)$.

MSC2000: primary 14H37; secondary 11R37, 11G20, 14H10.

Keywords: automorphisms, curves, p -groups, ray class fields, Artin–Schreier–Witt theory.

Work supported in part by the European Community's Sixth Framework Programme under Contract MRTN-CT-2006-035495.

An open question concerns the classification of full automorphism groups of compact Riemann surfaces C of fixed genus $g \geq 2$. This classification has been partially achieved for large automorphism groups G , “large” meaning that the order of G is greater than $4(g - 1)$ [Kulkarni 1997]. This lower bound imposes strict restrictions on the genus g_0 of the quotient curve C/G , namely $g_0 = 0$, on the number r of points of C/G ramified in C , namely $r \in \{3, 4\}$, and on the corresponding ramification indices; see [Kulkarni 1997; Breuer 2000, Lemma 3.18]. Following this work, Magaard et al. [2002] exhibited the list of large groups $\text{Aut}(C)$ of compact Riemann surfaces of genus g up to $g = 10$, determining in each case the dimension and number of components of the corresponding loci in the moduli space of genus g curves.

General results on Hurwitz spaces and other moduli spaces parametrizing deformations have been obtained in the case of characteristic zero and extended to positive characteristic $p > 0$ when p does not divide the order of the automorphism group; see [Bertin and Romagny 2008], for instance. For instance, if C is a compact Riemann surface with genus $g \geq 2$ and G an automorphism group of C , the deformations of the cover $\varphi : C \rightarrow C/G$ are parametrized by a moduli space of dimension $3g_0 - 3 + |\mathcal{B}| + \dim \text{Aut}(C/G - \mathcal{B})$, where g_0 is the genus of C/G and \mathcal{B} the branch locus of φ . By the Hurwitz genus formula, g_0 only depends on $|G|$, g , $|\mathcal{B}|$ and the orders of the inertia groups. All these results are no longer true in positive characteristic $p > 0$ when φ is wildly ramified. Likewise, in positive characteristic $p > 0$, the Hurwitz bound is no longer true for automorphism groups G whose order is not prime to p . The finiteness result still holds [Schmid 1938] but the Hurwitz linear bound is replaced with biquadratic bounds [Stichtenoth 1973a; 1973b]. These biquadratic bounds are optimal: so, in positive characteristic, the automorphism groups may be very large compared with the case of characteristic zero, as a result of wild ramification.

Wild ramification points also contribute to the dimension of the tangent space to the global infinitesimal deformation functor of a curve C together with an automorphism group G , and it is precisely this that makes computations difficult. Following [Bertin and Mézard 2000], in the case where G is cyclic of order p , Pries [2005] and Kontogeorgis [2007] have obtained lower and upper bounds for the dimension of the tangent space, with explicit computations in some special cases, in particular when G is an abelian p -group. (See also [Cornelissen and Kato 2003]).

To rigidify the situation in characteristic $p > 0$ as has been done in characteristic zero, one idea is to consider large automorphism p -groups. From [Nakajima 1987] we deduce that if G is a p -subgroup of $\text{Aut}_k(C)$ such that $|G| > 2pg/(p-1)$, the Hasse–Witt invariant of C is zero. The Deuring–Shafarevich formula (see [Bouw 2000], for instance) then implies that the genus of the quotient curve C/G is zero

and that the branch locus of the cover $C \rightarrow C/G$ is reduced to one point. From now on, we define a *big action* as a pair (C, G) where G is a p -subgroup of $\text{Aut}_k(C)$ such that $|G|/g > 2p/(p-1)$.

Outline of the paper. Let (C, G) be a big action with $g \geq 2$. As shown in [Lehr and Matignon 2005], there is a point of C , say ∞ , such that G is equal to the wild inertia subgroup G_1 of G at ∞ . Let G_2 be the second ramification group of G at ∞ in lower notation. The quotient curve C/G_2 is isomorphic to the projective line \mathbb{P}_k^1 and the quotient group G/G_2 acts as a group of translations of \mathbb{P}_k^1 fixing ∞ , through $X \rightarrow X + y$, where y runs over a subgroup V of k . In this way, the group G appears as an extension of G_2 by the p -elementary abelian group V via the exact sequence

$$0 \longrightarrow G_2 \longrightarrow G = G_1 \longrightarrow V \simeq (\mathbb{Z}/p\mathbb{Z})^v \longrightarrow 0.$$

The purpose of this paper is twofold: to give necessary conditions on G_2 for (C, G) to be a big action, and to display realizations of big actions with G_2 abelian of large exponent. We gather here the main results of the first part (Sections 2–5):

Theorem. *Let (C, G) be a big action with $g \geq 2$.*

1. *Let H be a subgroup of G . Then C/H has genus 0 if and only if $H \supset G_2$ (Lemma 2.4.1).*
2. *Let H be a normal subgroup of G such that $H \subsetneq G_2$. Then $(C/H, G/H)$ is a big action with second ramification group $(G/H)_2 = G_2/H$ (Lemma 2.4.2).*
3. *The group G_2 is equal to $D(G)$, the commutator subgroup of G (Theorem 2.7). In particular, G cannot be abelian.*
4. *The group G_2 cannot be cyclic unless G_2 has order p (Theorem 5.1).*
5. *If $|G|/g^2 \geq 4/(p^2 - 1)^2$, then G_2 is an elementary abelian p -group with order dividing p^3 (Proposition 4.1).*

These results highlight the major role played by G_2 in the study of big actions. They are also crucial in pursuing the classification of big actions initiated in [Lehr and Matignon 2005]. The companion paper [Rocher 2008a] is devoted to big actions with a p -elementary abelian G_2 , and its results led to the classification of the big actions satisfying $|G|/g^2 \geq 4/(p^2 - 1)^2$, in [Rocher 2008b].

After exploring restrictions on G_2 , the second part of the paper is devoted to examples of big actions with G_2 abelian, knowing that we do not know yet examples of big actions with a nonabelian G_2 .

In Section 6, following [Lauter 1999] and [Auer 1999], we consider the maximal abelian extension K_S^m of $K := \mathbb{F}_q(X)$ (where $q = p^e$) that is unramified outside $X = \infty$, completely split over the set S of the finite rational places and whose

conductor is smaller than $m\infty$, with $m \in \mathbb{N}$. Class field theory gives a precise description of the Galois group $G_S(m)$ of this extension. Moreover, it follows from the uniqueness and the maximality of K_S^m that the group of translations $X \mapsto X + y$ ($y \in \mathbb{F}_q$) extends to a p -group of \mathbb{F}_q -automorphisms of K_S^m , say $G(m)$, with the exact sequence

$$0 \longrightarrow G_S(m) \longrightarrow G(m) \longrightarrow \mathbb{F}_q \longrightarrow 0.$$

This provides examples of big actions whose $G_2 = G_S(m)$ is abelian of exponent as large as we want, but it also relates the problem of big actions to the search of algebraic curves with many rational points compared with their genera.

In Section 7 we use the Katz–Gabber theorem to highlight the link between big actions on curves and an analogous ramification condition for finite p -groups acting on $k((z))$.

Notation and preliminary remarks. Let k be an algebraically closed field of characteristic $p > 0$. We denote by F the Frobenius endomorphism for a k -algebra. Then \wp means the Frobenius operator minus identity. We denote by $k\{F\}$ the k -subspace of $k[X]$ generated by the polynomials $F^i(X)$, with $i \in \mathbb{N}$. It is a ring under the composition. Furthermore, for all α in k , $F\alpha = \alpha^p F$. The elements of $k\{F\}$ are the additive polynomials, i.e. the polynomials $P(X)$ of $k[X]$ such that for all α and β in k , $P(\alpha + \beta) = P(\alpha) + P(\beta)$. A separable polynomial is additive if and only if the set of its roots is a subgroup of k [Goss 1996, Chapter 1].

Let $f(X)$ be a polynomial of $k[X]$. There is a unique polynomial $\text{red}(f)(X)$ in $k[X]$, called the reduced representative of f , which is p -power free, (meaning that $\text{red}(f)(X) \in \bigoplus_{(i,p)=1} k X^i$) and such that $\text{red}(f)(X) = f(X) \bmod \wp(k[X])$. We say that the polynomial f is reduced mod $\wp(k[X])$ if and only if it coincides with its reduced representative $\text{red}(f)$. The equation $W^p - W = f(X)$ defines a p -cyclic étale cover of the affine line that we denote by C_f . Conversely, any p -cyclic étale cover of the affine line $\text{Spec } k[X]$ corresponds to a curve C_f where f is a polynomial of $k[X]$; see [Milne 1980, III.4.12, p. 127]. By Artin–Schreier theory, the covers C_f and $C_{\text{red}(f)}$ define the same p -cyclic covers of the affine line. The curve C_f is irreducible if and only if $\text{red}(f) \neq 0$.

Throughout the text, C always denotes a nonsingular smooth projective curve with genus g and $\text{Aut}_k(C)$ means its k -automorphism group. Our main references for ramification theory are [Serre 1968] and [Auer 1999].

2. First results on big actions

To pinpoint the background of our work, we begin by collecting and completing the first results on big actions already obtained in [Lehr and Matignon 2005]. A *big action* is a curve endowed with a big automorphism p -group. The first task is to recall what we mean by big.

Definition 2.1. Let G be a subgroup of $\text{Aut}_k(C)$. We say that the pair (C, G) is a big action if G is a finite p -group, if $g \neq 0$ and if

$$\frac{|G|}{g} > \frac{2p}{p-1}. \tag{2-1}$$

Proposition 2.2 [Lehr and Matignon 2005]. *Assume that (C, G) is a big action with $g \geq 2$. Then there is a point of C (say ∞) such that G is the wild inertia subgroup G_1 of G at ∞ . Moreover, the quotient C/G is isomorphic to the projective line \mathbb{P}_k^1 and the ramification locus (respectively branch locus) of the cover $\pi : C \rightarrow C/G$ is the point ∞ (respectively $\pi(\infty)$). For all $i \geq 0$, we denote by G_i the i -th lower ramification group of G at ∞ .*

1. G_2 is nontrivial and it is strictly included in G_1 .
2. The Hurwitz genus formula applied to $C \rightarrow C/G$ reads

$$2g = \sum_{i \geq 2} (|G_i| - 1). \tag{2-2}$$

Thus, (2-1) can be written as $|G| > 2g/(p-1)p$, with $2g/(p-1) \in \mathbb{N}^*$.

3. The quotient curve C/G_2 is isomorphic to the projective line \mathbb{P}_k^1 . Moreover, the quotient group G/G_2 acts as a group of translations of the affine line $C/G_2 - \{\infty\} = \text{Spec } k[X]$, through $X \mapsto X + y$, where y runs over a subgroup V of k . Then V is an \mathbb{F}_p -vector subspace of k . We denote by v its dimension. Thus, we obtain the exact sequence

$$0 \longrightarrow G_2 \longrightarrow G = G_1 \xrightarrow{\pi} V \simeq (\mathbb{Z}/p\mathbb{Z})^v \longrightarrow 0,$$

where $\pi : G \rightarrow V$ is defined by $g \mapsto g(X) - X$.

4. Let H be a normal subgroup of G such that $g_{C/H} > 0$. Then $(C/H, G/H)$ is also a big action. Moreover, the group G/H fixes the image of ∞ in the cover $C \rightarrow C/H$. In particular, if $g_{C/H} = 1$, then $p = 2$, C/H is birational to the curve $W^2 + W = X^3$ and G/H is isomorphic to Q_8 , the quaternion group of order 8 (see [Silverman 1986, Appendix A, Proposition 1.2]).

Remark 2.3. 1. For $g = 1$, one can find big actions (C, G) such that G is not included in a decomposition group of $\text{Aut}_k(C)$ as in Proposition 2.2.

2. Let (C, G) be a big action. Call L the function field of C and $k(X) = L^{G_2}$. As seen above, the Galois extension $L/k(X)$ is only ramified at $X = \infty$. Therefore, the support of the conductor of $L/k(X)$, as defined in [Serre 1968, chapitre 15, corollaire 2] reduces to the place ∞ . So, in what follows, we systematically confuse the conductor $m \infty$ with its degree m . In this case, one can also see m as the smallest integer $n > 0$ such that the n -th upper ramification group G^n of G at ∞ is trivial; see [Auer 2000, I.3].

The following lemma generalizes and completes the last part of Proposition 2.2.

Lemma 2.4. *Let G a finite p -subgroup of $\text{Aut}_k(C)$. We assume that the quotient curve C/G is isomorphic to \mathbb{P}_k^1 and that there is a point of C (say ∞) such that G is the wild inertia subgroup G_1 of G at ∞ . We also assume that the ramification locus of the cover $\pi : C \rightarrow C/G$ is the point ∞ , and the branch locus is $\pi(\infty)$. Let G_2 be the second ramification group of G at ∞ and H a subgroup of G . Then:*

1. C/H is isomorphic to \mathbb{P}_k^1 if and only if $H \supset G_2$.
2. In particular, if (C, G) is a big action with $g \geq 2$ and if H is a normal subgroup of G such that $H \subsetneq G_2$, then $g_{C/H} > 0$ and $(C/H, G/H)$ is also a big action. Moreover, its second ramification group is $(G/H)_2 = G_2/H$.

Proof. Applied to the cover $C \rightarrow C/G \simeq \mathbb{P}_k^1$, the Hurwitz genus formula (see for instance [Stichtenoth 1993]) yields $2(g - 1) = 2|G| (g_{C/G} - 1) + \sum_{i \geq 0} (|G_i| - 1)$. When applied to the cover $C \rightarrow C/H$, it yields $2(g - 1) = 2|H| (g_{C/H} - 1) + \sum_{i \geq 0} (|H \cap G_i| - 1)$. Since $H \subset G = G_0 = G_1$, it follows that

$$2|H|g_{C/H} = -2(|G| - |H|) + \sum_{i \geq 0} (|G_i| - |H \cap G_i|) = \sum_{i \geq 2} (|G_i| - |H \cap G_i|).$$

Therefore, $g_{C/H} = 0$ if and only if for all $i \geq 2$, $G_i = H \cap G_i$, i.e., $G_i \subset H$, which is equivalent to $G_2 \subset H$, proving 1.

Together with part 1, Proposition 2.2.4 shows that $(C/H, G/H)$ is a big action. Then $G = G_1 \supsetneq G_2$ and $G/H = (G/H)_1 \supsetneq (G/H)_2$. Since the first jump always coincides in lower and upper ramification, it follows that $G_2 = G^2$ and $(G/H)_2 = (G/H)^2$. By [Serre 1968, chapitre IV, proposition 14], we obtain $(G/H)_2 = (G/H)^2 = G^2H/H = G_2H/H = G_2/H$. □

The very first step in studying big actions is to give a precise description of them when $G_2 \simeq \mathbb{Z}/p\mathbb{Z}$. The following proposition collects and reformulates the results already obtained for this case.

Proposition 2.5 [Lehr and Matignon 2005, Propositions 5.5, 8.1, 8.3]. *Let (C, G) be a big action, with $g \geq 2$, such that $G_2 \simeq \mathbb{Z}/p\mathbb{Z}$.*

1. Then C is birational to the curve $C_f : W^p - W = f(X) = X S(X) + c X \in k[X]$, where S in $k\{F\}$ is an additive polynomial with degree $s \geq 1$ in F . If we denote by m the degree of f , then $m = 1 + p^s = i_0$, where $i_0 \geq 2$ is the integer such that

$$G = G_0 = G_1 \supsetneq G_2 = G_3 = \dots = G_{i_0} \supsetneq G_{i_0+1} = \dots$$

2. Write $S(F) = \sum_{j=0}^s a_j F^j$, with $a_s \neq 0$. Define (following [Elkies 1999b, Section 4]) the palindromic polynomial of f as the additive polynomial

$$\text{Ad}_f := a_s^{-p^s} F^s \left(\sum_{j=0}^s a_j F^j + F^{-j} a_j \right).$$

The set of roots of Ad_f , denoted by $Z(\text{Ad}_f)$, is an \mathbb{F}_p -vector subspace of k , isomorphic to $(\mathbb{Z}/p\mathbb{Z})^{2s}$. Moreover,

$$Z(\text{Ad}_f) = \{y \in k, f(X + y) - f(X) = 0 \pmod{\wp(k[X])}\}.$$

3. Let $A_{\infty,1}$ be the wild inertia subgroup of $\text{Aut}_k(C)$ at ∞ . Then $A_{\infty,1}$ is a central extension of $\mathbb{Z}/p\mathbb{Z}$ by the elementary abelian p -group $Z(\text{Ad}_f)$ which can be identified with a subgroup of translations $\{X \rightarrow X + y, y \in k\}$ of the affine line. Furthermore, if we denote by $Z(A_{\infty,1})$ the center of $A_{\infty,1}$ and by $D(A_{\infty,1})$ its commutator subgroup, $Z(A_{\infty,1}) = D(A_{\infty,1}) = \langle \sigma \rangle$, where $\sigma(X) = X$ and $\sigma(W) = W + 1$. Thus, we get the exact sequence

$$0 \longrightarrow Z(A_{\infty,1}) = D(A_{\infty,1}) \simeq \mathbb{Z}/p\mathbb{Z} \longrightarrow A_{\infty,1} \xrightarrow{\pi} Z(\text{Ad}_f) \simeq (\mathbb{Z}/p\mathbb{Z})^{2s} \longrightarrow 0,$$

where $\pi : A_{\infty,1} \rightarrow Z(\text{Ad}_f) \simeq (\mathbb{Z}/p\mathbb{Z})^{2s}$ is defined by $g \mapsto g(X) - X$. For $p > 2$, $A_{\infty,1}$ is the unique extraspecial group with exponent p and order p^{2s+1} . (The case $p = 2$ is more complicated; see [Lehr and Matignon 2005, 4.1]).

4. There exists an \mathbb{F}_p -vector space $V \subset Z(\text{Ad}_f) \simeq (\mathbb{Z}/p\mathbb{Z})^{2s}$ such that $G = \pi^{-1}(V) \subset A_{\infty,1}$ and such that we get the exact sequence

$$0 \longrightarrow G_2 \simeq \mathbb{Z}/p\mathbb{Z} \longrightarrow G \xrightarrow{\pi} V \longrightarrow 0.$$

Remark 2.6. Proposition 2.5 still holds for big actions (C, G) with $g = 1$ when G is included in a decomposition group of $\text{Aut}_k(C)$; see [Lehr and Matignon 2005, Proposition 8.3]. In particular, this is true for the pair $(C/H, G/H)$ when (C, G) is a big action with $g \geq 2$ and H a normal subgroup of G such that $g_{C/H} = 1$ (see Proposition 2.2.4).

Therefore, the key idea in studying big actions is to use Proposition 2.2.4 and Lemma 2.4.2 to go back to the well-known situation described above. This motivates the following result:

Theorem 2.7. Let (C, G) be a big action with $g \geq 2$. Let \mathcal{G} be a normal subgroup in G such that \mathcal{G} is strictly included in G_2 . There exists a group H , normal in G , such that $\mathcal{G} \subset H \subsetneq G_2$ and $[G_2 : H] = p$. In this case, $(C/H, G/H)$ enjoys the following properties.

1. The pair $(C/H, G/H)$ is a big action and the exact sequence

$$0 \longrightarrow G_2 \longrightarrow G \xrightarrow{\pi} V \longrightarrow 0$$

of Proposition 2.2 induces the following one:

$$0 \longrightarrow G_2/H = (G/H)_2 \simeq \mathbb{Z}/p\mathbb{Z} \longrightarrow G/H \xrightarrow{\pi} V \longrightarrow 0.$$

2. The curve C/H is birational to $C_f: W^p - W = f(X) = X S(X) + c X \in k[X]$, where S is an additive polynomial of degree $s \geq 1$ in F . Let Ad_f be the palindromic polynomial related to f (Proposition 2.5). Then $V \subset Z(\text{Ad}_f) \simeq (\mathbb{Z}/p\mathbb{Z})^{2s}$.
3. Let E be the wild inertia subgroup of $\text{Aut}_k(C/H)$ at ∞ . We denote by $D(E)$ its commutator subgroup of E and by $Z(E)$ its center. Then E is an extraspecial group of order p^{2s+1} and

$$0 \longrightarrow D(E) = Z(E) \simeq \mathbb{Z}/p\mathbb{Z} \longrightarrow E \xrightarrow{\pi} Z(\text{Ad}_f) \simeq (\mathbb{Z}/p\mathbb{Z})^{2s} \longrightarrow 0.$$

4. G/H is a normal subgroup in E . It follows that G_2 is equal to $D(G)$, the commutator subgroup of G , which is also equal to the Frattini subgroup of G .

Proof. The existence of the group H comes from [Suzuki 1982, Chapter 2, Theorem 1.12]. The first assertion now follows from Lemma 2.4.2. The second and third derive directly from Proposition 2.5.

We now prove part 4. By Proposition 2.5, $Z(E) = (G/H)_2 = G_2/H \subset G/H$. So, G/H is a subgroup of E containing $Z(E)$. Moreover, since $(\mathbb{Z}/p\mathbb{Z})^{2s}$ is abelian, $\pi(G/H)$ is normal in $E/Z(E)$. It follows that G/H is normal in E . We eventually show that $G_2 = D(G)$. Since G/G_2 is abelian, $D(G)$ is included in G_2 . Now assume that $D(G)$ is strictly included in G_2 . Then the first point applied to $\mathcal{G} = D(G)$ ensures the existence of a group H , normal in G , with $D(G) \subset H \subset G_2$, $[G_2 : H] = p$ and such that $(C/H, G/H)$ is a big action. Since $D(G) \subset H$, G/H is an abelian subgroup of E . As G/H is also a normal group in E , [Huppert 1967, Satz 13.7] implies $|G/H| \leq p^{s+1}$. Furthermore, by Proposition 2.5.1 (and Remark 2.6), C/H is birational to a curve $W^p - W = X S(X) + c X \in k[X]$, where S is an additive polynomial of $k[X]$ with degree p^s . It follows that $g_{C/H} = \frac{1}{2}(p - 1) p^s$. Combined with the bound on $|G/H|$, this gives $|G/H|/g_{C/H} \leq 2 p/(p-1)$, which contradicts condition (2-1) for the big action $(C/H, G/H)$. Hence $D(G) = G_2$.

It remains to prove the statement about the Frattini subgroup of G . As G is a p -group, its Frattini subgroup, $\text{Fratt}(G)$, is equal to $D(G)G^p$, where G^p means the subgroup generated by the p powers of elements of G [Leedham-Green and McKay 2002, Proposition 1.2.4]. As G/G_2 is an elementary abelian p -group, then $G^p = G_1^p \subset G_2 = D(G)$. As a consequence, $G_2 = D(G)G^p = \text{Fratt}(G)$. □

Remark 2.8. When applying Theorem 2.7 to $\mathcal{G} = G_{i_0+1}$, where i_0 is defined as in Proposition 2.5, one obtains [Lehr and Matignon 2005, Theorem 8.6(i)]. In particular, for all big actions (C, G) with $g \geq 2$, there exists a subgroup H of index p in G_2 , with H normal in G , such that $(C/H, G/H)$ is a big action with C/H

birational to $W^p - W = f(X) = X S(X) + c X \in k[X]$, where S is an additive polynomial of degree $s \geq 1$ in F . Note that, in this case, $i_0 = 1 + p^s$.

Since G_2 cannot be trivial for a big action, we gather from the last part of Theorem 2.7 the following result.

Corollary 2.9. *Let (C, G) be a big action with $g \geq 2$. Then G cannot be abelian.*

It is natural to wonder whether G_2 can be nonabelian. Although we do not know yet the answer to this question, we can mention a special case in which G_2 is always abelian, namely:

Corollary 2.10. *Let (C, G) be a big action with $g \geq 2$. If the order of G_2 divides p^3 , then G_2 is abelian.*

Proof. There is actually only one case to study, namely $|G_2| = p^3$. We denote by $Z(G_2)$ the center of G_2 . The case $|Z(G_2)| = 1$ is impossible since G_2 is a p -group. If $|Z(G_2)| = p$, then $Z(G_2)$ is cyclic. G_2 is a p -group, normal in G and included in $D(G)$ (see Theorem 2.7); hence, by [Suzuki 1986, Proposition 4.21, p. 75], G_2 is also cyclic, which contradicts the strict inclusion of $Z(G_2)$ in G_2 . If $|Z(G_2)| = p^2$, then $G_2/Z(G_2)$ is cyclic and G_2 is abelian, which leads to the same contradiction as above. This leaves only one possibility: $|Z(G_2)| = p^3$, which means that $G_2 = Z(G_2)$. □

Corollary 2.11. *Let (C, G) be a big action with $g \geq 2$. Let $A_{\infty,1}$ be the wild inertia subgroup of $\text{Aut}_k(C)$ at ∞ . Then $(C, A_{\infty,1})$ is a big action whose second lower ramification group is equal to $D(A_{\infty,1}) = D(G)$. In particular, G is equal to $A_{\infty,1}$ if and only if $|G/D(G)| = |A_{\infty,1}/D(A_{\infty,1})|$.*

Proof: As G is included in $A_{\infty,1}$, then $D(G) \subset D(A_{\infty,1})$. If the inclusion is strict, one can find a subgroup \mathcal{G} such that $G \subsetneq \mathcal{G} \subset A_{\infty,1}$, with $[\mathcal{G} : G] = p$; see [Suzuki 1982, Chapter 2, Theorem 19]. Note that $D(G) \subset D(\mathcal{G})$. We now prove that $D(G) \supset D(\mathcal{G})$. As $|G| \leq |\mathcal{G}|$, the pair (C, \mathcal{G}) is also a big action. So, by Theorem 2.7.4, $\mathcal{G}_2 = D(\mathcal{G})$. Since (C, G) is a big action, $g(C/D(G))$ vanishes by Proposition 2.2.3. It follows from Lemma 2.4.1 that $D(G) \supset \mathcal{G}_2 = D(\mathcal{G})$, hence $D(G) = D(\mathcal{G})$. The claim follows by reiterating the process. □

Remark 2.12. Let $(C, A_{\infty,1})$ be a big action as in Corollary 2.11. Then $A_{\infty,1}$ is a p -Sylow subgroup of $\text{Aut}_k(C)$. Moreover, we deduce from [Giulietti and Korchmáros 2007, Theorem 1.3] that $A_{\infty,1}$ is the unique p -Sylow subgroup of $\text{Aut}_k(C)$ except in four special cases: the hyperelliptic curves $W^{p^n} - W = X^2$ with $p > 2$, the Hermitian curves and the Deligne–Lusztig curves arising from the Suzuki groups and the Ree groups; see the equations in [Giulietti and Korchmáros 2007, Theorem 1.1].

3. Base change and big actions

Starting from a given big action (C, G) , we now display a way to produce a new one, (\tilde{C}, \tilde{G}) , with $\tilde{G}_2 \simeq G_2$ and $g_{\tilde{C}} = p^s g_C$. The chief tool is a base change associated with an additive polynomial map $\mathbb{P}_k^1 \xrightarrow{S} C/G_2 \simeq \mathbb{P}_k^1$.

Proposition 3.1. *Let (C, G) be a big action with $g \geq 2$. We denote by $L := k(C)$ the function field of the curve C , by $k(X) := L^{G_2}$ the subfield of L fixed by G_2 and by $k(T) := L^{G_1}$, with $T = \prod_{v \in V} (X - v)$. Write $X = S(Z)$, where $S(Z)$ is a separable additive polynomial of $k[Z]$ with degree p^s , $s \in \mathbb{N}$. Then:*

1. L and $k(Z)$ are linearly disjoint over $k(X)$.
2. Let \tilde{C} be the smooth projective curve over k with function field $k(\tilde{C}) := L[Z]$. Then $k(\tilde{C})/k(T)$ is a Galois extension with group $\tilde{G} \simeq G \times (\mathbb{Z}/p\mathbb{Z})^s$. Furthermore, $g_{\tilde{C}} = p^s g_C$. It follows that $|\tilde{G}|/g_{\tilde{C}} = |G|/g$. So, (\tilde{C}, \tilde{G}) is still a big action with second ramification group $\tilde{G}_2 \simeq G_2 \times \{0\} \subset G \times (\mathbb{Z}/p\mathbb{Z})^s$. This can be illustrated by the diagram

$$\begin{array}{ccc} C & \longleftarrow & \tilde{C} \\ \downarrow & & \downarrow \\ C/G_2 \simeq \mathbb{P}_k^1 & \xleftarrow{S} & \mathbb{P}_k^1 \end{array}$$

The proof requires two preliminary lemmas.

Lemma 3.2. *Let $K := k((z))$ be a formal power series field over k . Let K_1/K be a Galois extension whose group \mathcal{G} is a p -group. Let K_0/K be a cyclic extension of degree p . Assume that K_0 and K_1 are linearly disjoint over K . Put $L := K_0K_1$:*

$$\begin{array}{ccc} K_1 & \text{---} & L = K_0K_1 \\ \mathcal{G} \Big\downarrow & & \Big\downarrow \\ K & \text{---} & K_0 \end{array}$$

Suppose that the conductor of K_0/K (see Remark 2.3.2) is 2. Then L/K_1 also has conductor 2.

Proof. Consider a principal series of \mathcal{G} , that is, a sequence

$$\mathcal{G} = \mathcal{G}_0 \supseteq \mathcal{G}_1 \dots \supseteq \mathcal{G}_n = \{0\},$$

with \mathcal{G}_i normal in \mathcal{G} and $[\mathcal{G}_{i-1} : \mathcal{G}_i] = p$. One shows, by induction on i , that the conductor of each extension $K_0K_1^{\mathcal{G}_i}/K_1^{\mathcal{G}_i}$ is 2. Therefore, it is sufficient to prove the result for $\mathcal{G} \simeq \mathbb{Z}/p\mathbb{Z}$. By induction on i , it can be extended to the general case.

So, assume $\mathcal{G} \simeq \mathbb{Z}/p\mathbb{Z}$. Then $L/k((z))$ is a Galois extension with group $G \simeq (\mathbb{Z}/p\mathbb{Z})^2$. Write the ramification filtration of G in lower notation:

$$G = G_0 = \dots = G_{i_0} \supsetneq G_{i_0+1} = \dots = G_{i_1} \supsetneq G_{i_1+1} = \dots$$

First assume that $G_{i_0+1} = \{0\}$. An exercise shows that, for any subgroup H of index p in G , the extensions L/L^H (case (α)) and L^H/K (case (β)) are cyclic extensions of degree p , with conductor $i_0 + 1$. When applied to $H = \text{Gal}(L/K_0)$, case (β) gives $i_0 = 1$. Therefore, one concludes by applying case (α) to $H = \text{Gal}(L/K_1)$.

Now assume instead that $G_{i_0+1} \neq \{0\}$. As above, let H be a subgroup of index p in G . An exercise using the classical properties of ramification theory (see [Serre 1968, chapitre IV], for instance) shows that

- (a) If $H = G_{i_0+1}$, then L/L^H (resp. L^H/K) is a cyclic extension of degree p , with conductor $i_0 + i_1 + 1$ (resp. $i_0 + 1$).
- (b) If $H \neq G_{i_0+1}$, then L/L^H (resp. L^H/K) is a cyclic extension of degree p , with conductor $i_0 + 1$ (resp. $i_0 + (i_1/p) + 1$).

Apply this result to $H := \text{Gal}(L/K_0)$. Since K_0/K has conductor 2, it follows that $i_0 + 1 = 2$, so $i_0 = 1$ and $\text{Gal}(L/K_0) = G_{i_0+1}$. Therefore, $\text{Gal}(L/K_1) \neq G_{i_0+1}$ and we infer from case (b) that L/K_1 has conductor $i_0 + 1 = 2$. □

Lemma 3.3. *Let W be a finite \mathbb{F}_p -vector subspace of k . Let W_1 and W_2 be two \mathbb{F}_p -subvectors spaces of W such that $W = W_1 \oplus W_2$. Define $T := \prod_{w \in W} (Z - w)$ and $T_i := \prod_{w \in W_i} (Z - w)$, for i in $\{1, 2\}$. Then $k(T) \subset k(T_i) \subset k(Z)$. Moreover:*

1. *The extensions $k(T_1)/k(T)$ and $k(T_2)/k(T)$ are linearly disjoint over $k(T)$.*
2. *For all i in $\{1, 2\}$, $k(Z)/k(T)$ (resp. $k(Z)/k(T_i)$) is a Galois extension with group isomorphic to W (resp. W_i).*
3. *For all i in $\{1, 2\}$, $k(T_i)/k(T)$ is a Galois extension with group isomorphic to W/W_i .*

This induces the diagram

$$\begin{array}{ccc}
 k(T_1) & \xrightarrow{W_1} & k(Z) \\
 \left. \begin{array}{c} W/W_1 \\ \left| \right. \end{array} \right\} & & \left. \begin{array}{c} \left| \right. \\ W_2 \end{array} \right\} \\
 k(T) & \xrightarrow{W/W_2} & k(T_2)
 \end{array}$$

Proof. Use for example [Goss 1996, (1.8)]. □

Proof of Proposition 3.1. Statement 1 derives from Lemma 2.4.1. For 2, we put $W := S^{-1}(V)$, with V defined as in Proposition 2.2.3, and $W_1 := S^{-1}(\{0\})$; then $W_1 \simeq (\mathbb{Z}/p\mathbb{Z})^s$, since S is an additive separable polynomial of $k[Z]$ with degree p^s (see [Goss 1996, Chapter 1], for instance). Let W_2 be any \mathbb{F}_p -vector subspace of W such that $W = W_1 \oplus W_2$. Then Lemma 3.3 applied to the extension $k(Z)/k(T)$

induces the diagram

$$\begin{array}{ccc}
 L = k(C) & \xrightarrow{\quad\quad\quad} & k(\tilde{C}) \\
 G_2 \downarrow & & \downarrow \\
 L^{G_2} = k(X) = k(Z)^{W_1} & \xrightarrow{W_1} & k(Z) \\
 W/W_1 \downarrow & & \downarrow W_2 \\
 L^{G_1} = k(T) = k(Z)^W & \xrightarrow{W/W_2} & k(Z)^{W_2}
 \end{array}$$

In particular, Lemma 3.3 implies that $k(Z)^{W_1} \cap k(Z)^{W_2} = k(T)$. Since $k(C) \cap k(Z) = k(X)$ (see statement 1 of the proposition), we deduce that $k(C)$ and $k(Z)^{W_2}$ are linearly disjoint over $k(T)$. As $k(Z)^{W_2}/k(T)$ is a Galois extension with group $W/W_2 \simeq W_1 \simeq (\mathbb{Z}/p\mathbb{Z})^s$, it follows that $k(\tilde{C})/k(T)$ is a Galois extension with group $\tilde{G} \simeq \text{Gal}(k(C)/k(T)) \times \text{Gal}(k(Z)^{W_2}/k(T)) \simeq G \times (\mathbb{Z}/p\mathbb{Z})^s$.

Now, consider a flag of \mathbb{F}_p -vector subspaces of W_1 :

$$W_1 = W_1^{(1)} \supsetneq W_1^{(2)} \supsetneq \dots \supsetneq W_1^{(s+1)} = \{0\}$$

such that $[W_1^{(i-1)} : W_1^{(i)}] = p$. It induces the inclusions

$$k(Z) = k(Z)^{W_1^{(s+1)}} \supsetneq k(Z)^{W_1^{(s)}} \supsetneq \dots \supsetneq k(Z)^{W_1^{(1)}} = k(X).$$

We now prove the claim by induction on the integer $s \geq 1$, p^s being the degree of the additive polynomial S . Considering the flag above, it is sufficient to solve the case $s = 1$. Let K_1/K be the completion at ∞ of the extension $k(C)/k(X)$, whose group G_2 is a p -group and let K_0/K be the completion at ∞ of the cyclic extension of degree p and conductor 2: $k(Z)/k(X)$. To apply Lemma 3.2, we need to show that the two completions are linearly disjoint. Otherwise, $K_1 \cap K_0 = K_0$, which gives the inclusion $K \subset K_0 \subset K_1$. Consider a subgroup H of index p in G_2 such that $K_0 = K_1^H$. Let $k(X) \subset k(C)^H \subset k(C)$ be the corresponding extension of $k(X)$. Then $k(C)^H/k(X)$ is an étale p -cyclic cover of the affine line with conductor 2. It follows from the Hurwitz genus formula that the genus $g_{C/H}$ of the quotient curve C/H is 0, which contradicts Lemma 2.4.1. As a consequence, K_0 and K_1 are linearly disjoint over K and, by Lemma 3.2, the extension $k(\tilde{C})/k(C)$ has conductor 2. We deduce from the Hurwitz genus formula that $g_{\tilde{C}} = p g_C$. Finally, the last statement on \tilde{G}_2 is a consequence of Theorem 2.7.4. \square

Remark 3.4. Under the conditions of Proposition 3.1, it can happen that G is a p -Sylow subgroup of $\text{Aut}_k(C)$ without \tilde{G} being a p -Sylow subgroup of $\text{Aut}_k(\tilde{C})$.

Indeed, take $C : W^p - W = X^{1+p}$ and $S(Z) = Z^p - Z$. Then \tilde{C} is parametrized by $\tilde{W}^p - \tilde{W} = (Z^p - Z)(Z^{p^2} - Z^p) = -Z^2 + 2Z^{1+p} - Z^{1+p^2} \pmod{\wp(k[Z])}$. We denote by $A_{\infty,1}(C)$ (resp. $A_{\infty,1}(\tilde{C})$) the wild inertia subgroup of $\text{Aut}_k(C)$ (resp. $\text{Aut}_k(\tilde{C})$)

at $X = \infty$ (resp. $Z = \infty$). Note that $A_{\infty,1}(C)$ (resp. $A_{\infty,1}(\tilde{C})$) is a p -Sylow subgroup of $\text{Aut}_k(C)$ (resp. $\text{Aut}_k(\tilde{C})$). Take $G := A_{\infty,1}(C)$. From Proposition 2.5, we deduce that $|\tilde{G}| = p |G| = p |A_{\infty,1}(C)| = p^4$, whereas $|A_{\infty,1}(\tilde{C})| = p^5$.

4. A new step towards a classification of big actions

If big actions are defined through the value taken by the quotient $|G|/g$, it turns out that the key criterion to classify them is the value of another quotient, $|G|/g^2$. Indeed, the quotient $|G|/g^2$ has, to some extent, a sieve effect among big actions. If (C, G) is a big action, we first deduce from [Nakajima 1987, Theorem 1] that $|G|/g^2 \leq 4 p/(p - 1)^2$. In what follows, we pursue the work of Lehr and Matignon who describe big actions for the two highest possible values of this quotient, namely $|G|/g^2 = 4 p/(p - 1)^2$ and $|G|/g^2 = 4/(p - 1)^2$; see [Lehr and Matignon 2005, Theorem 8.6]. More precisely, we investigate the big actions (C, G) that satisfy

$$M := \frac{4}{(p^2 - 1)^2} \leq \frac{|G|}{g^2}. \tag{4-1}$$

The choice of the lower bound M can be explained as follows: as shown in the proof of [Lehr and Matignon 2005, Theorem 8.6], a lower bound M on the quotient $|G|/g^2$ produces an upper bound on the order of the second ramification group, namely

$$|G_2| \leq \frac{4}{M} \frac{|G_2/G_{i_0+1}|^2}{(|G_2/G_{i_0+1}| - 1)^2}, \tag{4-2}$$

where i_0 is defined as in Proposition 2.5. Therefore, we have to choose M small enough to obtain a wide range of possibilities for the quotient, but meanwhile large enough to get serious restrictions on the order of G_2 . The optimal bound seems to be $M := 4/(p^2 - 1)^2$, insofar as, for such a choice of M , the upper bound on G_2 implies that its order divides p^3 , and then that G_2 is abelian (Corollary 2.10).

Proposition 4.1. *Let (C, G) be a big action with $g \geq 2$ satisfying condition (4-1). Then the order of G_2 divides p^3 . It follows that G_2 is abelian.*

Proof. Put $p^m := |G_2/G_{i_0+1}|$, with $m \geq 1$, and

$$Q_m := \frac{4}{M} \frac{|G_2/G_{i_0+1}|}{(|G_2/G_{i_0+1}| - 1)^2} = \frac{4}{M} \frac{p^m}{(p^m - 1)^2}.$$

Then inequality (4-2) becomes $1 < |G_2| = p^m |G_{i_0+1}| \leq p^m Q_m$, which gives $1 \leq |G_{i_0+1}| \leq Q_m$. Since $(Q_m)_{m \geq 1}$ is a decreasing sequence with $Q_4 < 1$, we conclude that $m \in \{1, 2, 3\}$.

If $m = 3$, then $1 \leq |G_{i_0+1}| \leq Q_3 < p$. So $|G_{i_0+1}| = 1$ and $|G_2| = p^3$. If $m = 2$, then $1 \leq |G_{i_0+1}| \leq Q_2 = p^2$. So $|G_2| = p^2 |G_{i_0+1}|$, with $|G_{i_0+1}| \in \{1, p, p^2\}$. This leaves only one case to exclude, namely $|G_{i_0+1}| = p^2$. In this case, $|G_2| = p^4$ and

formula (2–2) yields a lower bound on the genus, namely $2g \geq (i_0 - 1)(p^4 - 1)$. Let s be the integer defined in Remark 2.8. Then $i_0 = 1 + p^s$. Besides, by Theorem 2.7, $V \subset (\mathbb{Z}/p\mathbb{Z})^{2s}$. Consequently, $|G| = |G_2||V| \leq p^{4+2s}$ and

$$\frac{|G|}{g^2} \leq \frac{4p^{4+2s}}{p^{2s}(p^4 - 1)^2} = \frac{4}{(p^2 - 1)^2} \frac{p^4}{(p^2 + 1)^2} < \frac{4}{(p^2 - 1)^2},$$

which contradicts inequality (4–1).

If $m = 1$, then $1 \leq |G_{i_0+1}| \leq Q_1$, with

$$Q_1 := p(p + 1)^2 < \begin{cases} p^4 & \text{if } p \geq 3, \\ p^5 & \text{if } p = 2. \end{cases}$$

Because G_{i_0+1} is a p -group, we get $1 \leq |G_{i_0+1}| \leq p^3$ if $p \geq 3$, or $1 \leq |G_{i_0+1}| \leq p^4$ if $p = 2$. Since $|G_2| = p|G_{i_0+1}|$, there are two cases to exclude: $|G_{i_0+1}| = p^{3+\epsilon}$, with $\epsilon = 0$ if $p \geq 3$ and $\epsilon \in \{0, 1\}$ if $p = 2$. Then $|G_2| = p^{4+\epsilon}$. If $\epsilon = 0$, we are in the same situation as in the previous case. If $\epsilon = 1$, (2–2) yields $2g \geq (i_0 - 1)(p^5 - 1)$. Since this case only occurs for $p = 2$, we eventually get the inequality

$$\frac{|G|}{g^2} \leq \frac{4p^{5+2s}}{p^{2s}(p^5 - 1)^2} = \frac{128}{961} < \frac{4}{9} = \frac{4}{(p^2 - 1)^2},$$

which contradicts condition (4–1). Therefore, the order of G_2 divides p^3 . Then we conclude from Corollary 2.10 that G_2 is abelian. \square

But we can even prove better:

Proposition 4.2. *Let (C, G) be a big action with $g \geq 2$ satisfying condition (4–1). Then G_2 is abelian with exponent p .*

Proof. By Proposition 4.1, G_2 is abelian, with order dividing p^3 . As a consequence, if G_2 has exponent greater than p , either G_2 is cyclic with order p^2 or p^3 , or G_2 is isomorphic to $\mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. We begin with a lemma excluding the second case. Note that one can find big actions (C, G) with G_2 abelian of exponent p^2 . Nevertheless, it requires the p -rank of G_2 to be large enough (see Section 6).

Lemma 4.3. *Let (C, G) be a big action with $g \geq 2$ satisfying condition (4–1). Then G_2 cannot be isomorphic to $\mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.*

Proof. Assume $G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$. The lower ramification filtration of G has one of the following forms:

1. $G = G_1 \supseteq G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+1} \simeq \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+i_1+1} = \{0\}$.
2. $G = G_1 \supseteq G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset G_{i_0+i_1+1} = \{0\}$.
3. $G = G_1 \supseteq G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset G_{i_0+i_1+1} \simeq \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+i_1+i_2+1} = \{0\}$.

$$4. G = G_1 \supsetneq G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+1} \simeq \mathbb{Z}/p^2\mathbb{Z} \supset G_{i_0+i_1+1} \simeq \mathbb{Z}/p\mathbb{Z} \supset G_{i_0+i_1+i_2+1} = \{0\}.$$

We now focus on the ramification filtration of G_2 , temporary denoted by H for convenience. For all $i \geq 0$, the lower ramification groups of H are $H_i = H \cap G_i$.

In case (i), the lower ramification of $H =: H_0$ reads

$$H_0 = \dots = H_{i_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+1} = \dots = H_{i_0+i_1} \simeq \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+i_1+1} = \{0\}.$$

Consider the upper ramification groups $H^{\nu_0} = H^{\varphi(i_0)} = H_{i_0}$ and $H^{\nu_1} = H^{\varphi(i_0+i_1)} = H_{i_0+i_1}$, where φ denotes the Herbrand function [Serre 1968, IV.3]. Then the ramification filtration in upper notation reads

$$H^0 = \dots = H^{\nu_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H^{\nu_0+1} = \dots = H^{\nu_1} \simeq \mathbb{Z}/p\mathbb{Z} \supset H^{\nu_1+1} = \{0\}.$$

Since H is abelian, it follows from the Hasse–Arf theorem (loc. cit.) that ν_0 and ν_1 are integers. Consequently, the equality

$$\varphi(m) + 1 = \frac{1}{|H_0|} \sum_{i=0}^m |H_i| \quad \text{for all } m \in \mathbb{N}$$

gives $\nu_0 = i_0$ and $\nu_1 = i_0 + i_1/p^2$. By [Marshall 1971, Theorem 6] we have $H^{\nu_0} \supsetneq H^{p\nu_0} \supset (H^{\nu_0})^p$ with $(H^{\nu_0})^p = H^p = G_2^p \simeq \mathbb{Z}/p\mathbb{Z}$. Thus, $H^{p\nu_0} \supset H^{\nu_1}$, which implies $p\nu_0 \leq \nu_1$ and $i_1 \geq p^2(p-1)i_0$. Then the Hurwitz genus formula applied to $C \rightarrow C/H \simeq \mathbb{P}_k^1$ yields a lower bound for the genus:

$$2g = (i_0 - 1)(|H| - 1) + i_1(|H_{i_0+1}| - 1) \geq (p - 1)(i_0 + 1)(p^3 + p + 1).$$

Let s be the integer defined in Remark 2.8. Then $i_0 = 1 + p^s$. Moreover, by Theorem 2.7, $|G| = |G_2| |V| \leq p^{3+2s}$. It follows that

$$\frac{|G|}{g^2} \leq \frac{4}{(p^2 - 1)^2} \frac{p^3(p + 1)^2}{(p^3 + p + 1)^2}.$$

Since $p^3(p + 1)^2 / (p^3 + p + 1)^2 < 1$ for $p \geq 2$, this contradicts condition (4–1).

In case (ii), the lower ramification filtration of $H = H_0$ reads

$$H_0 = \dots = H_{i_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+1} = \dots = H_{i_0+i_1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset H_{i_0+i_1+1} = \{0\}.$$

Keeping the notation of case (i), the upper ramification filtration is

$$H^0 = \dots = H^{\nu_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H^{\nu_0+1} = \dots = H^{\nu_1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset H^{\nu_1+1} = \{0\},$$

with $\nu_0 = \varphi(i_0) = i_0$ and $\nu_1 = \varphi(i_0 + i_1) = i_0 + i_1/p$. Once again, $H^{p\nu_0} \supset (H^{\nu_0})^p \simeq \mathbb{Z}/p\mathbb{Z}$ implies $H^{p\nu_0} \supset H^{\nu_1}$, which involves $p\nu_0 \leq \nu_1$ and $i_1 \geq i_0 p(p-1)$. Then

the Hurwitz genus formula yields

$$\begin{aligned} 2g &= (i_0 - 1)(|H| - 1) + i_1(|H_{i_0+1}| - 1) \\ &\geq (p - 1)p^s(p^3 + p^2 + 1) \geq (p - 1)p^s(p^3 + p + 1). \end{aligned}$$

Thus we get the same lower bound on the genus as in the preceding case, hence the same contradiction.

In case (iii), the lower ramification filtration of H becomes

$$\begin{aligned} H_{i_0} &\simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+1} = \dots \\ &= H_{i_0+i_1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset H_{i_0+i_1+1} = \dots = H_{i_0+i_1+i_2} \simeq \mathbb{Z}/p\mathbb{Z} \supset \{0\}. \end{aligned}$$

Keeping the same notation as above and introducing $H^{v_2} = H^{\varphi(i_0+i_1+i_2)} = H_{i_0+i_1+i_2}$, the upper ramification filtration is

$$\begin{aligned} H^{v_0} &\simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H^{v_0+1} = \dots \\ &= H^{v_1} \simeq (\mathbb{Z}/p\mathbb{Z})^2 \supset H^{v_1+1} = \dots = H^{v_2} \simeq \mathbb{Z}/p\mathbb{Z} \supset H^{v_2+1} = \{0\}, \end{aligned}$$

with $v_0 = \varphi(i_0) = i_0$, $v_1 = \varphi(i_0 + i_1) = i_0 + i_1/p$ and $v_2 = \varphi(i_0 + i_1 + i_2) = i_0 + i_1/p + i_2/p^2$. Since $H^{p v_0} \supset (H^{v_0})^p \simeq \mathbb{Z}/p\mathbb{Z}$, we obtain $H^{p v_0} \supset H^{v_2}$. Then $p v_0 \leq v_2$, which involves $p^2(p-1)i_0 \leq i_1 p + i_2$. With such inequalities, the Hurwitz genus formula gives a new lower bound for the genus, namely

$$\begin{aligned} 2g &= (i_0 - 1)(|H| - 1) + i_1(|H_{i_0+1}| - 1) + i_2(|H_{i_0+i_1+1}| - 1) \\ &\geq (p - 1)(p^s(p^2 + p + 1) + (p^s + 1)(p - 1)p^2). \end{aligned}$$

From the inequalities $2g \geq (p-1)(p^{3+s} + p^{1+s} + p^s + p^3 - p^2) \geq (p-1)p^s(p^3 + p)$, we infer that

$$\frac{|G|}{g^2} \leq \frac{4}{(p^2 - 1)^2} \frac{p^{2s+3}(p+1)^2}{p^{2s}(p^3 + p)^2} = \frac{4}{(p^2 - 1)^2} \frac{p(p+1)^2}{(p^2 + 1)^2}.$$

Since $p(p+1)^2/(p^2+1)^2 < 1$ for $p \geq 2$, this contradicts condition (4-1).

In case (iv), the lower ramification filtration of H , namely

$$\begin{aligned} H_{i_0} &\simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+1} = \dots \\ &= H_{i_0+i_1} \simeq (\mathbb{Z}/p^2\mathbb{Z}) \supset H_{i_0+i_1+1} = \dots = H_{i_0+i_1+i_2} \simeq \mathbb{Z}/p\mathbb{Z} \supset \{0\} \end{aligned}$$

induces the upper ramification filtration

$$\begin{aligned} H^{v_0} &\simeq \mathbb{Z}/p^2\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \supset H^{v_0+1} = \dots \\ &= H^{v_1} \simeq (\mathbb{Z}/p^2\mathbb{Z}) \supset H^{v_1+1} = \dots = H^{v_2} \simeq \mathbb{Z}/p\mathbb{Z} \supset H^{v_2+1} = \{0\}. \end{aligned}$$

This is almost the same situation as in case (iii), except that H_{i_0+1} is isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$ instead of $(\mathbb{Z}/p\mathbb{Z})^2$. But, since the only thing that plays a part in the proof is the order of H_{i_0+1} , which is the same in both cases, namely p^2 , we conclude with the same arguments as in case (iii). \square

Remark 4.4. The preceding method, based on the analysis of the ramification filtration of G_2 , fails to exclude the case $G_2 \simeq \mathbb{Z}/p^2\mathbb{Z}$ for a big action satisfying (4–1). Indeed, if $H := G_2 \simeq \mathbb{Z}/p^2\mathbb{Z}$, the lower ramification filtration of H ,

$$H_0 = \dots = H_{i_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \supset H_{i_0+1} = \dots = H_{i_0+i_1} \simeq \mathbb{Z}/p\mathbb{Z} \supset H_{i_0+i_1+1} = \{0\}$$

induces the upper ramification filtration

$$H^0 = \dots = H^{v_0} \simeq \mathbb{Z}/p^2\mathbb{Z} \supset H^{v_0+1} = \dots = H^{v_1} \simeq \mathbb{Z}/p\mathbb{Z} \supset H^{v_1+1} = \{0\}.$$

with $v_0 = \varphi(i_0) = i_0$ and $v_1 = \varphi(i_0+i_1) = i_0+i_1/p$. Since $H^{pv_0} \supset (H^{v_0})^p \simeq \mathbb{Z}/p\mathbb{Z}$, we obtain $p v_0 \leq v_1$, hence $i_1 \geq (p-1) p i_0$. Let s be the integer defined in Remark 2.8. The Hurwitz genus formula yields

$$\begin{aligned} 2g &= (i_0 - 1)(|H| - 1) + i_1(|H_{i_0+1}| - 1) \\ &\geq (p - 1)(p^s(p^2 + 1) + p^2 - p) \geq (p - 1)p^s(p^2 + 1). \end{aligned}$$

If we denote by v the dimension of the \mathbb{F}_p -vector space V , we ultimately get

$$\frac{|G|}{g^2} \leq \frac{4}{(p^2 - 1)^2} \frac{p^{2+v}(p + 1)^2}{p^{2s}(p^2 + 1)^2}.$$

In this case, condition (4–1) requires $p^{1+(v/2)-s}(p + 1) > p^2$. Since $v/2 \leq s$, this implies $p + 1 > p^{1+s-v/2} \geq p$, hence $v/2 = s$. This means that $V = Z(\text{Ad}_f)$, where f is the function defined in Remark 2.8 and Ad_f its palindromic polynomial as defined in Proposition 2.5. Therefore, one does not obtain yet any contradiction.

Accordingly, to exclude the cyclic cases $G_2 \simeq \mathbb{Z}/p^2\mathbb{Z}$ and $G_2 \simeq \mathbb{Z}/p^3\mathbb{Z}$ and thus complete the proof of Proposition 4.2, we need to shift from a ramification point of view on G_2 to the embedding problem $G_2 \subsetneq G_1$. This enables us to prove the more general result on big actions formulated later.

5. Big actions with a cyclic second ramification group G_2

The aim of this section is to prove that there does not exist any big action whose second ramification group G_2 is cyclic, except for the trivial case $G_2 \simeq \mathbb{Z}/p\mathbb{Z}$. For Witt vectors and Artin–Schreier–Witt theory, our main reference is [Bourbaki 1983, chapitre IX].

Theorem 5.1. *Let (C, G) be a big action. If $G_2 \simeq (\mathbb{Z}/p^n\mathbb{Z})$, then $n = 1$.*

Proof. Let (C, G) be a big action with $G_2 \simeq \mathbb{Z}/p^n\mathbb{Z}$. We proceed in steps.

1. *We first prove that we can assume $n = 2$.* Indeed, for $n > 2$, $\mathcal{H} := G_2^{p^{n-2}}$ is a normal subgroup in G , strictly included in G_2 . So Lemma 2.4.2 asserts that the pair $(C/\mathcal{H}, G/\mathcal{H})$ is a big action. Besides, the second lower ramification group of G/\mathcal{H} is isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$.

Notation and preparatory remarks. We denote by $L := k(C)$ the function field of C and by $k(X) := L^{G_2}$ the subfield of L fixed by G_2 . Following Artin–Schreier–Witt theory [Bourbaki 1983, chapitre IX, exercice 19], we define the $W_2(\mathbb{F}_p)$ -module

$$\tilde{A} := \frac{\wp(W_2(L)) \cap W_2(k(X))}{\wp(W_2(k(X)))},$$

where $W_2(L)$ denotes the ring of Witt vectors of length 2 with coordinates in L . The inclusion $k[X] \subset k(X)$ induces an injection

$$A := \frac{\wp(W_2(L)) \cap W_2(k[X])}{\wp(W_2(k[X]))} \hookrightarrow \tilde{A}.$$

Since L/L^{G_2} is étale outside $X = \infty$, it follows from [Milne 1980, III, 4.12] that we can identify A with \tilde{A} . Consider the Artin–Schreier–Witt pairing

$$\begin{aligned} G_2 \times A &\longrightarrow W_2(\mathbb{F}_p), \\ (g, \overline{\wp x}) &\longmapsto [g, \overline{\wp x}] := gx - x, \end{aligned}$$

where $g \in G_2 \subset \text{Aut}_k(L)$, $x \in L$ such that $\wp x \in k[X]$ and $\overline{\wp x}$ denotes the class of $\wp x \bmod \wp(k[X])$. This pairing is nondegenerate, which proves that, as a group, A is dual to G_2 .

As a \mathbb{Z} -module, A is generated by $(f_0(X), g_0(X))$ in $W_2(k[X])$, and then, $L = k(X, W_0, V_0)$ with $\wp(W_0, V_0) = (f_0(X), g_0(X))$. An exercise left to the reader shows that one can choose $f_0(X)$ and $g_0(X)$ reduced mod $\wp(k[X])$ (see the definition of a reduced polynomial on page 890). We denote by m_0 the degree of f_0 and by n_0 that of g_0 . Note that they are prime to p . The p -cyclic cover $L^{G_2^p}/L^{G_2}$ is parametrized by $W_0^p - W_0 = f_0(X)$. We deduce from Proposition 2.5 that $f_0(X) = XS(X) + cX$, where S is an additive polynomial with degree $s \geq 1$ in F . After an homothety on X , we can assume S to be monic. Furthermore, note that $s \geq 2$. Indeed, if $s = 1$, the inequalities $|G| \leq p^{2+2s} \leq p^4$ and $2g \geq (p-1)(p^s(p^2+1) + p^2 - p) = (p-1)(p^3 + p^2)$ of Remark 4.4 imply

$$\frac{|G|}{g} \leq \frac{2p}{p-1} \frac{p^3}{p^3 + p^2} < \frac{2p}{p-1},$$

which contradicts (2–1).

2. *The embedding problem.* Let V be the \mathbb{F}_p -vector space defined in Proposition 2.2.3. For any $y \in V$, the class of $(f_0(X + y), g_0(X + y))$ in A induces a new

generating system of A , which means that

$$\mathbb{Z}(f_0(X), g_0(X)) = \mathbb{Z}(f_0(X + y), g_0(X + y)) \pmod{\wp(W_2(k[X]))}. \quad (5-1)$$

Since A is isomorphic to $\mathbb{Z}/p^2\mathbb{Z}$, (5-1) ensures the existence of an integer $n(y)$ such that

$$(f_0(X + y), g_0(X + y)) = n(y) (f_0(X), g_0(X)) \pmod{\wp(W_2(k[X]))}, \quad (5-2)$$

where $n(y) := a_0(y) + b_0(y)p$ for integers $a_0(y)$ and $b_0(y)$ such that $0 < a_0(y) < p$ and $0 \leq b_0(y) < p$. We calculate $n(y) (f_0(X), g_0(X)) = a_0(y) (f_0(X), g_0(X)) + b_0(y)p (f_0(X), g_0(X))$. On the one hand, we have

$$a_0(y) (f_0(X), g_0(X)) = (a_0(y)f_0(X), a_0(y)g_0(X) + c(a_0(y))f_0(X)),$$

where $c(a_0(y))$ is given by the recursion

$$c(1) = 1, \quad c(i + 1) = c(i) + \frac{1}{p} (1 + i^p - (1 + i)^p) \pmod{p} \quad \text{for all } i \in \mathbb{N}.$$

On the other hand,

$$b_0(y)p (f_0(X), g_0(X)) = b_0(y)(0, f_0(X)^p) = (0, b_0(y)f_0(X)) \pmod{\wp(W_2(k[X]))}.$$

Consequently, (5-2) becomes

$$(f_0(X + y), g_0(X + y)) = (a_0(y)f_0(X), a_0(y)g_0(X) + \ell_0(y)f_0(X)) \pmod{\wp(W_2(k[X]))}, \quad (5-3)$$

where $\ell_0(y) := c(a_0(y)) + b_0(y)$. We notice that $a_0(y) = 1 \pmod{p}$ for all y in V . Indeed, the equality of the first coordinate of Witt vectors in (5-3) implies that $f_0(X + y) = a_0(y) f_0(X) \pmod{\wp(k[X])}$. Thus, by induction, $f_0(X + py) = a_0(y)^p f_0(X) \pmod{\wp(k[X])}$. Since V is an elementary abelian p -group we get $f_0(X + py) = f_0(X)$, which entails $a_0(y)^p = 1 \pmod{p}$ and $a_0(y) = 1 \pmod{p}$. So (5-3) becomes

$$(f_0(X + y), g_0(X + y)) = (f_0(X), g_0(X) + \ell_0(y)f_0(X)) + (P^p(X), Q^p(X)) - (P(X), Q(X)), \quad (5-4)$$

with $P(X)$ and $Q(X)$ polynomials of $k[X]$. In order to circumvent the problem related to the special formula giving the opposite of Witt vectors for $p = 2$, we would rather write (5-4) as

$$(f_0(X + y), g_0(X + y)) + (P(X), Q(X)) = (f_0(X), g_0(X) + \ell_0(y) f_0(X)) + (P(X)^p, Q(X)^p). \quad (5-5)$$

The first coordinate of (5-5) reads

$$f_0(X + y) + P(X) = f_0(X) + P(X)^p. \quad (5-6)$$

On the second coordinate of (5–5), the addition law in the ring of Witt vectors gives in $k[X]$ the equality

$$g_0(X + y) + Q(X) + \psi(f_0(X + y), P(X)) = g_0(X) + \ell_0(y) f_0(X) + Q(X)^p + \psi(f_0(X), P(X)^p), \quad (5-7)$$

where ψ is defined by

$$\begin{aligned} \psi(a, b) &:= \frac{1}{p} (a^p + b^p - (a + b)^p) = -\frac{1}{p} \sum_{i=1}^{p-1} \binom{p}{i} a^i b^{p-i} \\ &= \sum_{i=1}^{p-1} \frac{(-1)^i}{i} a^i b^{p-i} \pmod{p}. \end{aligned}$$

As a consequence, (5–7) gives

$$\Delta_y(g_0) := g_0(X + y) - g_0(X) = \ell_0(y) f_0(X) + \delta \pmod{\wp(k[X])}, \quad (5-8)$$

with

$$\begin{aligned} \delta &:= \psi(f_0(X), P(X)^p) - \psi(f_0(X + y), P(X)) \\ &= \sum_{i=1}^{p-1} \frac{(-1)^i}{i} \{f_0(X)^i P(X)^{p(p-i)} - f_0(X + y)^i P(X)^{p-i}\}. \end{aligned}$$

Lemma 5.2. *With the notation defined above, δ is equal to*

$$\delta = \sum_{i=1}^{p-1} \frac{(-1)^i}{i} y^{p-i} X^{i+p^s+1} + \text{lower-degree terms in } X. \quad (5-9)$$

Proof. We search for the monomials in δ that have degree at least $p^{s+1} + 1$ in X . We first focus on $f_0(X)^i P(X)^{p(p-i)}$. We can infer from equality (5–6) that $P(X)$ has degree p^{s-1} and that its leading coefficient is $y^{1/p}$. By [Lehr and Matignon 2005, proof of Proposition 8-1], $P(X) - P(0)$ is an additive polynomial. So we can write $P(X) = y^{1/p} X^{p^{s-1}} + P_1(X)$, where $P_1(X)$ is a polynomial of $k[X]$ of degree at most p^{s-2} . Then, for all i in $\{1, \dots, p - 1\}$, $f_0(X)^i P(X)^{p(p-i)} = f_0(X)^i (y X^{p^s} + P_1(X)^p)^{p-i} = f_0(X)^i \left(\sum_{j=0}^{p-i} \binom{p-i}{j} y^j X^{jp^s} P_1(X)^{p(p-i-j)}\right)$. Since $f_0(X)$ has degree $1 + p^s$, this gives in δ a monomial of degree at most $i(1 + p^s) + j p^s + p(p-i-j) p^{s-2} = p^s + (i+j)(p-1) p^{s-1} + i$. If $j \leq p-i-1$, this degree is at most $p^s + (p-1)^2 p^{s-1} + i = (p-1) p^s + p^{s-1} + i$, which is strictly less than $p^{s+1} + 1$, for $s \geq 2$ and $1 \leq i \leq p-1$. As a consequence, monomials of degree at least $p^{s+1} + 1$ can only occur when the index j is equal to $p-i$, namely in $f_0(X)^i y^{p-i} X^{p^s(p-i)}$. As $f_0(X) = X S(X) + c X$, where S is a monic additive polynomial of degree s in F , f_0 reads $f_0(X) = X^{1+p^s} + P_2(X)$ where $P_2(X)$ is a polynomial in $k[X]$ with degree at most $1 + p^{s-1}$. Then, for all i in $\{1, \dots, p - 1\}$, we have $f_0(X)^i y^{p-i} X^{p^s(p-i)} = y^{p-i} X^{p^s(p-i)} \left(\sum_{k=0}^i \binom{i}{k} X^{(1+p^s)k} P_2(X)^{i-k}\right)$. Accordingly, we get a monomial of degree at most to $p^s(p-i) + k(1+p^s) + (i-k)(1+p^{s-1})$, a number we can rewrite

as $p^s(p-i) + i(1+p^{s-1}) + k(p^s - p^{s-1})$. When $0 \leq k \leq i-1$, the maximal degree obtained in this way is $i + p^{s-1} - p^s + p^{s+1}$ which is strictly lower than $p^{s+1} + 1$. Therefore, for all i in $\{1, \dots, p-1\}$, the only contribution to take into account is $k = i$, which produces in δ the sum

$$\sum_{i=1}^{p-1} \frac{(-1)^i}{i} y^{p-i} X^{i+p^{s+1}}.$$

We now search for monomials with degree greater or equal to $p^{s+1} + 1$ in the second part of δ , namely $f_0(X+y)^i P(X)^{p-i}$. This has degree at most $i(1+p^s) + (p-i)p^{s-1} = ip^s + (p-i)p^{s-1} + i$, which is strictly less than $p^{s+1} + 1$, for $s \geq 2$ and $1 \leq i \leq p-1$. Therefore, $f_0(X+y)^i P(X)^{p-i}$ does not give any monomial in δ with degree greater or equal to $p^{s+1} + 1$. Thus, we get the expected formula. \square

3. We next show that $g_0(X)$ cannot be of the form $X \Sigma(X) + \gamma X$, with $\Sigma \in k\{F\}$ and $\gamma \in k$. Otherwise, the left-hand side of (5-8) reads $\Delta_y(g_0) := g_0(X+y) - g_0(X) = X \Sigma(y) + y \Sigma(X) + y \Sigma(y) + \gamma y$, which only gives a linear contribution in X after reduction mod $\wp(k[X])$. By Lemma 5.2, $\deg f_0 = 1 + p^s < \deg \delta = p^{s+1} + p - 1$, which involves that the degree of the right-hand side of (5-8) is $p - 1 + p^{s+1} > 1$, hence a contradiction.

Therefore, we can define an integer $a \leq n_0 = \deg g_0$ such that X^a is the monomial of $g_0(X)$ with highest degree which is not of the form $1 + p^n$, with $n \in \mathbb{N}$. Note that since g_0 is reduced mod $\wp(k[X])$, $a \not\equiv 0 \pmod p$. We also notice that the monomials in $g_0(X)$ with degree greater than a are of the form X^{1+p^n} ; hence, as explained above, they only give linear monomials in $\Delta_y(g_0) \pmod{\wp(k[X])}$. Therefore, after reduction mod $\wp(k[X])$, the degree of the left-hand side of (5-8) is at most $a - 1$. Since the degree of the right-hand side is $p^{s+1} + p - 1$, it follows that

$$a - 1 \geq p^{s+1} + p - 1. \tag{5-10}$$

4. We show that p divides $a - 1$. Assume that p does not divide $a - 1$. In this case, the monomial X^{a-1} is reduced mod $\wp(k[X])$. Since the monomials of $g_0(X)$ with degree greater than a only give a linear contribution in $\Delta_y(g_0) \pmod{\wp(k[X])}$, (5-8) reads as follows, for all y in V :

$$\begin{aligned} c_a(g_0) a y X^{a-1} + \text{lower-degree terms} \\ = -y X^{p^{s+1}+p-1} + \text{lower-degree terms} \pmod{\wp(k[X])}, \end{aligned}$$

where $c_a(g_0) \neq 0$ denotes the coefficient of X^a in g_0 . If $a - 1 > p^{s+1} + p - 1$, the coefficient $c_a(g_0) a y = 0$, for all y in V . Since $a \not\equiv 0 \pmod p$, it leads to $V = \{0\}$, so $G_1 = G_2$, which is impossible for a big action (see Proposition 2.2.1). We gather from (5-10) that $a - 1 = p^{s+1} + p - 1$, which contradicts $a \not\equiv 0 \pmod p$.

Thus p divides $a - 1$. So, we can write $a = 1 + \lambda p^t$, with $t > 0$, λ prime to p and $\lambda \geq 2$ because of the definition of a . We also define $j_0 := a - p^t = 1 + (\lambda - 1) p^t$. Note that $p j_0 > a$. Indeed,

$$p j_0 \leq a \Leftrightarrow p(1 + (\lambda - 1) p^t) \leq 1 + \lambda p^t \Leftrightarrow \lambda \leq \frac{1 - p + p^{t+1}}{p^t(p - 1)} = \frac{-1}{p^t} + \frac{p}{p - 1} < \frac{p}{p - 1} \leq 2,$$

which is impossible since $\lambda \geq 2$.

5. We determine the coefficient of X^{j_0} in the left hand-side of (5–8). Since p does not divide j_0 , the monomial X^{j_0} is reduced mod $\wp(k[X])$. On the left-hand side of (5–8), namely $\Delta_y(g_0) \bmod \wp(k[X])$, the monomial X^{j_0} comes from monomials of $g_0(X)$ of the form X^b , with b in $\{j_0 + 1, \dots, a\}$. As a matter of fact, the monomials of $g_0(X)$ with degree greater than a only give a linear contribution mod $\wp(k[X])$, whereas $j_0 = 1 + (\lambda - 1) p^t > 1$. For all $b \in \{j_0 + 1, \dots, a\}$, the monomial X^b of $g_0(X)$ generates $\binom{b}{j_0} y^{b-j_0} X^{j_0}$ in $\Delta_y(g_0)$. Since $p j_0 > a \geq b$ (see above), these monomials X^b do not produce any $X^{j_0 p^n}$, with $n \geq 1$, which would also give X^{j_0} after reduction mod $\wp(k[X])$. It follows that the coefficient of X^{j_0} in the left-hand side of (5–8) is $T(y)$ with $T(Y) := \sum_{b=j_0+1}^a c_b(g_0) \binom{b}{j_0} Y^{b-j_0}$, where $c_b(g_0)$ denotes the coefficient of X^b in $g_0(X)$. As the coefficient of Y^{a-j_0} in $T(Y)$ is $c_a(g_0) \binom{a}{j_0} = c_a(g_0) \binom{1+\lambda p^t}{1+(\lambda-1)p^t} \equiv c_a(g_0) \lambda \not\equiv 0 \pmod p$, the polynomial $T(Y)$ has degree $a - j_0 = p^t$.

6. We identify with the coefficient of X^{j_0} in the right-hand side of (5–8) and obtain a contradiction. We first assume that the monomial X^{j_0} does not occur in the right-hand side of (5–8). Then $T(y) = 0$ for all y in V , which means that V is included in the set of roots of T . Thus, $|V| \leq p^t$. To compute the genus g , put $M_0 := m_0$ and $M_1 := \max\{p m_0, n_0\}$. Then, by [Garuti 2002], the Hurwitz genus formula applied to $C \rightarrow C/G_2 \simeq \mathbb{P}_k^1$ yields

$$2(g - 1) = 2|G_2|(g_{C/G_2} - 1) + d = -2p^2 + d,$$

with $d := (p - 1)(M_0 + 1) + p(p - 1)(M_1 + 1)$. From $p m_0 = p(p^s + 1) = p^{s+1} + p$ and $p^{s+1} + p - 1 < n_0$, we infer $M_1 = n_0$. Moreover, since $n_0 \geq a = 1 + \lambda p^t \geq 1 + 2p^t > 2p^t$, we obtain the lower bound $2g = (p - 1)p(n_0 - 1 + p^{s-1}) \geq 2p^{t+1}(p - 1)$ for the genus. Since $|G| = |G_2||V| \leq p^{2+t}$, this entails

$$\frac{|G|}{g} \leq \frac{2p}{p-1} \frac{p^{1+t}}{2p^{1+t}} = \frac{1}{2} \frac{2p}{p-1},$$

which contradicts (2–1).

As a consequence, the monomial X^{j_0} appears in the right-hand side of (5–8), which implies that $j_0 \leq p^{s+1} + p - 1$. Using (5–10), we get $j_0 = 1 + (\lambda - 1) p^t \leq$

$p^{s+1} + p - 1 < a = 1 + \lambda p^t$. This yields

$$\lambda - 1 \leq p^{s+1-t} + \frac{p-2}{p^t} < \lambda. \tag{5-11}$$

If $s + 1 - t \leq -1$, since $t \geq 1$, (5-11) gives $\lambda - 1 \leq 1/p + (p - 2)/p < 1$, which contradicts $\lambda \geq 2$. It follows that $s + 1 - t \geq 0$. Then (5-11), combined with the inequalities $0 \leq (p - 2)/p^t < 1$, leads to $\lambda - 1 = p^{s+1-t}$. We gather that $j_0 = 1 + (\lambda - 1) p^t = 1 + p^{s+1} > \deg f_0 = 1 + p^s$. Therefore, in the right-hand side of (5-8), the monomial $X^{j_0} = X^{1+p^{s+1}}$ only occurs in δ . By Lemma 5.2, the coefficient of $X^{j_0} = X^{1+p^{s+1}}$ in δ is $-y^{p-1}$. By equating the coefficient of X^{j_0} in each side of (5-8), we get $T(y) = -y^{p-1}$, for all y in V . Put $\tilde{T}(Y) := T(Y) + Y^{p-1}$. Since $\deg T = p^t > p - 1$, the polynomial \tilde{T} has still degree p^t and satisfies $\tilde{T}(y) = 0$ for all y in V . Once again, it leads to $|V| \leq p^t$, which contradicts (2-1) as above. \square

Therefore, when (C, G) is a big action, $G_2 \simeq (\mathbb{Z}/p^n\mathbb{Z})$ implies $n = 1$. More generally, if G_2 is abelian of exponent p^n , with $n \geq 2$, there exists a subgroup H of index p in G_2^p , with H normal in G , such that the pair $(C/H, G/H)$ is a big action with $(G/H)_2 = G_2/H \simeq \mathbb{Z}/p^2\mathbb{Z} \times (\mathbb{Z}/p\mathbb{Z})^t$, with $t \in \mathbb{N}^*$. A natural question is to search for a lower bound on the p -rank t depending on the genus g of the curve. As seen in the proof of Theorem 5.1, the difficulty lies in the embedding problem, i.e. in finding an extension which is stable under the translations by V . In the next section, we exhibit big actions with G_2 abelian of exponent at least p^2 . In particular, we construct big actions (C, G) with $G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times (\mathbb{Z}/p\mathbb{Z})^t$ where $t = O(\log_p g)$.

6. Examples of big actions with G_2 abelian of exponent greater than p

In characteristic 0, an analogue of big actions is given by the actions of a finite group G on a compact Riemann surface C with genus $g_C \geq 2$ such that $|G| = 84(g_C - 1)$. Such a curve C is called a *Hurwitz curve* and such a group G a *Hurwitz group* [Conder 1990]. In particular, the lowest genus Hurwitz curves are the Klein’s quartic with $G \simeq \text{PSL}_2(\mathbb{F}_7)$ (cf. [Elkies 1999a]) and the Fricke–Macbeath curve with genus 7 and $G \simeq \text{PSL}_2(\mathbb{F}_8)$ [Macbeath 1965].

Let C be a Hurwitz curve with genus g_c . Let $n \geq 2$ be an integer and let C_n be the maximal unramified Galois cover whose group is abelian, with exponent n . The Galois group of the cover C_n/C is isomorphic to $(\mathbb{Z}/n\mathbb{Z})^{2g_c}$. We infer from the uniqueness of C_n that the \mathbb{C} -automorphisms of C have n^{2g_c} prolongations to C_n . Therefore, $g_{C_n} - 1 = n^{2g}(g_C - 1)$. Consequently, C_n is still a Hurwitz curve; see [Macbeath 1961].

Now let (C, G) be a big action. Then $C \rightarrow C/G$ is an étale cover of the affine line whose group is a p -group. From the Deuring–Shafarevich formula (see [Bouw

2000], for example), it follows that the Hasse–Witt invariant of C is zero. This means that there are no nontrivial connected étale Galois covers of C with group a p -group. Therefore, if we want to generalize the method mentioned above to produce Galois covers of C corresponding to big actions, it is necessary to introduce ramification. A means to do so is to consider ray class fields of function fields, as studied by K. Lauter [1999] and R. Auer [1999]. Since the cover $C \rightarrow C/G_2$ is an étale cover of the affine line $\text{Spec } k[X]$ totally ramified at ∞ , we focus on the special case of ray class fields of the rational function field $\mathbb{F}_q(X)$, where $q = p^e$ [Auer 1999, III.8]. Such ray class fields allow us to produce families of big actions (C, G) (where C is defined over $k = \mathbb{F}_p^{\text{alg}}$) with specific conditions imposed on ramification and endowed with an abelian G_2 of exponent as large as we want.

Definition 6.1 [Auer 1999, Part II]. Let $K := \mathbb{F}_q(X)$ be the rational function field, with $q = p^e$ and $e \in \mathbb{N}^*$. Let S be the set of all finite rational places, namely $\{(X - y), y \in \mathbb{F}_q\}$. Let $m \geq 0$ be an integer. Fix K^{alg} an algebraic closure of K in which all extensions of K are assumed to lie. We define $K_S^m \subset K^{\text{alg}}$ as the largest abelian extension L/K with conductor $\leq m\infty$, such that every place in S splits completely in L .

Remarks 6.2. 1. We define the splitting set of any finite Galois extension L/K , denoted by $S(L)$, as the set consisting of the places of K that split completely in L . If K_S^m/K is the extension defined in Definition 6.1, then $S \subset S(K_S^m)$.

2. In what follows, we only consider finite Galois extensions L/K that are unramified outside $X = \infty$ and (totally) ramified at $X = \infty$. Therefore, the support of the conductor of L/K is reduced to the place ∞ . So, we systematically confuse the conductor $m\infty$ with its degree m .
3. We could more generally define K_S^m for S a nonempty subset of the finite rational places, i.e. $S := \{(X - y), y \in V \subset \mathbb{F}_q\}$. However, to get big actions, it is necessary to consider the case where V is a subgroup of \mathbb{F}_q . In what follows, we focus on the case $V = \mathbb{F}_q$, as announced in Definition 6.1.

Remarks 6.3. We keep the notation of Definition 6.1.

1. The existence of the extension K_S^m/K is based on global class field theory; see [Auer 1999, Part II].
2. K_S^m/K is a finite abelian extension whose full constant field is \mathbb{F}_q .
3. The reason why Lauter and Auer are interested in such ray class fields is that they provide for examples of global function fields with many rational places, or what amounts to the same, of algebraic curves with many rational points. Indeed, let $C(m)/\mathbb{F}_q$ be the nonsingular projective curve with function field K_S^m . If we denote by $N_m := |C(m)(\mathbb{F}_q)|$ the number of \mathbb{F}_q -rational points on the

curve $C(m)$, then $N_m = 1 + q [K_S^m : K]$. The main difficulty lies in computing $[K_S^m : K]$. We first wonder when K_S^m coincide with K . Here are partial answers.

4. Let $q = p^e$, with $e \in \mathbb{N}$. If e is even, put $r := \sqrt{q}$ and if e is odd, put $r := \sqrt{qp}$. Then, for all i in $\{0, \dots, r + 1\}$, $K_S^i = K = \mathbb{F}_q(X)$; see [Auer 1999, III, Lemma 8.7 and formula (13)]. Note that the previous estimate $N_m = 1 + q [K_S^m : K]$, combined with the Hasse–Weil bound (see [Stichtenoth 1993, V.2.3], for instance), furnishes another proof of $K_S^i = K$ when $i < 1 + r$.
5. More generally, Lauter displays a method to compute the degree of the extension K_S^m/K via a formula giving the order of its Galois group $G_S(m)$ [Lauter 1999, Theorem 1]. Lauter’s proof starts from the following presentation of $G_S(m)$:

$$G_S(m) \simeq \frac{1 + Z \mathbb{F}_q \llbracket Z \rrbracket}{\langle 1 + Z^m \mathbb{F}_q \llbracket Z \rrbracket, 1 - yZ, y \in \mathbb{F}_q \rangle},$$

where $Z = X^{-1}$, which indicates that $G_S(m)$ is an abelian finite p -group. Then she transforms the multiplicative structure of the group into an additive group of generalized Witt vectors. In particular, she deduces from this theorem the smallest conductor m such that $G_S(m)$ has exponent strictly greater than p (see next proposition).

Proposition 6.4 [Lauter 1999, Proposition 4]. *We keep the notation defined above. If $q = p^e$, the smallest conductor m for which the group $G_S(m)$ is not of exponent p is $m_2 := p^{\lceil e/2 \rceil + 1} + p + 1$, where $\lceil \cdot \rceil$ is the ceiling function.*

We now emphasize the link with big actions. Let F be a function field with full constant field \mathbb{F}_q . Let C/\mathbb{F}_q be the smooth projective curve whose function field is F and $C^{\text{alg}} := C \times_{\mathbb{F}_q} k$ with $k = \mathbb{F}_p^{\text{alg}}$. If G is a finite p -subgroup of $\text{Aut}_{\mathbb{F}_q} C$, then G can be identified with a subgroup of $\text{Aut}_k C^{\text{alg}}$. In this case, (C^{alg}, G) is a big action if and only if $g_{C^{\text{alg}}} = g_C > 0$ and $|G|/g_C > 2p/(p-1)$. For convenience, in the sequel, we shall say that (C, G) is a big action if (C^{alg}, G) is a big action.

In what follows, we consider the curve $C(m)/\mathbb{F}_q$ whose function field is K_S^m and, starting from this, we construct a p -group $G(m)$ acting on $C(m)$ by extending the translations $X \rightarrow X + y$, with $y \in \mathbb{F}_q$. In particular, we obtain an upper bound for the genus of $C(m)$, which allows us to circumvent the problem related to the computation of the degree $[K_S^m : K]$ when checking whether $(C(m), G(m))$ is a big action.

Proposition 6.5. *We keep the notation defined above.*

1. *Let $C(m)/\mathbb{F}_q$ be the nonsingular projective curve with function field K_S^m . Then the group of translations $X \rightarrow X + y$, $y \in \mathbb{F}_q$, extends to a p -group of \mathbb{F}_q -automorphisms of $C(m)$, say $G(m)$, with the exact sequence*

$$0 \longrightarrow G_S(m) \longrightarrow G(m) \longrightarrow \mathbb{F}_q \longrightarrow 0.$$

2. Let L be an intermediate field of K_S^m/K . Assume $L = (K_S^m)^H$, i.e. the extension L/K is Galois with group: $G_S(m)/H$. For all $i \geq 0$, we define L^i as the i -th upper ramification field of L , i.e. the subfield of L fixed by the i -th upper ramification group of $G_S(m)/H$ at ∞ : $G_S^i(m)H/H$, where $G_S^i(m)$ denotes the i -th upper ramification group of $G_S(m)$ at ∞ . Then, for all $i \geq 0$,

$$L^i = L \cap K_S^i.$$

In particular, when $L = K_S^m$ and $i \leq m$, $L^i = K_S^i$, i.e. $G_S^i(m) = \text{Gal}(K_S^m/K_S^i)$.

3. Let L be an intermediate field of K_S^m/K . Define $n := \min\{i \in \mathbb{N}, L \subset K_S^i\}$. The genus of the extension L/K is given by the formula

$$g_L = 1 + [L : K](-1 + n/2) - \frac{1}{2} \sum_{j=0}^{n-1} [L \cap K_S^j : K],$$

where the sum is empty for $n = 0$. In particular, g_L vanishes if and only if $L \subset K_S^0$; in all other cases, $g_L < [L : K](-1 + n/2)$.

4. If $m \geq r + 2$, then $|G(m)|/g_{K_S^m} > q/(-1 + \frac{1}{2}m)$. It follows that if $q/(-1 + \frac{1}{2}m) \geq 2p/(p-1)$, the pair $(C(m), G(m))$ is a big action. In this case, the second lower ramification group $G_2(m)$ of $G(m)$ is equal to $G_S(m)$. In particular, with m_2 as in Proposition 6.4, if $p > 2$ and $e \geq 4$ or $p = 2$ and $e \geq 6$, the pair $(C(m_2), G(m_2))$ is a big action whose second ramification group $G_S(m_2)$ is abelian of exponent p^2 .

Proof. 1. The set S is globally invariant under the translations $X \mapsto X + y$, $y \in \mathbb{F}_q$. That is the same for ∞ , so the translations by \mathbb{F}_q do not change the conditions imposed on ramification. As a consequence, owing to the maximality and the unicity of K_S^m , they can be extended to \mathbb{F}_q -automorphisms of K_S^m . This proves the first assertion.

2. This follows directly from [Auer 1999, II, Theorem 5.8].

3. The genus formula is obtained by combining the preceding results, the Hurwitz genus formula and the discriminant formula [Auer 1999, I, 3.7]. Now assume that $n = 0$. Then $L \subset K_S^0 = \mathbb{F}_q(X)$ and $g_L = 0$. Conversely, assume $g_L = 0$. If $n \neq 0$, Remark 6.3.4 implies that $n \geq r + 2 \geq 3$. Using the preceding formula and Remark 6.3.4, $g_L = 0$ reads

$$2 + (n - 2)[L : K] = \sum_{j=0}^{n-1} [K_S^j \cap L : K] = 2 + \sum_{j=2}^{n-1} [K_S^j \cap L : K] \leq 2 + (n - 2)[L : K].$$

It follows that, for all j in $\{2, \dots, n - 1\}$, $K_S^j \cap L = L$. In particular, $L \subset K_S^2 = K_S^0$, hence a contradiction. Finally, since $n > 0$ implies $n \geq 3$ and since $K = K_S^0 = K_S^1$,

one notices that

$$g_L = [L : K](-1 + n/(2)) - \frac{1}{2} \sum_{j=2}^{n-1} [L \cap K_S^j : K] < [L : K](-1 + n/(2)).$$

Assume that $m \geq r + 2$. We gather from Remark 6.3.4 that $n := \min\{i \in \mathbb{N}, K_S^m \subset K_S^i\} \geq r + 2 \geq 3$. It follows from the third point that

$$g_{K_S^m} < [K_S^m : K](-1 + n/(2)) \leq [K_S^m : K](-1 + m/(2)).$$

As $|G(m)| = q[K_S^m : K]$, we deduce the expected inequality. In particular, when $q/(-1 + \frac{1}{2}m) > 2p/(p-1)$, the pair $(C(m), G(m))$ is a big action. It remains to show that, in this case, $G_2(m)$ is equal to $G_S(m)$. Lemma 2.4.2 first proves that $G_S(m) \supset G_2(m)$. Let $L := (K_S^m)^{G_2(m)}$ be the subfield of L fixed by $G_2(m)$ and define $n := \min\{i \in \mathbb{N}, L \subset K_S^i\}$. Assume $G_S(m) \supsetneq G_2(m)$. Then $L \supsetneq (K_S^m)^{G_S(m)} = K$. We infer from Remark 6.3.4 that $n \geq r + 2$, which proves, using the previous point, that $g_L > 0$. But, since $(C(m), G(m))$ is a big action, $C/G_2(m) \simeq \mathbb{P}_k^1$, so $g_L = 0$, hence a contradiction. We eventually explain the last statement. By Proposition 6.5.2, $G_S^{m_2-1}(m_2) = \text{Gal}(K_S^{m_2}/K_S^{m_2-1})$, which induces the exact sequence

$$0 \longrightarrow G_S^{m_2-1}(m_2) \longrightarrow G_S(m_2) \longrightarrow G_S(m_2 - 1) \longrightarrow 0.$$

We infer from Proposition 6.4 that $G_S(m_2 - 1)$ has exponent p whereas the exponent of $G_S(m_2)$ is at least p^2 . It follows that $G_S^{m_2-1}(m_2)$ cannot be trivial. Since $G_S^{m_2}(m_2) = \{0\}$ (use Proposition 6.5.2), we deduce from the elementary properties of the ramification groups that $G_S^{m_2-1}(m_2)$ is p -elementary abelian. Therefore, $G_S(m_2)$ has exponent smaller than p^2 and the claim follows. \square

Remark 6.6. Let N_m be the number of \mathbb{F}_q -rational points on the curve $C(m)$ as defined in Remark 6.3.3. Then $N_m = 1 + q |G_S(m)| = 1 + |G(m)|$. This highlights the equivalence of the two ratios: $|G(m)|/g_{C(m)}$ and $N_m/g_{C(m)}$. In particular, this equivalence emphasizes the link between the problem of big actions and the search of algebraic curves with many rational points.

As seen in Remark 6.3.4, $K_S^i = K$ for all i in $\{0, \dots, r + 1\}$, where $r = \sqrt{q}$ or \sqrt{qp} according to whether q is a square or not. The following extensions K_S^m , for $m \geq r + 2$, are partially parametrized, at least for the first ones, in [Auer 1999, Proposition 8.9]. The table on the next page gives a complete description of the extensions K_S^m for m varying from 0 to $m_2 = p^{\lceil e/2 \rceil + 1} + p + 1$, in the special case $p = 5$ and $e = 4$. This involves $q = p^e = 625$, $s = e/2 = 2$, $r = p^s = 25$ and $m_2 = 131$. The table below should suggest the general method to parametrize such extensions.

conductor m	$[K_S^m:K]$	new equations
$0 \leq m \leq r + 1 = 26$	1	
$27 \leq m \leq 2r + 1 = 51$	5^2	$W_0^r + W_0 = X^{1+r}$
$m = 2r + 2 = 52$	5^6	$W_1^q - W_1 = X^{2r} (X^q - X)$
$53 \leq m \leq 3r + 1 = 76$	5^8	$W_2^r + W_2 = X^{2(1+r)}$
$m = 3r + 2 = 77$	5^{12}	$W_3^q - W_3 = X^{3r} (X^q - X)$
$m = 3r + 3 = 78$	5^{16}	$W_4^q - W_4 = X^{3r} (X^{2q} - X^2)$
$79 \leq m \leq 4r + 1 = 101$	5^{18}	$W_5^r + W_5 = X^{3(1+r)}$
$m = 4r + 2 = 102$	5^{22}	$W_6^q - W_6 = X^{4r} (X^q - X)$
$m = 4r + 3 = 103$	5^{26}	$W_7^q - W_7 = X^{4r} (X^{2q} - X^2)$
$m = 4r + 4 = 104$	5^{30}	$W_8^q - W_8 = X^{4r} (X^{3q} - X^3)$
$105 \leq m \leq 5r + 1 = 126$	5^{32}	$W_9^r + W_9 = X^{4(1+r)}$
$m = 5r + 2 = 127$	5^{36}	$W_{10}^q - W_{10} = X^{5r} (X^q - X)$
$m = 5r + 3 = 128$	5^{40}	$W_{11}^q - W_{11} = X^{5r} (X^{2q} - X^2)$
$m = 5r + 4 = 129$	5^{44}	$W_{12}^q - W_{12} = X^{5r} (X^{3q} - X^3)$
$m = 5r + 5 = 130$	5^{48}	$W_{13}^q - W_{13} = X^{5r} (X^{4q} - X^4)$
$m = m_2 = 131$	5^{50}	$[W_0, W_{14}]^r + [W_0, W_{14}] = [X^{1+r}, 0]$

In this case,

$$\frac{|G(m_2)|}{g_{K_S^{m_2}}} \simeq 9, 6929 \dots \tag{6-1}$$

Comments on the construction of the table. For all i in $\{0, \dots, 14\}$, put $L_i := K(W_0, \dots, W_i)$.

1. We first prove that the splitting set of each extension $K(W_i)/K$ (see Remark 6.2.1) contains S . Indeed, fix y in \mathbb{F}_q and call P_y the corresponding place in S : $(X - y)$. We have to distinguish three cases. By [Stichtenoth 1993, Proposition VI. 4.1], P_y completely splits in the extension $K(W)/K$, where $W^r + W = X^{u(1+r)}$, with $1 \leq u \leq 4$, if the polynomial $T^r + T - y^{u(1+r)}$ has a root in K , which is true since $y^{u(1+r)} = (F^s + I) (\frac{1}{2} y^{u(1+r)})$. Likewise, P_y completely splits in the extension $K(W)/K$, where $W^q - W = X^{ur} (X^{vq} - X^v)$, with $1 \leq v < u \leq 5$, since $y^{vq} - y^v = 0$. Finally, P_y completely splits in the extension $K(W, \tilde{W})/K$, where $[W, \tilde{W}]^r + [W, \tilde{W}] = [X^{1+r}, 0]$, since

$$[y^{1+r}, 0] = (F^s + I) \left[\frac{1}{2} y^{1+r}, -\frac{2^p - 2}{4p} y^{(1+r)p} \right].$$

To conclude, we remark that $L_i = L_{i-1} K(W_i)$ for all i in $\{1, \dots, 14\}$. Then $S(L_i) = S(L_{i-1}) \cap S(K(W_i))$, by [Auer 1999, Corollary 3.2.b], which allows us to conclude, by induction on i , that the splitting set of each L_i contains S .

2. We now compute the conductor $m(K(W_i))$ of each extension $K(W_i)/K$. As above, we must distinguish three kinds of extensions. The extension $K(W)/K$, where $W^r + W = X^{u(1+r)}$, with $1 \leq u \leq 4$, has conductor $ur + u + 1$ [Auer 1999, Proposition 8.9a]. The extension $K(W)/K$, where now $W^q - W = X^{ur}(X^{vq} - X^v)$, with $1 \leq v < u \leq 5$, has conductor $ur + v + 1$ [Auer 1999, Proposition 8.9b]. Finally, the conductor of the extension $K(W, \tilde{W})/K$, where $[W, \tilde{W}]^r + [W, \tilde{W}] = [X^{1+r}, 0]$ is given by the formula $1 + \max\{p(1+r), -\infty\} = 1 + p + p^{s+1} = m_2$ [Garuti 2002, Theorem 1.1]. As a conclusion, since $m(L_i) = \max\{m(L_{i-1}), m(K(W_i))\}$ [Auer 1999, Corollary 3.2b], an induction on i allows us to obtain the expected conductor for L_i .
3. We obtain from 1 and 2 the inclusions $K(W_0) \subset K_S^{27}$, $K(W_0, W_1) \subset K_S^{52}, \dots$, $K(W_0, \dots, W_{14}) \subset K_S^{m_2}$. Equality is finally obtained by calculating the degree of each extension K_S^m/K via [Lauter 1999, Theorem 1] or [Auer 1999, (13), pp. 54–55]. □

We deduce from the foregoing an example of big actions with G_2 abelian of exponent p^2 , with a small p -rank. More precisely, we construct a subextension of $K_S^{m_2}$ with the commutative diagram

$$\begin{array}{ccccccc}
 0 & \longrightarrow & G_S(m_2) & \longrightarrow & G(m_2) & \longrightarrow & \mathbb{F}_q \longrightarrow 0 \\
 & & \varphi \downarrow & & \downarrow & & \parallel \\
 0 & \longrightarrow & H & \longrightarrow & G & \longrightarrow & \mathbb{F}_q \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \\
 & & 0 & & 0 & &
 \end{array}$$

such that the pair $(C(m_2)/\text{Ker}(\varphi), G)$ is a big action where $G_2 \simeq \mathbb{Z}/p^2\mathbb{Z} \times (\mathbb{Z}/p\mathbb{Z})^t$ with $t = O(\log_p g)$, g being the genus of the curve $C(m_2)/\text{Ker}(\varphi)$. Contrary to the previous case where the stability under the translations by \mathbb{F}_q was ensured by the maximality of $K_S^{m_2}$, the difficulty now lies in producing a system of equations defining a subextension of $K_S^{m_2}$ which remains globally invariant through the action of the group of translations $X \rightarrow X + y$, $y \in \mathbb{F}_q$. Write $q = p^e$. We have to distinguish the case e even and e odd.

Proposition 6.7. *Assume that $p > 2$. We keep the notation defined above. In particular, $K = \mathbb{F}_q(X)$ with $q = p^e$. Assume that $e = 2s$, with $s \geq 1$, and put $r := p^s$. Define*

$$\begin{aligned}
 f_0(X) &:= a X^{1+r} \text{ with } a \neq 0, a \in \Gamma := \{\gamma \in \mathbb{F}_q, \gamma^r + \gamma = 0\}, \\
 f_i(X) &:= X^{ir/p} (X^q - X) = X^{ip^{s-1}} (X^q - X) \text{ for all } i \in \{1, \dots, p-1\}.
 \end{aligned}$$

Let $L := K(W_i)_{0 \leq i \leq p}$ be the extension of K parametrized by the Artin–Schreier–Witt equations $W_0^p - W_0 = f_0(X)$, $W_i^q - W_i = f_i(X)$ for all $i \in \{1, \dots, p-1\}$, and $[W_0, W_p]^p - [W_0, W_p] = [f_0(X), 0]$.

For all i in $\{0, 1, \dots, p - 1\}$, put $L_i := K(W_0, \dots, W_i)$. Let C_L/\mathbb{F}_q be the nonsingular projective curve with function field L .

1. L is an abelian extension of K and every place in S completely splits in L . Moreover, $L_0 \subset K_S^{r+2}$, $L_i \subset K_S^{p^{s+1}+i+1}$ for all $i \in \{1, \dots, p - 1\}$, and $L \subset K_S^{m_2}$, where $m_2 = p^{s+1} + p + 1$ is the integer defined in Proposition 6.4. (See table on the next page.)

2. L/K has degree $[L : K] = p^{2+(p-1)e}$, and its Galois group G_L satisfies

$$G_L \simeq \mathbb{Z}/p^2\mathbb{Z} \times (\mathbb{Z}/p\mathbb{Z})^t \quad \text{with } t = (p - 1)e.$$

3. The extension L/K is stable under the translations $X \mapsto X + y$, with $y \in \mathbb{F}_q$. Therefore, the translations by \mathbb{F}_q extend to form a p -group of \mathbb{F}_q -automorphisms of L , say G , with the exact sequence

$$0 \longrightarrow G_L \longrightarrow G \longrightarrow \mathbb{F}_q \longrightarrow 0.$$

4. Let g_L be the genus of the extension L/K . Then

$$g_L = \frac{1}{2} \left(p^{2+2s(p-1)} (p^{s+1} + p - 1) - p^s (p^2 - p + 1) - p^{2s+1} \sum_{i=0}^{p-2} q^i \right).$$

In particular, when e grows large, $g_L \sim \frac{1}{2} p^{(2p-1)e/2+3}$ and $t = O(\log_p g_L)$.

5. For $s \geq 2$, (C_L, G) is a big action with $G_2 = G_L$. (Note that, for $p = 5$ and $e = 4$, one gets $|G|/g_L \simeq 9, 7049 \dots$, which is slightly bigger than the quotient obtained for the whole extension $K_S^{m_2}$ in (6-1).)

Proof. 1. Fix y in \mathbb{F}_q and call $P_y := (X - y)$, the corresponding place in S . As $f_i(y) = 0$ for all i in $\{1, \dots, p - 1\}$, the place P_y completely splits in each extension $K(W_i)$ with $W_i^q - W_i = f_i(X)$. Therefore, to prove that P_y completely splits in L , it is sufficient to show that $[f_0(y), 0] \in \wp(W_2(\mathbb{F}_q))$. By [Bourbaki 1983, chapitre IX, exercice 18], this is equivalent to show that $\text{Tr}([f_0(y), 0]) = 0$, where Tr means the trace map from $W_2(\mathbb{F}_q)$ to $W_2(\mathbb{F}_p)$. We first notice that, when y is in \mathbb{F}_q , $\gamma := f_0(y) = a y^{1+r}$ lies in Γ . It follows that

$$\begin{aligned} \text{Tr}([\gamma, 0]) &= \sum_{i=0}^{2s-1} F^i [\gamma, 0] = \sum_{i=0}^{s-1} [\gamma^{p^i}, 0] + \sum_{i=0}^{s-1} [\gamma^{r p^i}, 0] \\ &= \sum_{i=0}^{s-1} [\gamma^{p^i}, 0] + \sum_{i=0}^{s-1} [-\gamma^{p^i}, 0]. \end{aligned}$$

As $p > 2$, $[-\gamma^{p^i}, 0] = -[\gamma^{p^i}, 0]$ and $\text{Tr}([\gamma, 0]) = 0$. To establish the expected inclusions, it remains to compute the conductor of each extension L_i . First of all, [Auer 1999, I, exercice 3.3] together with [Stichtenoth 1993, Proposition III.7.10] shows that the conductor of L_0 is $r + 2$. Thus, $L_0 \subset K_S^{r+2}$. Moreover, since $f_i(X) =$

$X^{i+p^{s+1}} - X^{1+i p^{s-1}} \pmod{\wp(\mathbb{F}_q[X])}$, we infer from [Auer 1999, I, Exercise 3.3 and Corollary 3.2] that the conductor of L_i is $1 + i + p^{s+1}$. So, $L_i \subset K_S^{1+i+p^{s+1}}$. To complete the proof, it remains to show that L has conductor m_2 , which follows from [Garuti 2002] (see comments above).

The equations, conductor and degree of each extension L_i are as follows:

ext'n	conductor m	$[L_i:K]$	new equations
K	$0 \leq m \leq r+1 = p^s+1$	1	
L_0	$r+2 \leq m \leq p^{s+1}+1 = m_2-p$	p	$W_0^p - W_0 = f_0(X)$
L_1	$m = p^{s+1}+2 = m_2-(p-1)$	p^{1+e}	$W_1^q - W_1 = f_1(X)$
L_2	$m = p^{s+1}+3 = m_2-(p-2)$	p^{1+2e}	$W_2^q - W_2 = f_2(X)$
\dots	\dots	\dots	\dots
L_i	$m = p^{s+1}+i+1 = m_2-(p-i)$	p^{1+ie}	$W_i^q - W_i = f_i(X)$
\dots	\dots	\dots	\dots
L_{p-1}	$m = p^{s+1}+p = m_2-1$	$p^{1+(p-1)e}$	$W_{p-1}^q - W_{p-1} = f_{p-1}(X)$
L	$m = p^{s+1}+p+1 = m_2$	$p^{2+(p-1)e}$	$[W_0, W_p]^p - [W_0, W_p] = [f_0(X), 0]$

2. See preceding table.

3. Fix y in \mathbb{F}_q . Consider σ in $G(m_2)$ (defined as in Proposition 6.5) such that $\sigma(X) = X + y$.

(a) We prove that $\sigma(W_0) \in L_0$. Indeed, as $y \in \mathbb{F}_q$ and $a \in \Gamma = \{\gamma \in \mathbb{F}_q, \gamma^r + \gamma = 0\}$,

$$\begin{aligned} \wp(\sigma(W_0) - W_0) &= \sigma(\wp(W_0)) - \wp(W_0) = f_0(X + y) - f_0(X) \\ &= a y X^r + a y^r X + f_0(y) = -a^r y^{r^2} X^r + a y^r X + f_0(y) \\ &= \wp(P_y(X)) + f_0(y), \end{aligned}$$

where $P_y(X) := (I + F + F^2 + \dots + F^{s-1})(-a y^r X)$. Since $f_0(y) \in \wp(\mathbb{F}_q)$ (see proof of part 1), it follows that $\wp(P_y(X)) + f_0(y)$ belongs to $\wp(\mathbb{F}_q[X])$. Therefore, $\sigma(W_0) \in L_0 = \mathbb{F}_q(X, W_0)$.

(b) We now prove that $\sigma(W_i) \in L_i$ for all i in $\{1, \dots, p-1\}$. Indeed,

$$\begin{aligned} (F^e - \text{id})(\sigma(W_i) - W_i) &= \sigma(W_i^q - W_i) - (W_i^q - W_i) = f_i(X + y) - f_i(X) \\ &= (X + y)^{i p^{s-1}} (X^q - X) - X^{i p^{s-1}} (X^q - X) \\ &= (X^{p^{s-1}} + y^{p^{s-1}})^i (X^q - X) - X^{i p^{s-1}} (X^q - X) \\ &= \sum_{j=1}^{i-1} \binom{i}{j} y^{(i-j)p^{s-1}} f_j(X) \pmod{(F^e - \text{id})(\mathbb{F}_q[X])}, \end{aligned}$$

where the sum is empty for $i = 1$. It turn, the right-hand side equals

$$(F^e - \text{id}) \left(\sum_{j=1}^{i-1} \binom{i}{j} y^{(i-j)p^{s-i}} W_j \right) \pmod{(F^e - \text{id})(\mathbb{F}_q[X])}.$$

It follows that $\sigma(W_i) \in L_i = \mathbb{F}_q(X, W_0, W_1, \dots, W_i)$.

(c) We next show, using Remark 6.3.4, that $\sigma(W_p) \in L$. To this end, set

$$\Delta := \wp(\sigma[W_0, W_p] - [W_0, W_p]),$$

so

$$\Delta = \sigma(\wp([W_0, W_p])) - \wp([W_0, W_p]) = [f_0(X + y), 0] - [f_0(X), 0].$$

We know from the proof of part 1 that $[f_0(y), 0]$ lies in $\wp(W_2(\mathbb{F}_q))$. Then

$$\Delta = [f_0(X + y), 0] - [f_0(X), 0] - [f_0(y), 0] - [P_y(X), 0] + [P_y(X), 0]^p \pmod{\wp(W_2(\mathbb{F}_q[X]))},$$

with y in \mathbb{F}_q and P_y defined as above. Let $W(\mathbb{F}_q)$ be the ring of Witt vectors with coefficients in \mathbb{F}_q . Then, for any $y \in \mathbb{F}_q$, we denote by \tilde{y} the Witt vector $\tilde{y} := (y, 0, 0, \dots) \in W(k)$. For any $P(X) := \sum_{i=0}^s a_i X^i \in \mathbb{F}_q[X]$, set $\tilde{P}(X) := \sum_{i=0}^s \tilde{a}_i X^i \in W(\mathbb{F}_q)[X]$. Addition in the ring of Witt vectors yields

$$\Delta = [0, A] \pmod{\wp(W_2(\mathbb{F}_q[X]))},$$

where A is the reduction modulo p $W_2(\mathbb{F}_q)[X]$ of

$$\frac{1}{p} \{ \tilde{f}_0(X + \tilde{y})^p - \tilde{f}_0(X)^p - \tilde{f}_0(\tilde{y})^p + \tilde{P}_y(X)^p - \tilde{P}_y(X)^{p^2} - (\tilde{f}_0(X + \tilde{y}) - \tilde{f}_0(X) - \tilde{f}_0(\tilde{y}) - \tilde{P}_y(X) + \tilde{P}_y(X)^p)^p \}.$$

Since $\tilde{f}_0(X + \tilde{y}) - \tilde{f}_0(X) - \tilde{f}_0(\tilde{y}) + \tilde{P}_y(X) - \tilde{P}_y(X)^p = 0 \pmod{p W(\mathbb{F}_q)[X]}$, we get

$$A = \frac{1}{p} \{ \tilde{f}_0(X + \tilde{y})^p - \tilde{f}_0(X)^p - \tilde{f}_0(\tilde{y})^p + \tilde{P}_y(X)^p - \tilde{P}_y(X)^{p^2} \} \pmod{p W(\mathbb{F}_q)[X]}.$$

We observe that, modulo $\pmod{p^2 W(\mathbb{F}_q)[X]}$,

$$\begin{aligned} \tilde{f}_0(X + \tilde{y})^p &= \tilde{a}^p (X + \tilde{y})^p (X + \tilde{y})^{p^{s+1}} = \tilde{a}^p (X + \tilde{y})^p (X^{p^s} + \tilde{y}^{p^s})^p \\ &= \tilde{a}^p \sum_{i=0}^p \sum_{j=0}^p \binom{p}{i} \binom{p}{j} X^{j+ip^s} \tilde{y}^{p-j+p^s(p-i)} \end{aligned}$$

Since $\binom{p}{i} \binom{p}{j} = 0 \pmod{p^2}$ when $0 < i < p$ and $0 < j < p$, one obtains

$$\tilde{f}_0(X + \tilde{y})^p - \tilde{f}_0(X)^p - \tilde{f}_0(\tilde{y})^p = \tilde{a}^p \sum_{(i,j) \in I} \binom{p}{i} \binom{p}{j} X^{j+ip^s} \tilde{y}^{p-j+p^s(p-i)} \pmod{p^2 W(\mathbb{F}_q)[X]},$$

where I is the set of $(i, j) \in \{0, 1, \dots, p\}^2 \setminus \{(0, 0), (p, p)\}$ such that

$$ij = 0 \pmod{p}, (i, j).$$

Again modulo $\pmod{p^2 W(\mathbb{F}_q)[X]}$, we have

$$\begin{aligned} \tilde{P}_y(X)^p - \tilde{P}_y(X)^{p^2} &= \left(\sum_{i=0}^{s-1} (-\tilde{a} \tilde{y}^r X)^{p^i} \right)^p - \left(\sum_{i=0}^{s-1} (-\tilde{a} \tilde{y}^r X)^{p^i} \right)^{p^2} \\ &= \left(\sum_{i=0}^{s-1} (-\tilde{a} \tilde{y}^r X)^{p^i} \right)^p - \left(\sum_{i=0}^{s-1} (-\tilde{a} \tilde{y}^r X)^{p^{i+1}} \right)^p \\ &= -\tilde{a}^p \tilde{y}^{rp} X^p + \tilde{a}^{rp} \tilde{y}^{r^2 p} X^{pr} + p \tilde{T}_y(X), \end{aligned}$$

with $\tilde{T}_y(X) \in W(\mathbb{F}_q)[X]$. Since $y \in \mathbb{F}_q$ and $a \in \Gamma$, we get

$$\tilde{P}_y(X)^p - \tilde{P}_y(X)^{p^2} = -\tilde{a}^p \tilde{y}^{rp} X^p - \tilde{a}^{rp} \tilde{y}^{r^2 p} X^{pr} + p \tilde{T}_y(X) \pmod{p^2 W(\mathbb{F}_q)[X]}.$$

As a consequence,

$$A = \tilde{a}^p \sum_{(i,j) \in I_1} \frac{1}{p} \binom{p}{i} \binom{p}{j} X^{j+ip^s} \tilde{y}^{p-j+p^s(p-i)} + \tilde{T}_y(X) \pmod{p \wp(\mathbb{F}_q[X])},$$

where $I_1 = I \setminus \{(0, p), (p, 0)\}$. Thus

$$A = a^p \sum_{(i,j) \in I_1} \frac{1}{p} \binom{p}{i} \binom{p}{j} X^{j+ip^s} y^{p-j+p^s(p-i)} + T_y(X),$$

with $T_y \in \mathbb{F}_q[X]$. We first consider the sum. Since, for $j = 0$, $j = p$ and $i = 0$, one gets monomials whose degree (possibly after reduction $\pmod{\wp(\mathbb{F}_q[X])}$) is lower than $1 + p^s$, one can write

$$A = a^p \sum_{j=1}^{p-1} \frac{1}{p} \binom{p}{j} X^{j+p^{s+1}} y^{p-j} + R_y(X) + T_y(X) \pmod{\wp(\mathbb{F}_q[X])},$$

where $R_y(X)$ is a polynomial of $\mathbb{F}_q[X]$ of degree lower than $1 + p^s = 1 + r$. We now focus on the polynomial $T_y(X) \in \mathbb{F}_q[X]$. It is made of monomials of the forms $X^{i_0+i_1 p+\dots+i_{s-1} p^{s-1}}$, with $i_0 + i_1 + \dots + i_{s-1} = p$, and $X^{i_1 p+\dots+i_s p^s}$, with $i_1 + i_2 + \dots + i_s = p$. Since $X^{i_1 p+\dots+i_s p^s} = X^{i_1+\dots+i_s p^{s-1}} \pmod{\wp(\mathbb{F}_q[X])}$, it follows that T_y does not have any monomial with degree higher than $1 + p^s$ after reduction

mod $\wp(\mathbb{F}_q[X])$. Hence

$$A = a^p \sum_{j=1}^{p-1} \frac{1}{p} \binom{p}{j} X^{j+p^{s+1}} y^{p-j} + R_y^{[1]}(X) \pmod{\wp(\mathbb{F}_q[X])},$$

where $R_y^{[1]}(X)$ is a polynomial of $\mathbb{F}_q[X]$ with degree strictly lower than $1 + r$. Since $f_j(X) = X^{j+p^{s+1}} - X^{1+jp^{s-1}} \pmod{\wp(\mathbb{F}_q[X])}$ for all j in $\{1, \dots, p-1\}$, we conclude that

$$A = a^p \sum_{j=1}^{p-1} \frac{1}{p} \binom{p}{j} y^{p-j} f_j(X) + R_y^{[2]}(X) \pmod{\wp(\mathbb{F}_q[X])},$$

where $R_y^{[2]}(X)$ is a polynomial of $\mathbb{F}_q[X]$ of degree lower than $1 + r$. Then

$$A = \sum_{j=1}^{p-1} c_j(y) f_j(X) + R_y^{[2]}(X) \pmod{\wp(\mathbb{F}_q[X])},$$

with $c_j(y) := a^p \frac{1}{p} \binom{p}{j} y^{p-j} \in \mathbb{F}_q$. It follows that, modulo $\wp(\mathbb{F}_q[X])$,

$$A = \sum_{j=1}^{p-1} (F^e - \text{id})(c_j(y) W_j) + R_y^{[2]}(X) = (F - \text{id}) \sum_{j=1}^{p-1} P_j(W_j) + R_y^{[2]}(X),$$

where $P_j(W_j) = (\text{id} + F + \dots + F^{e-1})(c_j(y) W_j) \in \mathbb{F}_q[W_j]$. We gather that

$$\begin{aligned} \wp(\sigma[W_0, W_p] - [W_0, W_p]) \\ = \wp\left(\left[0, \sum_{j=1}^{p-1} P_j(W_j)\right]\right) + [0, R_y^{[2]}(X)] \pmod{\wp(W_2(\mathbb{F}_q[X]))}. \end{aligned}$$

As a consequence, $[0, R_y^{[2]}(X)]$ lies in $\wp(W_2(K_S^{m_2}))$, so there exists $V \in K_S^{m_2}$ such that $V^p - V = R_y^{[2]}(X)$. Accordingly, $K(V)$ is a K -subextension of $K_S^{m_2}$ with conductor $1 + \deg(R_y^{[2]}(X)) \leq 1 + r$. In particular, $K(V) \subset K_S^{r+1} = K = \mathbb{F}_q(X)$, which implies that $R_y^{[2]}(X) \in \wp(K)$. Therefore,

$$\wp(\sigma[W_0, W_p] - [W_0, W_p]) = \wp\left(\left[0, \sum_{j=1}^{p-1} P_j(W_j)\right]\right) \pmod{\wp(W_2(K))},$$

which allows us to conclude that $\sigma(W_p)$ is in $L = K(W_0, W_1, \dots, W_p)$. This finishes the proof of Proposition 6.7.3.

4. Since $L \subset K_S^{m_2}$ and $L \not\subset K_S^{m_2-1}$, the formula in Proposition 6.5.3 yields

$$\begin{aligned} g_L &= 1 + [L : K] \left(-1 + \frac{m_2}{2} \right) - \frac{1}{2} \sum_{j=0}^{m_2-1} [L \cap K_S^j : K] \\ &= 1 + p^{2+(p-1)e} \left(-1 + \frac{p^{s+1} + p + 1}{2} \right) \\ &\quad - \frac{1}{2} \left(r + 2 + (m_2 - p - (r + 2) + 1)p + \sum_{i=1}^{p-1} p^{1+ie} \right) \\ &= \frac{1}{2} p^{2+(p-1)e} (p^{s+1} + p - 1) - \frac{1}{2} \left(p^s + p^{s+2} - p^{s+1} + \sum_{i=1}^{p-1} p^{1+i2s} \right) \\ &= \frac{1}{2} p^{2+(p-1)e} (p^{s+1} + p - 1) - \frac{1}{2} p^s (p^2 - p + 1) - \frac{1}{2} p^{2s+1} (1 + q + q^2 + \dots + q^{p-2}). \end{aligned}$$

5. See Proposition 6.5.4. □

Remark 6.8. For $p = 2$, the equations given in Proposition 6.7 become

$$\begin{aligned} W_0^p - W_0 &= f_0(X) := X^{1+r}, \\ W_1^q - W_1 &= f_1(X) := X^{p^{s-1}} (X^q - X), \\ [W_0, W_2]^p - [W_0, W_2] &= [f_0(X), 0]. \end{aligned}$$

This last equation is no longer totally split over \mathbb{F}_q . One can circumvent this by replacing it with

$$[W_0, W_2]^p - [W_0, W_2] = [c^r X^{1+r}, 0] - [c X^{1+r}, 0] \quad \text{with} \quad c^r + c = 1.$$

In this case, we obtain the same results as in Proposition 6.7. The proof is left to the reader.

Proposition 6.7 can be generalized to construct a big action whose second ramification group G_2 is abelian of exponent as large as we want.

Proposition 6.9. *We keep the notation of Proposition 6.7. In particular, $q = p^e$, with $p > 2$, $e = 2s$ and $s \geq 1$. Let $n \geq 2$. Put $m_n := 1 + p^{n-1} (1 + p^s)$. If $q/(-1 + m_n/2) > 2p/(p-1)$, the pair $(C(m_n), G(m_n))$, as defined in Proposition 6.5, is a big action with a second ramification group $G_S(m_n)$ abelian of exponent at least p^n .*

Proof. Proposition 6.5.4 first ensures that $(C(m_n), G(m_n))$ is a big action. Consider the p^n -cyclic extension $K(W_1, \dots, W_n)/K$ parametrized with Witt vectors of length n as

$$[W_1, \dots, W_n]^p - [W_1, \dots, W_n] = [f_0(X), 0, \dots, 0],$$

where $f_0(X) = a X^{1+r}$ is defined as in Proposition 6.7, i.e., $r = p^s$, $a^r + a = 0$, $a \neq 0$. The same proof as in Proposition 6.7.1 shows that all places of S completely split in $K(W_1, \dots, W_n)$. Moreover, by [Garuti 2002] (Theorem 1.1) the conductor of the extension $K(W_1, \dots, W_n)$ is $1 + \max\{p^{n-1}(1 + p^s), 0\} = m_n$. It follows that $K(W_1, \dots, W_n)$ is included in $K_S^{m_n}$. Therefore, $G_S(m_n)$ has a quotient of exponent p^n and the claim follows. \square

The next proposition is an analogue of Proposition 6.7 in the case where e is odd. We do not spell out the proof, which is in the main similar to the proof of Proposition 6.7. Note that, contrary to the case where e is even, the equations still work for $p = 2$.

Proposition 6.10. *We keep the notation defined above. In particular, $K = \mathbb{F}_q(X)$ with $q = p^e$. Assume that $e = 2s - 1$, with $s \geq 2$, and put $r := \sqrt{qp} = p^s$. We define*

$$f_i(X) = X^{ir/p} (X^q - X) = X^{ip^{s-1}} (X^q - X) \quad \text{for all } i \in \{1, \dots, p - 1\},$$

$$g_i(X) = X^{ir/p^2} (X^q - X) = X^{ip^{s-2}} (X^q - X) \quad \text{for all } i \in \{1, \dots, p - 1\}.$$

Let $L := K(W_i, V_j)_{1 \leq i \leq p, 1 \leq j \leq p-1}$ be the extension of K parametrized by the Artin-Schreier-Witt equations

$$W_i^q - W_i = f_i(X) \text{ for all } i \in \{1, \dots, p - 1\},$$

$$V_j^q - V_j = g_j(X) \text{ for all } j \in \{1, \dots, p - 1\},$$

$$[W_1, W_p]^p - [W_1, W_p] = [X^{1+p^s}, 0] - [X^{1+p^{s-1}}, 0].$$

Finally, put $L_{i,0} := K(W_k)_{1 \leq k \leq i}$ and $L_{p-1,j} := K(W_i, V_k)_{1 \leq i \leq p-1, 1 \leq k \leq j}$, for all i and j in $\{1, \dots, p - 1\}$.

1. L is an abelian extension of K such that every place in S completely splits in L , satisfying $L_{i,0} \subset K_S^{p^s+i+1}$ for all $i, j \in \{1, \dots, p - 1\}$, $L_{p-1,j} \subset K_S^{p^{s+1}+j+1}$, and $L \subset K_S^{m_2}$, where $m_2 = p^{s+1} + p + 1$ is the integer defined in Proposition 6.4. (See table below on the next page.)
2. The extension L/K has degree $[L : K] = p^{2(p-1)e+1}$. Let G_L be its Galois group. Then

$$G_L \simeq \mathbb{Z}/p^2\mathbb{Z} \times (\mathbb{Z}/p\mathbb{Z})^t \quad \text{with } t = 2(p - 1)e - 1.$$

3. The extension L/K is stable under the translations $X \mapsto X + y$, with $y \in \mathbb{F}_q$. Therefore, the translations by \mathbb{F}_q extend to form a p -group of \mathbb{F}_q -automorphisms of L , say G , with the exact sequence

$$0 \longrightarrow G_L \longrightarrow G \longrightarrow \mathbb{F}_q \longrightarrow 0.$$

4. Let g_L be the genus of the extension L/K . Then

$$g_L = \frac{1}{2} \left(p^{1+2(p-1)e} (p^{s+1} + p - 1) - p^{(p-1)e} (p^{s+1} - p^s - p + 1) - p^s + p^e \sum_{i=0}^{2p-3} q^i \right).$$

In particular, when e grows large, $g_L \sim \frac{1}{2} p^{2+4s(p-1)+s}$ and $t = O(\log_p g_L)$.

We gather here the conductors, degrees and equations of each extension:

ext'n	conductor m	$[L_{i,j}:K]$	new equations
K	$0 \leq m \leq r+1 = p^s+1$	1	
$L_{1,0}$	$m = r+2 = p^s+2$	p^e	$W_1^q - W_1 = f_1(X)$
\dots	\dots	\dots	\dots
$L_{i,0}$	$m = p^s+i+1$	p^{ie}	$W_i^q - W_i = f_i(X)$
\dots	\dots	\dots	\dots
$L_{p-1,0}$	$p^s+p \leq m \leq p^{s+1}+1$	$p^{(p-1)e}$	$W_{p-1}^q - W_{p-1} = f_{p-1}(X)$
$L_{p-1,1}$	$m = p^{s+1}+2 = m_2 - (p-1)$	p^{pe}	$V_1^q - V_1 = g_1(X)$
\dots	\dots	\dots	\dots
$L_{p-1,j}$	$m = p^{s+1}+j+1 = m_2 - (p-j)$	$p^{(p+j-1)e}$	$V_j^q - V_j = g_j(X)$
\dots	\dots	\dots	\dots
$L_{p-1,p-1}$	$m = p^{s+1}+p = m_2 - 1$	$p^{2(p-1)e}$	$V_{p-1}^q - V_{p-1} = g_{p-1}(X)$
L	$m = p^{s+1}+p+1 = m_2$	$p^{1+2(p-1)e}$	$[W_1, W_p]^p - [W_1, W_p] = [X^{1+p^s}, 0] - [X^{1+p^{s-1}}, 0]$

7. A local approach to big actions

Let (C, G) be a big action. We recall that there exists a point $\infty \in C$ such that G is equal to $G_1(\infty)$ the wild inertia subgroup of G at ∞ , which means that the cover $\pi : C \rightarrow C/G$ is totally ramified at ∞ . Moreover, the quotient curve C/G is isomorphic to the projective line \mathbb{P}_k^1 and π is étale above the affine line $\mathbb{A}_k^1 = \mathbb{P}_k^1 - \pi(\infty) = \text{Spec } k[T]$. The inclusion $k[T] \subset k((T^{-1}))$ induces a Galois extension $k(C) \otimes_{k(T)} k((T^{-1})) =: k((Z))$ over $k((T^{-1}))$, with group equal to G and ramification groups in lower notation equal to $G_i := G_i(\infty)$. Then the genus of C is given by (2-2) as $g = \frac{1}{2} \sum_{i \geq 2} (|G_i| - 1) > 0$. It follows that

$$\frac{|G|}{\sum_{i \geq 2} (|G_i| - 1)} = \frac{|G|}{2g} > \frac{p}{p-1}.$$

This leads to:

Definition 7.1. A local big action is any pair $(k((Z)), G)$ where G is a finite p -subgroup of $\text{Aut}_k(k((Z)))$ whose ramification groups in lower notation at ∞ satisfy

the inequalities

$$g(G) := \frac{1}{2} \sum_{i \geq 2} (|G_i| - 1) > 0 \quad \text{and} \quad \frac{|G|}{g(G)} > \frac{2p}{p-1}.$$

It follows from the Katz–Gabber Theorem (see [Katz 1986, Theorem 1.4.1] or [Gille 2000, corollaire 1.9]) that big actions (C, G) and local big actions $(k((Z)), G)$ are in one-to-one correspondence via the following functor induced by the inclusion $k[T] \subset k((T^{-1}))$:

$$\left\{ \begin{array}{l} \text{finite étale Galois covers} \\ \text{of } \text{Spec } k[T] \\ \text{with Galois group a } p\text{-group} \end{array} \right\} \longrightarrow \left\{ \begin{array}{l} \text{finite étale Galois covers} \\ \text{of } \text{Spec } k((T^{-1})) \\ \text{with Galois group a } p\text{-group} \end{array} \right\}$$

Thus we can infer from the global point of view properties related to local extensions that would be difficult to prove directly. For instance, if $(k((Z)), G)$ is a local big action, we deduce that G_2 is strictly included in G_1 . Moreover, we obtain

$$\frac{|G|}{g(G)^2} \leq \frac{4p}{(p-1)^2}.$$

References

- [Auer 1999] R. Auer, *Ray class fields of global function fields with many rational places*, Ph.D. thesis, University of Oldenburg, 1999, Available at <http://www.bis.uni-oldenburg.de/dissertation/fb06.html>.
- [Auer 2000] R. Auer, “Ray class fields of global function fields with many rational places”, *Acta Arith.* **95**:2 (2000), 97–122. MR 2002e:11162 Zbl 0963.11067
- [Bertin and Mézard 2000] J. Bertin and A. Mézard, “Déformations formelles des revêtements sauvagement ramifiés de courbes algébriques”, *Invent. Math.* **141**:1 (2000), 195–238. MR 2001f:14023 Zbl 0993.14014
- [Bertin and Romagny 2008] J. Bertin and M. Romagny, “Champs d’Hurwitz”, preprint, 2008, Available at http://people.math.jussieu.fr/~romagny/champs_de_hurwitz.pdf.
- [Bourbaki 1983] N. Bourbaki, *Algèbre commutative, Chapitres 8 et 9*, Masson, Paris, 1983.
- [Bouw 2000] I. I. Bouw, “The p -rank of curves and covers of curves”, pp. 267–277 in *Courbes semi-stables et groupe fondamental en géométrie algébrique* (Luminy, 1998), edited by J.-B. Bost et al., Progr. Math. **187**, Birkhäuser, Basel, 2000. MR 2001j:14042 Zbl 0979.14015
- [Breuer 2000] T. Breuer, *Characters and automorphism groups of compact Riemann surfaces*, London Mathematical Society Lecture Note Series **280**, Cambridge University Press, Cambridge, 2000. MR 2002i:14034 Zbl 0952.30001
- [Conder 1990] M. Conder, “Hurwitz groups: a brief survey”, *Bull. Amer. Math. Soc. (N.S.)* **23**:2 (1990), 359–370. MR 91d:20032 Zbl 0716.20015
- [Cornelissen and Kato 2003] G. Cornelissen and F. Kato, “Equivariant deformation of Mumford curves and of ordinary curves in positive characteristic”, *Duke Math. J.* **116**:3 (2003), 431–470. MR 2004c:14044 Zbl 1092.14032

- [Elkies 1999a] N. D. Elkies, “The Klein quartic in number theory”, pp. 51–101 in *The eightfold way*, edited by S. Levy, Math. Sci. Res. Inst. Publ. **35**, Cambridge Univ. Press, Cambridge, 1999. MR 2001a:11103 Zbl 0991.11032
- [Elkies 1999b] N. D. Elkies, “Linearized algebra and finite groups of Lie type, I: Linear and symplectic groups”, pp. 77–107 in *Applications of curves over finite fields* (Seattle, 1997), edited by M. D. Fried, Contemp. Math. **245**, Amer. Math. Soc., Providence, RI, 1999. MR 2001a:20082 Zbl 0976.20032
- [Garuti 2002] M. A. Garuti, “Linear systems attached to cyclic inertia”, pp. 377–386 in *Arithmetic fundamental groups and noncommutative algebra* (Berkeley, 1999), edited by M. D. Fried and Y. Ihara, Proc. Sympos. Pure Math. **70**, Amer. Math. Soc., Providence, RI, 2002. MR 2003i:14014 Zbl 1072.14017
- [Gille 2000] P. Gille, “Le groupe fondamental sauvage d’une courbe affine en caractéristique $p > 0$ ”, pp. 217–231 in *Courbes semi-stables et groupe fondamental en géométrie algébrique* (Luminy, 1998), edited by J.-B. Bost et al., Progr. Math. **187**, Birkhäuser, Basel, 2000. MR 2002a:14027 Zbl 0978.14034
- [Giulietti and Korchmáros 2007] M. Giulietti and G. Korchmáros, “On large automorphism groups of algebraic curves in positive characteristic”, preprint, 2007. arXiv 0706.2320
- [Goss 1996] D. Goss, *Basic structures of function field arithmetic*, Ergebnisse der Math. (3) **35**, Springer, Berlin, 1996. MR 97i:11062 Zbl 0874.11004
- [Huppert 1967] B. Huppert, *Endliche Gruppen, I*, Grundlehren der Math. Wiss. **134**, Springer, Berlin, 1967. MR 37 #302 Zbl 0217.07201
- [Hurwitz 1892] A. Hurwitz, “Ueber algebraische Gebilde mit eindeutigen Transformationen in sich”, *Math. Ann.* **41**:3 (1892), 403–442. MR 1510753
- [Katz 1986] N. M. Katz, “Local-to-global extensions of representations of fundamental groups”, *Ann. Inst. Fourier (Grenoble)* **36**:4 (1986), 69–106. MR 88a:14032 Zbl 0564.14013
- [Kontogeorgis 2007] A. Kontogeorgis, “On the tangent space of the deformation functor of curves with automorphisms”, *Algebra Number Theory* **1**:2 (2007), 119–161. MR 2008j:14056
- [Kulkarni 1997] R. S. Kulkarni, “Riemann surfaces admitting large automorphism groups”, pp. 63–79 in *Extremal Riemann surfaces* (San Francisco, 1995), edited by J. R. Quine and P. Sarnak, Contemp. Math. **201**, Amer. Math. Soc., Providence, RI, 1997. MR 98g:30070 Zbl 0863.30050
- [Lauter 1999] K. Lauter, “A formula for constructing curves over finite fields with many rational points”, *J. Number Theory* **74**:1 (1999), 56–72. MR 99k:11088 Zbl 1044.11054
- [Leedham-Green and McKay 2002] C. R. Leedham-Green and S. McKay, *The structure of groups of prime power order*, London Mathematical Society Monographs. New Series **27**, Oxford University Press, Oxford, 2002. Oxford Science Publications. MR 2003f:20028 Zbl 1008.20001
- [Lehr and Matignon 2005] C. Lehr and M. Matignon, “Automorphism groups for p -cyclic covers of the affine line”, *Compos. Math.* **141**:5 (2005), 1213–1237. MR 2006f:14029 Zbl 1083.14028
- [Macbeath 1961] A. M. Macbeath, “On a theorem of Hurwitz”, *Proc. Glasgow Math. Assoc.* **5** (1961), 90–96. MR 26 #4244 Zbl 0134.16603
- [Macbeath 1965] A. M. Macbeath, “On a curve of genus 7”, *Proc. London Math. Soc.* (3) **15** (1965), 527–542. MR 31 #1605 Zbl 0146.42705
- [Magaard et al. 2002] K. Magaard, T. Shaska, S. Shpectorov, and H. Völklein, “The locus of curves with prescribed automorphism group”, pp. 112–141 in *Communications in arithmetic fundamental groups* (Kyoto, 1999/2001), Sūrikaiseikikenkyūsho Kōkyūroku **1267**, 2002. MR 1954371 arXiv math.0205314

- [Marshall 1971] M. A. Marshall, “Ramification groups of abelian local field extensions”, *Canad. J. Math.* **23** (1971), 271–281. MR 43 #189 Zbl 0211.06704
- [Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, Princeton, 1980. MR 81j:14002 Zbl 0433.14012
- [Nakajima 1987] S. Nakajima, “ p -ranks and automorphism groups of algebraic curves”, *Trans. Amer. Math. Soc.* **303**:2 (1987), 595–607. MR 88h:14037 Zbl 0644.14010
- [Pries 2005] R. Pries, “Equiramified deformations of covers in positive characteristic”, preprint, 2005. arXiv math.AG/0403056
- [Rocher 2008a] M. Rocher, “Large p -group actions with a p -elementary abelian derived group”, preprint, 2008. To appear in *J. Algebra*. arXiv 0801.3834
- [Rocher 2008b] M. Rocher, “Large p -group actions with $|G|/g^2 \geq 4/(p^2 - 1)^2$ ”, preprint, 2008. arXiv 0804.3494
- [Schmid 1938] H. L. Schmid, “Über die Automorphismen eines algebraischen Funktionenkörpers von Primzahlcharakteristik.”, *J. Reine Angew. Math.* **179** (1938), 5–15. Zbl 0019.00301
- [Serre 1968] J.-P. Serre, *Corps locaux*, 2nd ed., Hermann, Paris, 1968. Translated as *Local fields*, Springer, New York, 1979. MR 50 #7096 Zbl 0174.24301
- [Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1986. MR 87g:11070 Zbl 0585.14026
- [Stichtenoth 1973a] H. Stichtenoth, “Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik, I: Eine Abschätzung der Ordnung der Automorphismengruppe”, *Arch. Math. (Basel)* **24** (1973), 527–544. MR 49 #2749 Zbl 0282.14006
- [Stichtenoth 1973b] H. Stichtenoth, “Über die Automorphismengruppe eines algebraischen Funktionenkörpers von Primzahlcharakteristik, II: Ein spezieller Typ von Funktionenkörpern”, *Arch. Math. (Basel)* **24** (1973), 615–631. MR 53 #8068 Zbl 0282.14007
- [Stichtenoth 1993] H. Stichtenoth, *Algebraic function fields and codes*, Springer, Berlin, 1993. MR 94k:14016 Zbl 0816.14011
- [Suzuki 1982] M. Suzuki, *Group theory, I*, Grundlehren der Math. Wiss. **247**, Springer, Berlin, 1982. MR 82k:20001c Zbl 0472.20001
- [Suzuki 1986] M. Suzuki, *Group theory, II*, Grundlehren der Math. Wiss. **248**, Springer, New York, 1986. MR 87e:20001 Zbl 0586.20001

Communicated by Bjorn Poonen

Received 2008-02-01 Revised 2008-08-14 Accepted 2008-09-17

Michel.Matignon@math.u-bordeaux1.fr

*Institut de Mathématiques de Bordeaux,
Université de Bordeaux 1, 351 cours de la Libération,
33405 Talence Cedex, France
<http://www.math.u-bordeaux1.fr/~matignon/>*

Magali.Rocher@math.u-bordeaux1.fr

*Institut de Mathématiques de Bordeaux,
Université de Bordeaux 1, 351 cours de la Libération,
33405 Talence Cedex, France
<http://www.math.u-bordeaux.fr/~mrocher/>*

Inner derivations of alternative algebras over commutative rings

Ottmar Loos, Holger P. Petersson and Michel L. Racine

Erhard Neher zum 60. Geburtstag gewidmet

We define Lie multiplication derivations of an arbitrary non-associative algebra A over any commutative ring and, following an approach due to K. McCrimmon, describe them completely if A is alternative. Using this description, we propose a new definition of inner derivations for alternative algebras, among which Schafer's standard derivations and McCrimmon's associator derivations occupy a special place, the latter being particularly useful to resolve difficulties in characteristic 3. We also show that octonion algebras over any commutative ring have only associator derivations.

Introduction

There are many important properties satisfied by inner but not in general by all derivations of Lie, associative or (linear) Jordan algebras. A particularly important one may be described as follows.

Let $f: A \rightarrow B$ be a homomorphism of non-associative algebras and D a derivation of A . We say a derivation D' of B is *f-related to D* if

$$f(D(a)) = D'(f(a))$$

for all $a \in A$. In general, there will be no such D' . The situation is better for inner derivations, which satisfy the following

Mapping Principle. *Given a homomorphism $f: A \rightarrow B$ of algebras (Lie, associative or Jordan), every inner derivation D of A admits an inner derivation D' of B that is *f-related to D* .*

MSC2000: primary 17D05; secondary 17A36, 17A45, 17B40.

Keywords: inner derivations, alternative algebras, derivation functors, composition algebras, automorphisms.

At the authors' request, this paper was not edited by the publisher's production staff.

Indeed, an inner derivation of A can be naturally expressed in terms of left and right multiplication operators. This suggests and yields an inner derivation D' of B in a natural way which is f -related to D .

Properties of this kind tie up nicely with the fact that, under suitable regularity conditions, all derivations of A are inner.

This satisfactory state of affairs has led Schafer [19] (see also [20, II, §3, p. 21]) to propose a notion of inner derivations for arbitrary non-associative algebras over a field F that reduces to the usual one when dealing with Lie or unital associative (resp. Jordan) algebras [19]. Moreover, inner derivations in his sense always form an ideal in the full derivation algebra, so it follows also in this generality that all derivations are inner provided (i) non-zero inner derivations exist and (ii) the derivation algebra is simple (as a Lie algebra).

While (i) is a harmless condition rarely causing any difficulties, (ii) is a much more delicate one. Moreover, it points to a strong link between Lie theory and non-associative algebras in general that has dominated the scene for decades. For example, the interest in derivations of *alternative* algebras grew out of the fundamental observation, due to various authors, most notably É. Cartan [6], Jacobson [9; 11], Bannow [2] and Alberca-Elduque-Martín-Navarro [1], that the derivations of an octonion (= Cayley) algebra C over F form a central simple Lie algebra of type G_2 if and only if F has characteristic not 3; in particular, all derivations of C are inner in this case. Going one step further, the proof of [1, Prop. 1] may be combined with a base field extension argument to show that an octonion algebra over any field (possibly of characteristic 3) has only inner derivations.

In spite of these remarkable advances, a particularly annoying deficiency of Schafer's approach remains: again in the setting of alternative algebras, inner derivations in his sense fail to satisfy the Mapping Principle. Already implicit in Schafer's own work on the subject (cf. [20, p. 78]), this deficiency comes into full view through their characteristic-free description in McCrimmon's unpublished monograph on alternative algebras [15] that only quite recently has been made accessible to the mathematical public.

In view of the preceding circumstances, E. Neher has suggested to relinquish altogether the idea of a universal definition for inner derivations of arbitrary non-associative algebras. Instead, he argued, they should be defined, as in the old days, for each relevant class of non-associative algebras individually, always taking into account the special requirements of the theory at hand. In the present paper, Neher's suggestions will be implemented for the class of alternative algebras over an arbitrary commutative ring k . The basic concepts and results of the paper may be summarized as follows.

A slight modification of Schafer's original approach will lead us in Section 1 to what we call *Lie multiplication derivations*, which turn out to be the same as inner

derivations in the sense of Schafer when dealing with unital algebras (Remark to Prop. 1.4) but not in general (Example 1.5). The Lie multiplication derivations always form an ideal in the full derivation algebra and specialize to inner derivations (in the usual sense) of associative and linear Jordan algebras even when these fail to have a unit. We then proceed to show that the Lie multiplication algebra of a non-associative k -algebra A commutes with flat base change if A is finitely spanned as a k -module (Cor. 1.10). The same conclusion holds for the algebra of multiplication derivations if A is also projective as a k -module and its automorphism group is smooth as a group scheme (Cor. 1.12).

In Section 2, we follow McCrimmon [15, A5.2] to describe the Lie multiplication derivations of an alternative k -algebra A (Thm. 2.3). It follows immediately from this description that they do not in general satisfy the Mapping Principle. For this reason, we define *inner derivations* of A by a condition that is more restrictive than the one of just being a Lie multiplication derivation and automatically ensures the validity of the Mapping Principle (2.5). Adapting McCrimmon's terminology (loc. cit.) to the present set-up, we also introduce a few subclasses of inner derivations that turn out to be useful later on. Among them, *associator derivations* (2.5(a)), having the form $\sum [L_{a_i}, R_{b_i}]$ where $a_i, b_i \in A$ satisfy $\sum [a_i, b_i] = 0$, and *standard derivations* (2.5(b)), which are sums of operators $D_{a,b} = [L_a, L_b] + [L_a, R_b] + [R_a, R_b]$ for $a, b \in A$, seem to be of particular importance. Standard derivations made their first appearance in the work of Schafer [19] and, historically, constitute the oldest class of derivations known for arbitrary alternative algebras. Associator derivations, on the other hand, which can be transformed quite easily into standard ones if $3A = A$ (Prop. 2.7(b)), are apparently best suited for dealing with difficulties in characteristic 3, for which alternative algebras are notorious. The aforementioned definitions give rise to various ideals in the full derivation algebra that all commute with flat base change provided the algebra itself is finitely spanned as a k -module (Prop. 2.9).

Let again $f: A \rightarrow B$ be a homomorphism of algebras and D an inner derivation of A . The derivation D' of B furnished by the Mapping Principle will in general not be uniquely determined by D , so we don't have a natural map from inner derivations of A to those of B : the inner derivations of A do not depend functorially on A . In many examples, functoriality can be achieved at the cost of replacing the inner derivation algebra by a suitable central extension. This problem is addressed in Section 3. We introduce the notion of *derivation functor* and show that the ideals of standard, associator and commutator derivations are all induced by suitable derivation functors. These derivation functors commute with flat base change, and the standard derivation functor even with arbitrary base change (Proposition 3.12), without finiteness assumptions on the underlying algebra.

In Section 4, we take up the study of octonion algebras over commutative rings. They will be introduced here in a “rational” manner, i.e., without the need of changing scalars, along the lines of [17]. We define a *splitting* of any octonion algebra C over k as an isomorphism from Z onto C , where $Z = \text{Zor}(k)$ stands for the split octonion algebra of ordinary Zorn vector matrices over k . We then proceed to show, using the notion of splitting datum (4.7), that the functor assigning to each unital commutative associative k -algebra R the set of splittings of $C \otimes_k R$ over R is a smooth affine torsor in the étale topology whose structure group is the automorphism group scheme of Z (Thm. 4.10). As immediate consequences, we conclude that octonion algebras, just like Azumaya algebras, become split after a faithfully flat (even étale) extension (Cor. 4.11) and that their automorphism group schemes are smooth (Cor. 4.12). In particular, our definition of octonion algebras is equivalent to the one given by Thakur [22] over base rings containing $\frac{1}{2}$.

In the final section of the paper, the preceding results are applied to show that an octonion algebra C over an arbitrary commutative ring has only associator derivations (Thm. 5.1). This theorem is new even when the base ring is a field. After a reduction to the case where C is reduced, the proof consists in a careful analysis of the $\mathbb{Z}/3\mathbb{Z}$ -grading given on the derivation algebra (Example 5.4, Prop. 5.5) by an elementary idempotent of C (cf. 4.4).

Notations. Throughout we fix an arbitrary commutative ring k . Unadorned tensor products will always be taken over k . We write $\text{Spec}(k)$ for prime spectrum of k , i.e., for the totality of all prime ideals in k , equipped with the Zariski topology. The category of commutative associative k -algebras with 1 will be denoted by $k\text{-alg}$. For $R \in k\text{-alg}$, a k -module M and $x \in M$, we abbreviate $M_R = M \otimes R$ as R -modules and $x_R = x \otimes 1_R \in M_R$; we also write f_R for the R -linear extension of a k -linear map f between k -modules. The standard terminology of non-associative algebras (including notation) will be used as in Schafer [20], except that (linear) operators will always act on the left so, e.g., the equations $L_a b = ab = R_b a$ describe left and right multiplications in a non-associative algebra A , and the associator of elements $a, b, c \in A$ (resp. the commutator of a, b) will be indicated by $[a, b, c] = (ab)c - a(bc)$ (resp. $[a, b] = ab - ba$). The symbols \mathbb{N} and \mathbb{Z} denote the positive natural numbers and the rational integers, respectively.

1. Lie multiplication derivations

In this section, we fix an arbitrary non-associative algebra A over k . We do not assume that A has a unit.

1.1. The Lie multiplication algebra. The Lie algebra defined on the k -module $\text{End}_k(A)$ by the usual commutator of linear maps will be denoted by $\mathfrak{gl}(A)$. The

subalgebra of $\mathfrak{gl}(A)$ generated by all left and right multiplication operators of arbitrary elements in A is called the *Lie multiplication algebra* of A , denoted by $\mathfrak{L}(A)$. For example, if A is associative, then $\mathfrak{L}(A) = L_A + R_A$. Or, if A is a (linear) Jordan algebra over a ring containing $\frac{1}{2}$, then $\mathfrak{L}(A) = L_A + [L_A, L_A]$.

1.2. Derivations. Recall that a *derivation* of A is a linear map $D: A \rightarrow A$ satisfying one (hence all) of the following equivalent relations, for all $x, y \in A$:

$$D(xy) = (Dx)y + x(Dy), \tag{1-1}$$

$$[D, L_x] = L_{Dx},$$

$$[D, R_y] = R_{Dy}. \tag{1-2}$$

The derivations of A form a Lie algebra (more precisely, a subalgebra of $\mathfrak{gl}(A)$), denoted by $\text{Der}(A)$. The elements of $\text{Der}(A)$ also act on commutators and associators in a derivation-like manner, i.e., we have

$$D([x, y]) = [Dx, y] + [x, Dy], \tag{1-3}$$

$$D([x, y, z]) = [Dx, y, z] + [x, Dy, z] + [x, y, Dz] \tag{1-4}$$

for all $x, y, z \in A$. By (1-1), (1-2), $\text{Der}(A)$ acts on $\mathfrak{L}(A)$ through the adjoint representation of $\mathfrak{gl}(A)$, i.e.,

$$[\text{Der}(A), \mathfrak{L}(A)] \subseteq \mathfrak{L}(A). \tag{1-5}$$

1.3. The ideal of Lie multiplication derivations. We write $\hat{A} = k1 \oplus A$ for the algebra obtained by adjoining a unit $1 = 1_{\hat{A}}$ to A and \hat{L}, \hat{R} for the left, right multiplication, respectively, of \hat{A} . The relations $\hat{L}_{\alpha 1+a} = \alpha \text{Id}_{\hat{A}} + \hat{L}_a, \hat{R}_{\alpha 1+a} = \alpha \text{Id}_{\hat{A}} + \hat{R}_a$ ($\alpha \in k, a \in A$) show $\mathfrak{L}(\hat{A}) = k\text{Id}_{\hat{A}} + \hat{\mathfrak{L}}(A)$, where $\hat{\mathfrak{L}}(A)$ stands for the subalgebra of $\mathfrak{gl}(\hat{A})$ generated by $\hat{L}_A \cup \hat{R}_A$. Observe that there are no natural maps $\mathfrak{L}(A) \rightarrow \mathfrak{L}(\hat{A})$ satisfying $L_a \mapsto \hat{L}_a$ (resp. $R_a \mapsto \hat{R}_a$) unless $aA = \{0\}$ (resp. $Aa = \{0\}$) implies $a = 0$. But since $A \subseteq \hat{A}$ is an ideal, we obtain the inclusions

$$\mathfrak{L}(\hat{A}) \subseteq \mathfrak{g} := \{f \in \mathfrak{gl}(\hat{A}) \mid f(A) \subseteq A\}, \tag{1-6}$$

$$\hat{\mathfrak{L}}(A) \subseteq \mathfrak{g}' := \{f \in \mathfrak{gl}(\hat{A}) \mid f(\hat{A}) \subseteq A\} \subseteq \mathfrak{g}$$

as subalgebras of $\mathfrak{gl}(\hat{A})$, and the restriction homomorphism $\rho: \mathfrak{g} \rightarrow \mathfrak{gl}(A)$ satisfies $\rho(\hat{L}_a) = L_a, \rho(\hat{R}_a) = R_a$ for all $a \in A$, hence $\rho(\hat{\mathfrak{L}}(A)) = \mathfrak{L}(A)$.

There is a natural embedding $\text{Der}(A) \rightarrow \text{Der}(\hat{A}), D \mapsto \hat{D}$ of Lie algebras, where \hat{D} stands for the unique linear extension of $D \in \text{Der}(A)$ to \hat{A} given by $\hat{D}1 = 0$. It follows from (1-5) applied to \hat{A} in place of A that

$$\text{LMDer}(A) := \{D \in \text{Der}(A) \mid \hat{D} \in \mathfrak{L}(\hat{A})\} \subseteq \text{Der}(A) \tag{1-7}$$

is an ideal. The elements of $\text{LMDer}(A)$ are called *Lie multiplication derivations* of A . Indeed, as we will now see, they all belong to the Lie multiplication algebra of A and may thus be expressed as Lie polynomials in left and right multiplication operators by suitable elements of A .

1.4. Proposition. *The inclusion*

$$\text{LMDer}(A) \subseteq \mathfrak{L}(A) \cap \text{Der}(A)$$

always holds; it may be strengthened to the equality

$$\text{LMDer}(A) = \mathfrak{L}(A) \cap \text{Der}(A)$$

if A has a unit.

Proof. For the first part of the proposition, we must show $D \in \mathfrak{L}(A)$ for all $D \in \text{LMDer}(A)$. To this end, using 1.3, we decompose $\hat{D} \in \mathfrak{L}(\hat{A})$ as $\hat{D} = \alpha \text{Id}_{\hat{A}} + D'$ with $\alpha \in k$, $D' \in \hat{\mathfrak{L}}(A)$ and obtain $0 = \hat{D}1_{\hat{A}} = \alpha 1_{\hat{A}} + D'1_{\hat{A}}$, where the second summand on the right by (1-6) belongs to A . This implies $\alpha = 0$, $\hat{D} = D' \in \hat{\mathfrak{L}}(A)$, hence $D = \rho(\hat{D}) \in \rho(\hat{\mathfrak{L}}(A)) = \mathfrak{L}(A)$, as claimed.

For the second part, we assume A has a unit 1_A , put $e = 1_{\hat{A}} - 1_A$ and conclude $\hat{A} = ke \oplus A$ as a direct sum of ideals. This implies $\mathfrak{L}(\hat{A}) = k\text{Id}_{ke} \oplus \mathfrak{L}(A)$, the right-hand side being diagonally embedded into

$$\mathfrak{gl}(\hat{A}) = \begin{pmatrix} k \cdot \text{Id}_{ke} & \text{Hom}_k(A, ke) \\ \text{Hom}_k(ke, A) & \mathfrak{gl}(A) \end{pmatrix}. \tag{1-8}$$

On the other hand, given $D \in \text{Der}(A)$, we obtain $\hat{D}e = 0$ since D kills 1_A , and this amounts to $\text{Der}(A)^\wedge = \{\hat{D} \mid D \in \text{Der}(A)\} = \{0\} \oplus \text{Der}(A)$, the right-hand side again being embedded diagonally into (1-8). The assertion follows by comparing the decompositions for $\mathfrak{L}(\hat{A})$ and $\text{Der}(A)^\wedge$. □

Remark. Comparing Prop. 1.4 with [20, p. 21], we conclude that the Lie multiplication derivations of A and its inner derivations in the sense of Schafer are the same if A has a unit. In general, however, this need not be so, as may be seen from the following example.

1.5. Example. Equality does not always hold in Prop. 1.4. To see this, suppose A is associative. We first claim

$$\text{LMDer}(A) = \text{InDer}_{\text{ass}}(A) = \{L_a - R_a \mid a \in A\},$$

i.e., that the Lie multiplication derivations of A and its inner derivations (in the usual sense) are the same. As inner derivations of A obviously belong to

$\text{LMDer}(A)$, we need only worry about the converse, so let $D \in \text{LMDer}(A)$. Observing $\hat{D} \in \mathfrak{L}(\hat{A}) = k1_{\hat{A}} + \hat{L}_A + \hat{R}_A$ and $\hat{D}(1_{\hat{A}}) = 0$, we obtain $\hat{D} = \alpha 1_{\hat{A}} + \hat{L}_a + \hat{R}_b$ for some $\alpha \in k$, $a, b \in A$, hence $0 = \alpha 1_{\hat{A}} + (a + b)$, which yields $\alpha = 0$, $b = -a$, $D = L_a - R_a$, as claimed.

On the other hand, a derivation of A belonging to $\mathfrak{L}(A) = L_A + R_A$ need not be inner. To see this, suppose A is also commutative. Then there are no inner derivations other than zero, while $L_z \in L_A \subseteq \mathfrak{L}(A)$ for $z \in A$ is easily seen to be a derivation if and only if $AzA = \{0\}$, which in the absence of a unit element does not imply $L_z = 0$.

Remark. An analogous argument also works for a linear Jordan algebra J over k (with $\frac{1}{2} \in k$) since $D \in \text{LMDer}(J)$ implies $\hat{D} \in \mathfrak{L}(\hat{J}) = \hat{L}_J + [\hat{L}_J, \hat{L}_J] = k\text{Id}_J + \hat{L}_J + [\hat{L}_J, \hat{L}_J]$, $\hat{D}(1_J) = 0$, hence $\hat{D} = \alpha \text{Id}_J + \hat{L}_a + \sum[\hat{L}_{a_i}, \hat{L}_{b_i}]$ for some $\alpha \in k$, $a, a_i, b_i \in J$, and from $0 = \hat{D}(1_J) = \alpha 1_J + a$ we conclude $D = \sum[L_{a_i}, L_{b_i}]$. Thus the Lie multiplication derivations of J are just the inner ones in the usual sense:

$$\text{LMDer}(J) = \text{InDer}_{\text{Jord}}(J) = [L_J, L_J].$$

Not so, however, in the case of Lie algebras. The idea of defining Lie multiplication derivations by passing to the algebra \hat{A} seems to work well only when dealing with varieties of algebras that are stable under adjoining a unit.

Our principal objective in the present section will be to show that Lie multiplication derivations are well behaved under suitable scalar extensions. We begin by treating the analogous question for the Lie multiplication algebra $\mathfrak{L}(A)$.

1.6. Flat k -algebras. Let $R \in k\text{-alg}$ be a flat k -algebra, so R is flat as a k -module, equivalently, the assignment $M \mapsto M_R$ gives an exact functor from k -modules to R -modules. For a k -module M and a k -submodule $N \subseteq M$ with inclusion $i: N \rightarrow M$, we can and always will identify $N_R \subseteq M_R$ as an R -submodule via the injection $i_R: N_R \rightarrow M_R$.

The following easy lemma collects a few properties of flat k -algebras that are surely well known but seem to lack a convenient reference.

1.7. Lemma. *Conventions being as in 1.6, let $f: M \rightarrow M'$ be a k -linear map of k -modules and let $N, P \subseteq M, N' \subseteq M'$ be arbitrary k -submodules.*

- (a) $\text{Ker}(f)_R = \text{Ker}(f_R), \quad \text{Im}(f)_R = \text{Im}(f_R).$
- (b) $f(N)_R = f_R(N_R), \quad f^{-1}(N')_R = f_R^{-1}(N'_R).$
- (c) $(N \cap P)_R = N_R \cap P_R.$

(d) For every family $(N_\alpha)_{\alpha \in I}$ of k -submodules in M we have

$$\left(\sum_{\alpha \in I} N_\alpha \right)_R = \sum_{\alpha \in I} (N_\alpha)_R \subseteq M_R.$$

(e) If N is generated as a k -module by a family $(x_\alpha)_{\alpha \in I}$ of elements in M , then $N_R \subseteq M_R$ is generated as an R -module by the family $(x_{\alpha R})_{\alpha \in I}$ of elements in M_R .

Proof. By flatness, the functor $-\otimes R$ preserves kernels and co-kernels, which yields

(a). The first (resp. second) part of (b) follows by applying (a) to $f|_N : N \rightarrow M'$ (resp. to $\pi \circ f : M \rightarrow M'/N', \pi : M' \rightarrow M'/N'$ being the canonical projection).

In (c) we apply (b) with $N' = P$ to the natural embedding $i : N \rightarrow M$. For (d), we consider the canonical map $\bigoplus_{\alpha \in I} N_\alpha \rightarrow M$ determined by the inclusions $N_\alpha \rightarrow M$ and apply (a). Finally, in (e), we let M^0 be a free k -module with basis $(e_\alpha)_{\alpha \in I}$ and apply (a) to the k -linear map $M^0 \rightarrow M, e_\alpha \mapsto x_\alpha, \alpha \in I$. □

1.8. Proposition. *Conventions being as in 1.6, let $R \in k\text{-alg}$ be a flat k -algebra and write B for the k -subalgebra of A generated by a family $(x_\alpha)_{\alpha \in I}$ of elements in A . Then B_R is the R -subalgebra of A_R generated by the family $(x_{\alpha R})_{\alpha \in I}$ of elements in A_R .*

Proof. We denote by $(y_\beta)_{\beta \in J}$ the family of non-associative monomials built in A over the family $(x_\alpha)_{\alpha \in I}$. Then B is generated as a k -module by $(y_\beta)_{\beta \in J}$. By Lemma 1.7(e), $B_R \subseteq A_R$ is therefore generated as an R -module by the family $(y_{\beta R})_{\beta \in J}$, which consists precisely of the non-associative monomials built in A_R over the family $(x_{\alpha R})_{\alpha \in I}$. Thus B_R is generated as an R -algebra by $(x_{\alpha R})_{\alpha \in I}$. □

1.9. Finitely generated modules: base change of endomorphisms. Let M be a k -module and $S \in k\text{-alg}$ a unital commutative associative k -algebra. Then the natural map

$$\text{End}_k(M) \longrightarrow \text{End}_S(M_S), \quad f \longmapsto f_S,$$

extends to a homomorphism $\text{End}_k(M)_S \rightarrow \text{End}_S(M_S)$ of S -algebras, which is injective if S is a flat k -algebra and M is finitely generated [3, I, §2, Prop. 11]; we will then identify $\text{End}_k(M)_S \subseteq \text{End}_S(M_S)$ as an S -subalgebra accordingly. Under this identification, we even have equality $\text{End}_k(M)_S = \text{End}_S(M_S)$ if M is also projective [4, II, §5, Prop. 7]; in fact, equality then holds for *any* $S \in k\text{-alg}$.

1.10. Corollary. *If A is finitely generated as a k -module, its Lie multiplication algebra is stable under flat base change: For all flat k -algebras $S \in k\text{-alg}$, we have $\mathfrak{L}(A)_S = \mathfrak{L}(A_S)$ after the identifications of 1.9.*

Proof. $\mathfrak{L}(A)$ is generated by $L_A \cup R_A$ as a k -algebra. But $(L_a)_S = L_{a_S}, (R_a)_S = R_{a_S}$ for all $a \in A$. Hence, by Prop. 1.8, $\mathfrak{L}(A)_S$ and $\mathfrak{L}(A_S)$ are both generated as S -algebras by $L_{A_S} \cup R_{A_S}$. □

Remark. In this generality, Cor. 1.10 is due to E. Neher (oral communication), who also pointed out that exactly the same argument yields exactly the same conclusion for the ordinary multiplication algebra in place of the Lie multiplication algebra $\mathfrak{L}(A)$.

1.11. Affine group schemes. Writing **grp** for the category of groups, we let \mathbf{G} be an affine group scheme over k [7, II, §1, n° 1], so $\mathbf{G} : k\text{-alg} \rightarrow \mathbf{grp}$ is a functor represented by some commutative associative k -algebra with 1. We write $\text{Lie}(\mathbf{G})$ for its Lie algebra [7, II, §4, 4.8] and recall from loc. cit. that, if \mathbf{G} is smooth [7, I, §4, n° 4], $\text{Lie}(\mathbf{G})$ commutes with base change, so $\text{Lie}(\mathbf{G})_R \cong \text{Lie}(\mathbf{G}_R)$ canonically, for all $R \in k\text{-alg}$.

In this paper, we will be interested in the following special case. Assume A is finitely generated projective as a k -module and consider its automorphism group scheme by defining

$$\mathbf{Aut}(A) : k\text{-alg} \longrightarrow \mathbf{grp}, \quad R \longmapsto \mathbf{Aut}(A)(R) := \text{Aut}(A_R).$$

Then its Lie algebra is $\text{Der}(A)$ [7, II, §4, 2.3], so assuming that $\mathbf{Aut}(A)$ is smooth forces $\text{Der}(A)$ to commute with base change: $\text{Der}(A)_R = \text{Der}(A_R)$ for all $R \in k\text{-alg}$.

1.12. Corollary. *If A is finitely generated and projective as a k -module and $\mathbf{Aut}(A)$ is smooth as an affine group scheme, then $\text{LMDer}(A)$ commutes with flat base change: $\text{LMDer}(A)_R = \text{LMDer}(A_R)$ for all flat k -algebras $R \in k\text{-alg}$.*

Proof. This follows immediately from (1-7), Lemma 1.7, Cor. 1.10 and 1.11. \square

1.13. Nucleus and centre. We close this section by reminding the reader of the *nucleus* of A , which is defined by

$$\text{Nuc}(A) := \{x \in A \mid [x, A, A] = [A, x, A] = [A, A, x] = \{0\}\}. \tag{1-9}$$

It is an associative subalgebra of A and even a unital one if A contains an identity element. By (1-4), the nucleus is stable under derivations, i.e.,

$$\text{Der}(A) \text{Nuc}(A) \subseteq \text{Nuc}(A).$$

Recall also that the *centre* of A , denoted by $\text{Cent}(A)$, consists of those elements x in the nucleus satisfying $[A, x] = 0$. It is a commutative associative subalgebra of A but may collapse to zero unless A is unital and not zero.

1.14. Proposition. *If A is finitely generated as a k -module, then its nucleus and its centre both commute with flat base change: $\text{Nuc}(A)_R = \text{Nuc}(A_R)$, $\text{Cent}(A)_R = \text{Cent}(A_R)$ for all flat k -algebras $R \in k\text{-alg}$.*

Proof. Assume that the elements a_1, \dots, a_m span A as a k -module and, for $1 \leq i, j \leq m$, consider the linear maps $L_{ij}, M_{ij}, R_{ij}, C_i : A \rightarrow A$ defined respectively

by $x \mapsto [a_i, a_j, x], [a_i, x, a_j], [x, a_i, a_j], [a_i, x]$. Intersecting the kernels of the L_{ij}, M_{ij}, R_{ij} gives the nucleus of A , whose intersection with the kernels of the C_i in turn gives the centre of A . Hence the assertion follows from Lemma 1.7(a),(c). \square

2. Alternative algebras: inner derivations

We now specialize A to a (possibly non-unital) alternative algebra over k .

2.1. Some useful identities. A is alternative if and only if the associator $[x, y, z] = (xy)z - x(yz)$ is an alternating function of its arguments. Hence an element $x \in A$ belongs to the nucleus if and only if one of the relations defining the nucleus (1-9) is fulfilled, and we have the left and right alternative laws

$$x(xy) = x^2y, \tag{2-1}$$

$$(yx)x = yx^2 \tag{2-2}$$

as well as flexibility

$$x(yx) = (xy)x =: xyx, \tag{2-3}$$

for all $x, y \in A$. We also recall the left, middle and right Moufang identities

$$x(y(xz)) = (xyx)z, \quad (xy)(zx) = x(yz)x, \quad ((zx)y)x = z(xy)x \tag{2-4}$$

for all $x, y, z \in A$.

We now derive a number of identities that will play an important role in the explicit description of Lie multiplication derivations. The following relations hold for all $a, b, c, x, y \in A$:

$$[L_a, L_b] = L_{[a,b]} - 2[L_a, R_b], \tag{2-5}$$

$$[R_a, R_b] = -R_{[a,b]} - 2[L_a, R_b], \tag{2-6}$$

$$[[L_a, R_b], L_c] = L_{[a,b,c]} - [L_{[a,b]}, R_c], \tag{2-7}$$

$$[[L_a, R_b], R_c] = R_{[a,b,c]} - [L_c, R_{[a,b]}], \tag{2-8}$$

$$L_a(xy) = ((L_a + R_a)x)y - x(L_a y), \tag{2-9}$$

$$R_a(xy) = -(R_a x)y + x((L_a + R_a)y), \tag{2-10}$$

$$(L_a - R_a)(xy) = ((L_a - R_a)x)y + x((L_a - R_a)y) + [x, 3a, y], \tag{2-11}$$

$$[L_a, R_b]x = [a, b, x] = (ab)x - a(bx) = b(ax) - (ba)x = x(ba) - (xb)a, \tag{2-12}$$

$$[L_a, R_b](xy) = ([L_a, R_b]x)y + x([L_a, R_b]y) + [x, [a, b], y]. \tag{2-13}$$

Proof. Identities (2-5), (2-6) may be found in Schafer [20, (3.68), (3.67)]. While his proof is carried out over fields, it works equally well over the commutative ring k . Ignoring (2-7), (2-8) for the moment, (2-9) (resp. (2-10)) follows immediately by linearizing (2-1) (resp. (2-2)). Subtracting (2-10) from (2-9) yields (2-11). To establish (2-12), one simply observes $[L_a, R_b]x = -[a, x, b] = [a, b, x] = -[b, a, x] = -[x, b, a]$ by alternativity. (2-13) is slightly more troublesome. By (2-12), the left-hand side may be written as

$$[L_a, R_b](xy) = (ab)(xy) - a(b(xy)). \tag{2-14}$$

Linearizing (2-1), the first term on the right becomes

$$(ab)(xy) = ((ab)x + x(ab))y - x((ab)y). \tag{2-15}$$

To the second term, we apply the linearized left Moufang identity and obtain

$$a(b(xy)) = ((ab)x)y + ((xb)a)y - x(b(ay)). \tag{2-16}$$

Subtracting (2-16) from (2-15) and observing (2-14), (2-12) implies

$$\begin{aligned} [L_a, R_b](xy) &= (x(ab) - (xb)a)y - x((ab)y - b(ay)) \\ &= (x(ba) - (xb)a)y + (x[a, b])y + x(b(ay) - (ba)y) - x([a, b]y) \\ &= ([L_a, R_b]x)y + x([L_a, R_b]y) + [x, [a, b], y], \end{aligned}$$

which is (2-13). To prove (2-7), we set $x = c$ in (2-13), view the result as a linear map in y and observe $[c, [a, b], y] = -[[a, b], c, y] = -[L_{[a,b]}, R_c]y$ by (2-12). Finally, (2-8) follows by reading (2-7) in the opposite algebra of A . □

2.2. Proposition (McCrimmon [15, A5, 2.15]). *The Lie multiplication algebra of A is*

$$\mathfrak{L}(A) = L_A + R_A + [L_A, R_A].$$

Proof. By (2-5)–(2-8) above, it suffices to show $[[L_A, R_A], [L_A, R_A]] \subseteq L_A + R_A + [L_A, R_A]$, which follows from the Jacobi identity by applying (2-7) and (2-8) twice. □

Remark. Prop. 2.2 is due to Schafer [19, Thm. 5] if k is a field of characteristic not 2.

2.3. Theorem (cf. McCrimmon [15, A5, 2.16]). *D is a Lie multiplication derivation of A if and only if it has the form*

$$D = L_a - R_a + \sum_{i=1}^m [L_{a_i}, R_{b_i}] \tag{2-17}$$

for some $m \in \mathbb{N}$, $a, a_i, b_i \in A$ ($1 \leq i \leq m$) satisfying

$$3a + \sum_{i=1}^m [a_i, b_i] \in \text{Nuc}(A). \tag{2-18}$$

Proof. If D has the form (2-17), then

$$D(xy) - (Dx)y - x(Dy) = \left[x, 3a + \sum_{i=1}^m [a_i, b_i], y \right]$$

for all $x, y \in A$ by (2-11), (2-13), so (2-18) is equivalent to D being a derivation of A .

Now suppose D satisfies (2-17), (2-18). Then $D \in \text{Der}(A)$ and $\hat{D} = \hat{L}_a - \hat{R}_a + \sum [\hat{L}_{a_i}, \hat{R}_{b_i}]$ since the right-hand side kills 1. This implies $\hat{D} \in \mathcal{L}(\hat{A})$, hence $D \in \text{LMDer}(A)$ by (1-7). Conversely, let D be a Lie multiplication derivation of A . By Prop. 2.2,

$$\hat{D} = \hat{L}_{\hat{a}} + \hat{R}_{\hat{b}} + \sum_{i=1}^m [\hat{L}_{\hat{a}_i}, \hat{R}_{\hat{b}_i}]$$

for some $m \in \mathbb{N}$, $\hat{a}, \hat{a}_i, \hat{b}_i \in \hat{A}$ ($1 \leq i \leq m$), where $\hat{D}1 = 0$ implies $\hat{b} = -\hat{a}$, hence (2-17) with a, a_i, b_i being the A -components of $\hat{a}, \hat{a}_i, \hat{b}_i$, respectively. But then (2-18) drops out automatically since D was assumed to be a derivation. \square

When it comes to applications of Thm. 2.3, the following more concise description of Lie multiplication derivations turns out to be useful.

2.4. Derivations and exterior powers. We introduce the notation

$$W(A) := A \oplus \bigwedge^2 A.$$

Our description will be based on two linear maps defined on $W(A)$. The first one is

$$s = s_A: W(A) \longrightarrow A, \quad s(a \oplus (b \wedge c)) := 3a + [b, c] \quad (a, b, c \in A).$$

To define the second one, we note that the flexible law (2-3) makes the bilinear expression $[L_a, R_b]$ alternating in $a, b \in A$ and thus leads to a linear map

$$\Delta = \Delta_A: W(A) \longrightarrow \mathfrak{gl}(A), \quad \Delta_{a \oplus (b \wedge c)} := \Delta(a \oplus (b \wedge c)) := L_a - R_a + [L_b, R_c]$$

for $a, b, c \in A$. With these notations, Thm. 2.3 implies

$$\text{LMDer}(A) = \{ \Delta_x \mid x \in W(A), s(x) \in \text{Nuc}(A) \}. \tag{2-19}$$

Now observe that every $g \in \mathfrak{gl}(A)$ induces a linear map

$$g^\dagger: W(A) \longrightarrow W(A), \quad g^\dagger(a \oplus (b \wedge c)) = g(a) \oplus (g(b) \wedge c + b \wedge g(c)) \tag{2-20}$$

for $a, b, c \in A$. Clearly, the assignment $g \mapsto g^\dagger$ determines an embedding $\mathfrak{gl}(A) \rightarrow \mathfrak{gl}(W(A))$ of Lie algebras, and (1-1), (1-2), (2-20) are easily seen to imply

$$[D, \Delta_x] = \Delta_{D^\dagger x}, \quad [D^\dagger, \Delta_x^\dagger] = \Delta_{D^\dagger x}^\dagger \tag{2-21}$$

for $D \in \text{Der}(A)$, $x \in W(A)$, while (1-3) and (2-20) imply that the diagram

$$\begin{array}{ccc} W(A) & \xrightarrow{D^\dagger} & W(A) \\ s \downarrow & & \downarrow s \\ A & \xrightarrow{D} & A \end{array} \tag{2-22}$$

commutes.

Remark. Lie multiplication derivations of alternative algebras as described in Thm. 2.3 do not in general satisfy the Mapping Principle of the introduction since a homomorphism $A \rightarrow B$ may not map the nucleus of A into the nucleus of B , so if (2-18) holds for elements $a, a_i, b_i \in A$, it may no longer do so for their images in B . For this reason, we will introduce *inner* derivations of alternative algebras as a special type of Lie multiplication derivations where such unpleasantness can be ruled out. McCrimmon [15, A5.2, p. 24] gets around this difficulty in a slightly different manner, by means of his notion of *strictly inner* derivations.

2.5. Classes of inner derivations. With the terminology of 2.4, the elements of

$$\text{InDer}_{\text{alt}}(A) := \{\Delta_x \mid x \in W(A), s(x) = 0\} \tag{2-23}$$

are called *inner derivations* of A . Thus the difference to (2-19) is that $s(x)$ is required to be zero instead of in the nucleus. In more explicit terms, the inner derivations of A are precisely the linear maps

$$L_a - R_a + \sum_{i=1}^m [L_{a_i}, R_{b_i}], \tag{2-24}$$

where $m \in \mathbb{N}$ and $a, a_i, b_i \in A$ ($1 \leq i \leq m$) satisfy the relation

$$3a + \sum_{i=1}^m [a_i, b_i] = 0. \tag{2-25}$$

Inner derivations obviously satisfy the Mapping Principle. Adapting the terminology of McCrimmon [15] to the present set-up, and identifying A and $\bigwedge^2 A$

canonically with submodules of $W(A) = A \oplus \bigwedge^2 A$ throughout the rest of the paper, we now introduce the following three classes of inner derivations.

(a) *Associator derivations.* These are the elements of

$$\text{AssDer}(A) := \{ \Delta_u \mid u \in \bigwedge^2 A, s(u) = 0 \} \subseteq \text{InDer}_{\text{alt}}(A). \tag{2-26}$$

They have the form $\sum [L_{a_i}, R_{b_i}]$, where $a_i, b_i \in A$ satisfy $\sum [a_i, b_i] = 0$, so by (2-12), they act on $x \in A$ as $x \mapsto \sum [a_i, b_i, x]$, i.e., as a sum of associators, hence the name. In particular, $\text{AssDer}(A) = \{0\}$ if A is associative. On the other hand, as we shall see in Thm. 5.1 below, associator derivations play an important role in octonion algebras.

(b) *Standard derivations.* These are the elements of

$$\text{StanDer}(A) := \{ \Delta_{s(u) \oplus (-3u)} \mid u \in \bigwedge^2 A \} \subseteq \text{InDer}_{\text{alt}}(A). \tag{2-27}$$

As a k -module, $\text{StanDer}(A)$ is spanned by the elements

$$D_{a,b} = L_{[a,b]} - R_{[a,b]} - 3[L_a, R_b] = [L_a, L_b] + [L_a, R_b] + [R_a, R_b] \tag{2-28}$$

for $a, b \in A$, the last equation being a consequence of (2-5), (2-6). Standard derivations have the advantage of being parametrized by the full k -module $\bigwedge^2 A$, with no further constraints on the parameters involved. On the other hand, $\text{StanDer}(A) = \{0\}$ if A is commutative, since this is well known to imply $3[A, A, A] = \{0\}$, hence $D_{a,b} = 0$ for all $a, b \in A$ by (2-12) and (2-28); cf. Prop. 2.7 below for a more precise statement.

(c) *Commutator derivations.* These are the elements of

$$\text{ComDer}(A) := \{ \Delta_a \mid a \in A, 3a = 0 \} = \{ L_a - R_a \mid a \in A, 3a = 0 \} \subseteq \text{InDer}_{\text{alt}}(A).$$

They appear only in the presence of 3-torsion. Note $\Delta_a x = [a, x]$ for $a, x \in A$, justifying the chosen terminology.

2.6. Proposition. *In the terminology of 2.5,*

$$3\Delta_x = \Delta_{s(x)} + \Delta_{s(-u) \oplus 3u} \tag{2-29}$$

for all $x = a \oplus u \in W(A)$ satisfying $s(x) \in \text{Nuc}(A)$. In particular,

$$3\text{LMDer}(A) \subseteq \{ L_a - R_a \mid a \in \text{Nuc}(A) \} + \text{StanDer}(A) \subseteq \text{LMDer}(A). \tag{2-30}$$

Proof. (2-29) is obvious by (2-19) and the definition of s ; it immediately implies (2-30) since the first summand on the right of (2-29) is the nuclear derivation $L_{a'} - R_{a'}$, $a' \in \text{Nuc}(A)$, while the second one by (2-27) is a standard derivation.

Remark. If $3A = A$, we obtain

$$\text{LMDer}(A) = \{L_a - R_a \mid a \in \text{Nuc}(A)\} + \text{StanDer}(A),$$

hence Schafer's classical description [20, pp. 76–78] of Lie multiplication derivations of unital alternative algebras over fields of characteristic not 2 or 3, see also McCrimmon [15, A5, 2.17].

2.7. Proposition. *In the terminology of 2.5, the following statements hold.*

- (a) $\text{InDer}_{\text{alt}}(A)$, $\text{AssDer}(A)$, $\text{StanDer}(A)$, $\text{ComDer}(A)$ are all ideals in the full derivation algebra of A .
- (b) $\text{AssDer}(A)A \subseteq [A, A, A]$, $3\text{AssDer}(A) \subseteq \text{StanDer}(A)$.
- (c) $(\text{StanDer}(A) + \text{ComDer}(A))A \subseteq [A, A]$.
- (d) If $3[A, A] = [A, A]$, then

$$\text{InDer}_{\text{alt}}(A) = \text{AssDer}(A) + \text{StanDer}(A) + \text{ComDer}(A).$$

- (e) If $3A = A$, then

$$\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A) + \text{ComDer}(A).$$

- (f) If $\frac{1}{3} \in k$, then

$$\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A), \quad \text{ComDer}(A) = \{0\}.$$

- (g) If $3A = \{0\}$, then

$$\text{InDer}_{\text{alt}}(A) = \text{AssDer}(A) + \text{ComDer}(A).$$

Proof. (a) follows immediately from (2-21), (2-22).

(b) The first part has already been observed in 2.5(a), while the second one follows from (2-29) in the special case $a = s(u) = 0$.

(c) The relation $\text{ComDer}(A)A \subseteq [A, A]$ is obvious. Since (2-28) and (2-12) imply $\text{StanDer}(A)A \subseteq [A, A] + 3[A, A, A]$, it remains to show $3[A, A, A] \subseteq [A, A]$ which is probably known; we include a proof for convenience. Modulo $[A, A]$ we have by (2-11) that $3[x, a, y] \equiv [x, a]y + x[y, a] \equiv y[x, a] + x[y, a]$, which is $\equiv 0$, being the bilinearization of $x[x, a] = x(xa) - x(ax) = x(xa) - (xa)x$ (by (2-3)) $= [x, xa] \in [A, A]$.

(d), (e) Suppose $x = a \oplus u \in W(A)$ satisfies $s(x) = 0$ (cf. (2-23)). Since $s(\bigwedge^2 A) = [A, A]$, the hypothesis in (d) leads to an element $w \in \bigwedge^2 A$ such that $s(u) = 3s(w)$, so $v := u - 3w$ satisfies $s(v) = 0$. On the other hand, the hypothesis in (e) leads to an element $w \in \bigwedge^2 A$ such that $u = 3w$, so again $s(u) = 3s(w)$, but

this time even $v := u - 3w = 0$. In any event, setting $b = a + s(w)$, we conclude $3b = 3a + 3s(w) = 3a + s(u) = s(x) = 0$. Moreover,

$$\Delta_x = \Delta_v + \Delta_{(-s(w)) \oplus 3w} + \Delta_b \in \text{AssDer}(A) + \text{StanDer}(A) + \text{ComDer}(A).$$

Hence (d) and (e) hold. Now (f) follows immediately from (e) since $\text{ComDer}(A) = \{0\}$ in the absence of 3-torsion, and (g) is a consequence of (2-23). \square

2.8. Example. Let A be an associative k -algebra. Using (2-24), (2-25), one checks easily that

$$\text{InDer}_{\text{alt}}(A) \subseteq \text{InDer}_{\text{ass}}(A).$$

However, equality does not hold in general. To see this, suppose k contains $\frac{1}{3}$. Then $\text{InDer}_{\text{alt}}(A) = \text{StanDer}(A)$ by Prop. 2.7, while (2-28) reduces to $D_{a,b} = L_{[a,b]} - R_{[a,b]}$ for all $a, b \in A$. Hence we obtain

$$L_x - R_x \in \text{InDer}_{\text{ass}}(A) \setminus \text{InDer}_{\text{alt}}(A) \tag{2-31}$$

for any $x \in A$ that does not belong to $Z + [A, A]$, Z being the centre of A . More specifically, let k be a field of characteristic $p > 0$, $p \neq 3$, and put $A = \text{Mat}_p(k)$, the algebra of $p \times p$ matrices with entries in k . Then $[A, A]$, being the kernel of the trace, contains $Z = k \cdot 1_A$, so any $x \in A$ with non-zero trace will satisfy (2-31).

In the presence of the alternative law, we can improve and expand Cor. 1.12 considerably.

2.9. Proposition. *If A is finitely generated as a k -module, then the Lie algebras $\text{LMDer}(A)$, $\text{InDer}_{\text{alt}}(A)$, $\text{AssDer}(A)$, $\text{StanDer}(A)$, $\text{ComDer}(A)$ all commute with flat base change: For all flat k -algebras $R \in k\text{-alg}$, we have*

$$\text{LMDer}(A)_R = \text{LMDer}(A_R), \tag{2-32}$$

$$\text{InDer}_{\text{alt}}(A)_R = \text{InDer}(A_R), \tag{2-33}$$

$$\text{AssDer}(A)_R = \text{AssDer}(A_R), \tag{2-34}$$

$$\text{StanDer}(A)_R = \text{StanDer}(A_R), \tag{2-35}$$

$$\text{ComDer}(A)_R = \text{ComDer}(A_R). \tag{2-36}$$

Proof. Since taking exterior powers commutes with flat (even arbitrary) base change [4, III, §7, Prop. 8], so do the linear maps Δ_A and s_A . Furthermore, (2-19),(2-23) yield

$$\text{LMDer}(A) = \Delta_A \left(s_A^{-1}(\text{Nuc}(A)) \right), \quad \text{InDer}_{\text{alt}}(A) = \Delta_A(\text{Ker}(s_A)).$$

Hence (2-32),(2-33) follow from Lemma 1.7(a),(b) and Prop. 1.14, while an analogous argument yields (2-36). After the identifications of 1.9, we obtain $(D_{a,b})_R =$

D_{a_R, b_R} for all $a, b \in A$, and (2-35) follows from Lemma 1.7(e). It remains to prove (2-34). To do so, we put $K_A := \text{Ker}(s_A) \cap \wedge^2 A$ (which commutes with flat base change since s_A does), $T_A = \text{Ker}(\Delta_A)$, note that Δ by (2-26) restricts to a linear surjection $\Delta_0: K_A \rightarrow \text{AssDer}(A)$, and obtain a commutative diagram

$$\begin{array}{ccccccc}
 & & 0 & & 0 & & \\
 & & \downarrow & & \downarrow & & \\
 0 & \longrightarrow & T_A & \longrightarrow & K_A & \xrightarrow{\Delta_0} & \text{AssDer}(A) \longrightarrow 0 \\
 & & \downarrow & & \downarrow & & \downarrow \\
 & & \wedge^2 A & \xrightarrow{\Delta} & \text{End}_k(A) & & \\
 & & \downarrow s & & & & \\
 & & [A, A] & & & & \\
 & & \downarrow & & & & \\
 & & 0 & & & &
 \end{array}$$

with exact rows and columns. Tensoring with R , we end up with (2-34). □

3. Making inner derivations functorial

In this section, we address the lack of functoriality of the inner derivation algebra mentioned in the introduction. Before delving into the general categorical setup, it may be helpful to consider the following example.

3.1. Associative algebras. Every associative algebra A defines a Lie algebra A^- having underlying k -module A and Lie product $[x, y] = xy - yx$, and A^- depends functorially on A : every homomorphism $f: A \rightarrow B$ induces a Lie algebra homomorphism $f^- = f: A^- \rightarrow B^-$. The inner derivations of A are the derivations $D_x = L_x - R_x$ and the map sending $x \in A^-$ to $D_x \in \text{Der}(A)$ is a Lie algebra homomorphism $\alpha: A^- \rightarrow \text{Der}(A)$, i.e., an action of A^- on A by derivations, whose kernel is central in (indeed, equals the centre of) A^- . Moreover, the formula

$$f(D_x(a)) = D_{f^-(x)}(f(a)) \tag{3-1}$$

holds for all $x \in A^-$, $a \in A$, so the derivation $D_{f^-(x)}$ of B is f -related to the derivation D_x of A . But note $D_x = 0$ for central $x \in A$, yet $D_{f^-(x)}$ need not be zero if $f^-(x)$ is not central in B , showing why $D_x \mapsto D_{f^-(x)}$ is in general not a well defined map from $\text{InDer}_{\text{ass}}(A)$ to $\text{InDer}_{\text{ass}}(B)$.

3.2. Categories of algebras. Abstracting from the previous example, we replace associative algebras by an arbitrary category of non-associative algebras. To discuss questions of base change, it is convenient to consider not only algebras over a fixed base ring, but over all possible commutative rings, more generally, over a subcategory \mathcal{R} of all commutative rings. Thus let \mathcal{A} be a category of *algebras over \mathcal{R}* in the following sense: objects of \mathcal{A} are pairs (k, A) where $k \in \mathcal{R}$ and A is a non-associative k -algebra. Morphisms $(k, A) \rightarrow (l, B)$ of \mathcal{A} are pairs (τ, f) where $\tau: k \rightarrow l$ is a morphism of \mathcal{R} and $f: A \rightarrow B$ is a τ -semilinear map from the k -module A to the l -module B preserving the algebraic structure. Examples of this situation abound: associative, alternative or Jordan algebras, unital or not, Lie algebras, etc.

The projection onto the first component $\Pi: \mathcal{A} \rightarrow \mathcal{R}$ is a functor, and we denote by \mathcal{A}_k the fibre of Π over $k \in \mathcal{R}$, that is, the subcategory of \mathcal{A} with objects algebras over the fixed ring k , and morphisms k -linear maps, thus of the form $(\text{Id}_k, f): (k, A) \rightarrow (k, B)$. It is often inconvenient to indicate explicitly the base ring k of an object of \mathcal{A} . Thus we frequently write simply $A \in \mathcal{A}_k$ or even $A \in \mathcal{A}$ and $k = \Pi(A)$ instead of $(k, A) \in \mathcal{A}$, or employ a phrase like “let A be a k -algebra in \mathcal{A} ”. Similarly, a morphism of \mathcal{A} will often be written as $f: A \rightarrow B$, with $\Pi(f) = \tau: \Pi(A) = k \rightarrow \Pi(B) = l$ the corresponding homomorphism of the respective base rings.

3.3. The derivation category. Let \mathcal{A} be a category of (non-associative) algebras over \mathcal{R} as before and let Lie be the category of Lie algebras over \mathcal{R} . We define a category $\text{Der}(\mathcal{A})$ over \mathcal{A} as follows.

A *derivation action* of a Lie algebra $\mathfrak{g} \in \text{Lie}_k$ on an algebra $A \in \mathcal{A}_k$ is a homomorphism $\alpha: \mathfrak{g} \rightarrow \text{Der}(A)$ of k -Lie algebras. We write a derivation action as a quadruple $(k, A, \mathfrak{g}, \alpha)$ or simply as (A, α) . (Since A is not uniquely determined by $\text{Der}(A)$, the algebra A must be explicitly indicated. On the other hand, \mathfrak{g} is the domain of definition of α , so a derivation action is determined by A and α). Now construct a category $\text{Der}(\mathcal{A})$, called the *derivation category of \mathcal{A}* , whose objects are the derivation actions, and whose morphisms are defined as follows.

Let $f: A \rightarrow B$ be a morphism of \mathcal{A} and $\tau = \Pi(f): k \rightarrow l$ the corresponding morphism of \mathcal{R} . An *f -derivation from A to B* is a τ -semilinear map $d: A \rightarrow B$ such that

$$d(aa') = d(a)f(a') + f(a)d(a'),$$

for all $a, a' \in A$. Denote the set of f -derivations from A to B by $\text{Der}_f(A, B)$. This is in a natural way an l -module by defining $(sd)(a) = sd(a)$ for $a \in A$ and $s \in l$. Given derivations $D \in \text{Der}(A)$ and $D' \in \text{Der}(B)$, the maps $f_*(D) := f \circ D$ and $f^*(D') := D' \circ f$ both belong to $\text{Der}_f(A, B)$.

A *morphism* from (A, α) to (B, β) is now defined as a pair (f, φ) where $f: A \rightarrow B$ and $\varphi: \mathfrak{g} \rightarrow \mathfrak{h}$ are morphisms of \mathcal{A} and Lie , respectively, satisfying $\Pi(f) = \Pi(\varphi)$, and making the following diagram commutative:

$$\begin{array}{ccc}
 \mathfrak{g} & \xrightarrow{\varphi} & \mathfrak{h} \\
 \alpha \downarrow & & \downarrow \beta \\
 \text{Der}(A) & \xrightarrow{f_*} \text{Der}_f(A, B) \xleftarrow{f^*} & \text{Der}(B)
 \end{array}$$

Explicitly, this means:

$$f(\alpha(X) \cdot a) = \beta(\varphi(X)) \cdot f(a),$$

for all $X \in \mathfrak{g}$ and all $a \in A$. There is again a functor $\mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{R}$ given by $(A, \alpha) \mapsto \Pi(A)$ and $(f, \varphi) \mapsto \Pi(f)$. Moreover, the projections $(A, \alpha) \mapsto A$ and $(f, \varphi) \mapsto f$ define a functor $P_1: \mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{A}$. Similarly, the projections $(A, \alpha) \mapsto \mathfrak{g}$ and $(f, \varphi) \mapsto \varphi$ define a functor $P_2: \mathcal{D}er(\mathcal{A}) \rightarrow Lie$.

3.4. Derivation functors. Let \mathcal{A} be a category of algebras over \mathcal{R} as before. A *derivation functor* is a functor $F: \mathcal{A} \rightarrow \mathcal{D}er(\mathcal{A})$ commuting with the projections onto \mathcal{R} which is a section of the projection $P_1: \mathcal{D}er(\mathcal{A}) \rightarrow \mathcal{A}$ in the sense that $P_1 \circ F = Id_{\mathcal{A}}$.

In more detail, this means the following: for every $A \in \mathcal{A}_k$, we have a derivation action ρ_A of a k -Lie algebra $\mathfrak{d}(A)$ on A , and for every homomorphism $f: A \rightarrow B$ of algebras in \mathcal{A} we have a morphism $\mathfrak{d}(f): \mathfrak{d}(A) \rightarrow \mathfrak{d}(B)$ of Lie algebras, semilinear with respect to $\Pi(f)$, and compatible with the actions in the sense that the diagram

$$\begin{array}{ccc}
 \mathfrak{d}(A) & \xrightarrow{\mathfrak{d}(f)} & \mathfrak{d}(B) \\
 \rho_A \downarrow & & \downarrow \rho_B \\
 \text{Der}(A) & \xrightarrow{f_*} \text{Der}_f(A, B) \xleftarrow{f^*} & \text{Der}(B)
 \end{array} \tag{3-2}$$

is commutative. As before, this means

$$f(\rho_A(X) \cdot a) = \rho_B(\mathfrak{d}(f)(X)) \cdot f(a) \quad (X \in \mathfrak{d}(A), a \in A).$$

In particular, $\mathfrak{d}: \mathcal{A} \rightarrow Lie$ is a functor from \mathcal{A} to Lie . It is tempting to say that ρ is a natural transformation from \mathfrak{d} to the “functor” Der , but $\text{Der}(A)$ does not depend functorially on A . As a substitute, for every morphism $f: A \rightarrow B$ in \mathcal{A} , commutativity of (3-2) is equivalent to $(f, \mathfrak{d}(f)): (A, \rho_A) \rightarrow (B, \rho_B)$ being a morphism of $\mathcal{D}er(A)$. By abuse of notation, we will often write $F = (\mathfrak{d}, \rho)$ for a derivation functor.

Suppose $F = (\mathfrak{d}, \rho)$ and $F' = (\mathfrak{d}', \rho')$ are derivation functors. A *morphism* from F to F' is a natural transformation of functors. This amounts to Lie algebra homomorphisms $h_A: \mathfrak{d}(A) \rightarrow \mathfrak{d}'(A)$ for all $A \in \mathcal{A}$ defining a natural transformation $\mathfrak{d} \rightarrow \mathfrak{d}'$ and making the diagrams

$$\begin{array}{ccc}
 \mathfrak{d}(A) & \xrightarrow{h_A} & \mathfrak{d}'(A) \\
 \searrow \rho_A & & \swarrow \rho'_A \\
 & \text{Der}(A) &
 \end{array}$$

commutative.

A derivation functor F is called *inner* if ρ_A maps $\mathfrak{d}(A)$ into the Lie multiplication algebra of A , and *central* if the kernel of ρ_A is central in $\mathfrak{d}(A)$, for all $A \in \mathcal{A}$.

In the special case where $f = g \in \text{Aut}(A)$, the functoriality of F implies that $\text{Aut}(A)$ acts on the Lie algebra $\mathfrak{d}(A)$ by automorphisms and the map ρ_A is equivariant with respect to this action on the one hand, and with respect to conjugation of $\text{Aut}(A)$ on $\text{Der}(A)$ on the other, because (3-2) now says

$$g \circ \rho_A(X) \circ g^{-1} = \rho_A(\mathfrak{d}(g)(X)),$$

for all $g \in \text{Aut}(A)$ and $X \in \mathfrak{d}(A)$.

If we assume that F commutes with flat base change (see 3.10), which holds in all standard examples, then by extending k to the dual numbers $k(\varepsilon)$ and specializing $g = \text{Id} + \varepsilon D$ for $D \in \text{Der}(A)$, we obtain an action of $\text{Der}(A)$ on $\mathfrak{d}(A)$ by derivations and ρ_A is equivariant with respect to this action and the adjoint representation of $\text{Der}(A)$ on itself. In particular, $\mathfrak{d}(A)$ is then an ideal in $\text{Der}(A)$.

3.5. Remarks and examples. (a) The definition of a derivation functor does not tie the Lie algebras $\mathfrak{d}(A)$ very closely to the derivations of A : the kernels of ρ_A can be arbitrarily big. For example, let \mathfrak{l} be a fixed Lie algebra over \mathbb{Z} and define F by $\mathfrak{d}(A) = \mathfrak{l} \otimes_{\mathbb{Z}} k$ and $\rho_A = 0$ for all $A \in \mathcal{A}_k$ and $k \in \mathcal{R}$. The requirement that F be central cuts down the kernels to some extent; it is satisfied in all the examples treated below. The same is true of the condition of innerness.

(b) The example of associative algebras treated in 3.1 yields a derivation functor F with $\mathfrak{d}(A) = A^-$ and $\rho_A(x) = D_x$. The commutativity of (3-2) is formula (3-1), and F is inner and central.

(c) Let $\mathcal{A} = \mathcal{L}ie$. Then a natural choice of F is $\mathfrak{d} = \text{Id}_{\mathcal{L}ie}$ and ρ the adjoint representation. Again F is inner and central.

We show next that the classes of inner derivations of alternative algebras introduced in 2.5 come from derivation functors as well. The construction rests on the

following simple lemma which is essentially contained in [13, 2.1]. We include a proof for the convenience of the reader.

3.6. Lemma. *Let \mathfrak{g} be a k -Lie algebra, M a k - \mathfrak{g} -module, and let $\phi: M \rightarrow \mathfrak{g}$ be a homomorphism of left k - \mathfrak{g} -modules, where \mathfrak{g} acts on itself by the adjoint representation. Write the action of an element $x \in \mathfrak{g}$ on $u \in M$ as $x_M \cdot u$, and define a non-associative product $\{u, v\}$ on M by*

$$\{u, v\} := \phi(u)_M \cdot v.$$

(a) *The map $\phi: M \rightarrow \mathfrak{g}$ is a homomorphism of non-associative algebras:*

$$\phi(\{u, v\}) = [\phi(u), \phi(v)]. \tag{3-3}$$

(b) *\mathfrak{g} acts by derivations of the product $\{-, -\}$. The Jacobi identity holds in the following form on M :*

$$\{u, \{v, w\}\} - \{v, \{u, w\}\} = \{\{u, v\}, w\}, \tag{3-4}$$

so M is a left Leibniz algebra [13].

(c) *Let Q be the k -linear span of all squares $\{u, u\}$, $u \in M$, and let $Z = \text{Ker}(\phi) \subseteq M$. Then Q and Z are stable under the action of \mathfrak{g} , and*

$$Q \subseteq Z, \quad \{Z, M\} = 0, \quad \{M, Z\} \subseteq Q. \tag{3-5}$$

(d) *The product $\{-, -\}$ induces a Lie algebra structure on $\mathfrak{h} := M/Q$ and ϕ induces a Lie algebra homomorphism $\bar{\phi}: \mathfrak{h} \rightarrow \mathfrak{g}$ whose kernel Z/Q is central in \mathfrak{h} . The action of \mathfrak{g} on M induces an action of \mathfrak{g} on \mathfrak{h} by derivations.*

Proof. (a) Put $x = \phi(u) \in \mathfrak{g}$. Since ϕ is a homomorphism of \mathfrak{g} -modules, $\phi(\{u, v\}) = \phi(\phi(u)_M \cdot v) = \phi(x_M \cdot v) = [x, \phi(v)] = [\phi(u), \phi(v)]$.

(b) Let $x \in \mathfrak{g}$ and $v, w \in M$. Then

$$\begin{aligned} x_M \cdot \{v, w\} &= x_M \cdot (\phi(v)_M \cdot w) = [x_M, \phi(v)_M] \cdot w + \phi(v)_M \cdot (x_M \cdot w) \\ &= [x, \phi(v)]_M \cdot w + \phi(v)_M \cdot (x_M \cdot w) = \phi(x_M \cdot v)_M \cdot w + \{v, x_M \cdot w\} \\ &= \{x_M \cdot v, w\} + \{v, x_M \cdot w\}. \end{aligned}$$

Now (3-4) follows by specializing $x = \phi(u)$.

(c) Since ϕ is a homomorphism of \mathfrak{g} -modules, it is clear that Z is stable under \mathfrak{g} , and (b) implies that Q is stable under \mathfrak{g} as well. The inclusion $Q \subseteq Z$ follows from (3-3) and the fact that the Lie product in \mathfrak{g} is alternating. Let $u \in Z$ and $v \in M$. Then $\phi(u) = 0$, hence also $\{u, v\} = \phi(u)_M \cdot v = 0$ which proves $\{Z, M\} = 0$. Moreover, $\{v, u\} = \{u, v\} + \{v, u\} = \{u + u, v + v\} - \{u, u\} - \{v, v\} \in Q$, so $\{M, Z\} \subseteq Q$.

(d) Since $Q \subseteq Z$, (3-5) implies $\{Q, M\} + \{M, Q\} \subset Q$, so Q is an ideal of $\{-, -\}$. As it contains all squares, the product induced on M/Q is alternating. Now (3-4) shows that $\mathfrak{h} = M/Q$ is a Lie algebra, and $\bar{\phi}$ is a Lie algebra homomorphism by (3-3). Finally, $\{Z, M\} = 0$ implies that $\text{Ker}(\bar{\phi}) = Z/Q$ is central in \mathfrak{h} , and it follows from (b) that \mathfrak{g} acts on \mathfrak{h} by derivations. \square

3.7. The inner derivation functor of alternative algebras. Let A be an alternative algebra over k . Recall from 2.4 that $W(A) := A \oplus \bigwedge^2 A$ is a $\mathfrak{gl}(A)$ -module under the action

$$\mathfrak{gl}(A) \times W(A) \longrightarrow W(A), \quad (g, x) \longmapsto g \cdot x := g^\dagger(x).$$

For a τ -semilinear homomorphism $f: A \rightarrow B$ of alternative algebras over k and l , respectively, the map $W(f) := f \oplus \bigwedge^2 f: W(A) \rightarrow W(B)$ is again τ -semilinear, and one checks easily that the relations

$$f \circ s = s \circ W(f), \quad W(f) \circ \Delta_x^\dagger = \Delta_{W(f)(x)}^\dagger \circ W(f), \quad f \circ \Delta_x = \Delta_{W(f)(x)} \circ f \tag{3-6}$$

hold for all $x \in W(A)$.

We consider the k -submodule

$$W_{\text{in}}(A) := \text{Ker}(s) = \{x \in W(A) \mid s(x) = 0\}.$$

of the $\mathfrak{gl}(A)$ -module $W(A)$, which by (2-22) remains stable under $\text{Der}(A)$, hence may be regarded canonically as a $\text{Der}(A)$ -module. Define $\phi: W_{\text{in}}(A) \rightarrow \text{Der}(A)$ by $\phi(x) := \Delta_x$ for $x \in W_{\text{in}}(A)$. Then (2-21) shows that ϕ is a homomorphism of $\text{Der}(A)$ -modules. Applying Lemma 3.6 therefore yields a Leibniz algebra $W_{\text{in}}(A)$, a Lie algebra $\mathfrak{d}_{\text{in}}(A) = W_{\text{in}}(A)/Q$ and a homomorphism $\rho_{\text{in},A} = \bar{\phi}: \mathfrak{d}_{\text{in}}(A) \rightarrow \text{Der}(A)$ whose image, by (2-23), is precisely $\text{InDer}_{\text{alt}}(A)$.

Returning to the τ -semilinear homomorphism $f: A \rightarrow B$, we conclude from (3-6) that $W(f)$ sends $W_{\text{in}}(A)$ to $W_{\text{in}}(B)$ and hence induces a τ -semilinear map $W_{\text{in}}(f): W_{\text{in}}(A) \rightarrow W_{\text{in}}(B)$, which, again by (3-6), is a homomorphism of Leibniz algebras, inducing canonically a Lie algebra homomorphism $\mathfrak{d}_{\text{in}}(f): \mathfrak{d}_{\text{in}}(A) \rightarrow \mathfrak{d}_{\text{in}}(B)$. The commutativity of (3-2) is a consequence of (3-6), so we have defined a derivation functor F_{in} for alternative algebras, the *inner derivation functor*.

3.8. The associator and commutator derivation functors. Consider the submodule $W_{\text{ass}}(A) = W_{\text{in}}(A) \cap \bigwedge^2 A$ of $W_{\text{in}}(A)$, which is in fact a $\text{Der}(A)$ -stable subalgebra of the Leibniz algebra $W_{\text{in}}(A)$, so the construction of 3.7 can be performed mutatis mutandis on $W_{\text{ass}}(A)$ and yields a derivation functor F_{ass} , called the *associator derivation functor* of alternative algebras. The inclusions $W_{\text{ass}}(A) \rightarrow W_{\text{in}}(A)$ induce homomorphisms $\mathfrak{d}_{\text{ass}}(A) \rightarrow \mathfrak{d}_{\text{in}}(A)$ of Lie algebras (in general no longer

injective) which are compatible with the representations ρ and ρ_{ass} and with morphisms of \mathcal{A} . Thus we have a natural transformation $F_{\text{ass}} \rightarrow F_{\text{in}}$ of derivation functors.

Similarly, let $W_{\text{com}}(A) = W_{\text{in}}(A) \cap A = {}_3A$, the 3-torsion elements of A . As before, $W_{\text{com}}(A)$ is a subalgebra of $W_{\text{in}}(A)$ and in fact is already a Lie algebra, because $\{a, a\} = \Delta_a \cdot a = [a, a] = 0$. We obtain a derivation functor $F_{\text{com}} = (\mathfrak{d}_{\text{com}}, \rho_{\text{com}})$, the *commutator derivation functor*, where $\mathfrak{d}_{\text{com}}(A) = {}_3A$ with Lie bracket $[x, y] = xy - yx$, and $\rho_{\text{com}}: {}_3A \rightarrow \text{Der}(A)$ given by the commutator: $x \mapsto (a \mapsto [x, a])$. Again, there is a natural transformation $F_{\text{com}} \rightarrow F_{\text{in}}$ induced from the inclusions $W_{\text{com}}(A) \rightarrow W(A)$.

3.9. Standard derivation functors. By definition (2.5(b)), the standard derivations of an alternative algebra A are of the form $\Delta_{s(u) \oplus (-3u)}$, $u \in \bigwedge^2 A$. Thus they can be parametrized by all of $\bigwedge^2 A$ or by the image in $W(A)$ of $\bigwedge^2 A$ under the map $\zeta: u \mapsto s(u) \oplus (-3u)$. This gives rise to two standard derivation functors as follows.

First let $M := \bigwedge^2 A$ and define $\phi: M \rightarrow \text{Der}(A)$ by $\phi(a \wedge b) = [L_a, L_b] + [L_a, R_b] + [R_a, R_b]$ as in (2-28). It follows from the formulas in 2.4 that ϕ is equivariant with respect to the action of $\text{Der}(A)$ on M and on itself by the adjoint representation. Hence Lemma 3.6 yields a Lie algebra $\mathfrak{d}_{\text{st}}(A) = M/Q$ and a homomorphism $\rho_{\text{st}}(A): \mathfrak{d}_{\text{st}}(A) \rightarrow \text{Der}(A)$ with central kernel and image $\text{StanDer}(A)$. Also, for $f: A \rightarrow B$ a homomorphism of alternative algebras, we have

$$f \circ \phi(u) = \phi\left(\bigwedge^2 f(u)\right) \circ f.$$

Hence $\bigwedge^2 f$ is a homomorphism of Leibniz algebras, and induces a homomorphism $\mathfrak{d}_{\text{st}}(f): \mathfrak{d}_{\text{st}}(A) \rightarrow \mathfrak{d}_{\text{st}}(B)$, compatible with the representations ρ_{st} . This defines a derivation functor $F_{\text{st}} = (\mathfrak{d}_{\text{st}}, \rho_{\text{st}})$, the *standard derivation functor* of alternative algebras.

Next, imitating the procedure of 3.8, we have a derivation functor induced from

$$W_{\text{st}}(A) = \{\zeta(u) \mid u \in \bigwedge^2 A\} \subseteq W_{\text{in}}(A),$$

denoted $\bar{F}_{\text{st}} = (\bar{\mathfrak{d}}_{\text{st}}, \bar{\rho}_{\text{st}})$. As before, the inclusions $W_{\text{st}}(A) \rightarrow W_{\text{in}}(A)$ induce a morphism $\bar{F}_{\text{st}} \rightarrow F_{\text{in}}$. The map ζ induces a morphism $h: F_{\text{st}} \rightarrow \bar{F}_{\text{st}}$ with the property that $h_A: \mathfrak{d}_{\text{st}}(A) \rightarrow \bar{\mathfrak{d}}_{\text{st}}(A)$ is always surjective. In general, however, h_A is not injective. For example, if A is commutative then $\text{StanDer}(A) = \{0\}$, $\mathfrak{d}_{\text{st}}(A) = \bigwedge^2 A$ and $\bar{\mathfrak{d}}_{\text{st}}(A) = W_{\text{st}}(A) \cong 3 \bigwedge^2 A$ (abelian Lie algebras), and h_A is multiplication by -3 . Thus \bar{F}_{st} is closer to the standard derivations in the sense that the kernels of

$\bar{\rho}_{\text{st}}(A)$ are smaller. On the other hand, F_{st} commutes with arbitrary base change, whereas \bar{F}_{st} does so only for flat base change, see Prop. 3.12 below for details.

3.10. Base change. Let \mathcal{A} be a category of algebras over \mathcal{R} as in 3.2. We say that \mathcal{A} admits base change if for every $A \in \mathcal{A}_k$ and every homomorphism $\tau : k \rightarrow R$ of \mathcal{R} the R -algebra $A_R = A \otimes_k R$ (with the naturally extended algebraic structure) belongs to \mathcal{A}_R . (In more precise categorical language, this says that \mathcal{A} is a co-fibred category over \mathcal{R} .) This is true for all the examples considered in this paper, in particular for the category of Lie algebras. If \mathcal{A} admits base change then so does $\text{Der}(\mathcal{A})$: Indeed, for a morphism $\tau : k \rightarrow R$ of \mathcal{R} and a derivation action α of $\mathfrak{g} \in \text{Lie}_k$ on $A \in \mathcal{A}_k$, it is easily seen that

$$\alpha_R := \text{can} \circ (\rho \otimes \text{Id}_R) : \mathfrak{g} \otimes_k R \rightarrow \text{Der}(A) \otimes_k R \rightarrow \text{Der}(A \otimes_k R).$$

is a derivation action of $\mathfrak{g}_R = \mathfrak{g} \otimes_k R$ on A_R , called the *base change of α with respect to $\tau : k \rightarrow R$* .

Suppose \mathcal{A} admits base change and $F : \mathcal{A} \rightarrow \text{Der}(\mathcal{A})$ is a derivation functor. We say F commutes with base change if for all morphisms $\tau : k \rightarrow R$ of \mathcal{R} there are natural isomorphisms

$$g(\tau) : \mathfrak{d}(A)_R \xrightarrow{\cong} \mathfrak{d}(A_R), \tag{3-7}$$

making the diagrams

$$\begin{array}{ccc} \mathfrak{d}(A)_R & \xrightarrow{g(\tau)} & \mathfrak{d}(A_R) \\ & \cong & \\ & \searrow & \swarrow \\ (\rho_A)_R & & \rho_{A_R} \\ & \text{Der}(A_R) & \end{array} \tag{3-8}$$

commutative. (Naturality means that the $g(\tau)$ behave in the expected way with respect to composition of morphisms in \mathcal{R} and the usual canonical isomorphisms between repeated tensor products. A more precise formulation would require the formalism of fibred categories.)

One sees immediately that the usual inner derivation functor of associative algebras, see 3.5(b), commutes with arbitrary base change. We will now show that the standard derivation functor F_{st} of alternative algebras commutes with arbitrary base change, and that the other derivation functors of alternative algebras introduced earlier commute with flat base change. Let us emphasize that this does not improve Proposition 2.9 since we are not dealing with the algebras $\rho_A(\mathfrak{d}(A))$ of inner derivations of the respective type themselves, but with the more abstractly defined Lie algebras $\mathfrak{d}(A)$. We begin with a lemma.

3.11. Lemma. *Let M, N be k -modules and let $q: M \rightarrow N$ be a quadratic map. Let $R \in k\text{-alg}$ and let $q_R: M_R \rightarrow N_R$ be the extension of q to a quadratic map of R -modules, cf. [18, Proposition 2.1]. Let $q(M) \subseteq N$ be the k -linear span of $\{q(x) : x \in M\}$ and define $q_R(M_R) \subseteq N_R$ analogously. Finally, let $\iota: q(M) \rightarrow N$ be the inclusion map. Then the base extension $\iota_R: q(M) \otimes R \rightarrow N_R$ has image $q_R(M_R)$.*

Proof. For $y = \sum x_i \otimes r_i \in M_R$, we have

$$q_R(y) = \sum q(x_i) \otimes r_i^2 + \sum_{i < j} q(x_i, x_j) \otimes r_i r_j, \tag{3-9}$$

where $q(-, -)$ is the polar map of q . On the other hand, let us denote a typical spanning element of $q(M) \otimes R$ by $q(x) \tilde{\otimes} r$, to distinguish the tensor product in $q(M) \otimes R$ (where $q(M)$ is taken as a k -module in its own right) from the tensor product in $N \otimes R$. Then $\iota_R(q(x) \tilde{\otimes} r) = q(x) \otimes r = q_R(x \otimes 1_R)r$. This implies that indeed $\iota_R(q(M) \otimes R) \subseteq q_R(M_R)$. Moreover, (3-9) shows that every $q_R(y)$ belongs to the image of ι_R , proving the lemma.

3.12. Proposition. (a) *The standard derivation functor F_{st} of alternative algebras commutes with arbitrary base change.*

(b) *The derivation functors $F_{\text{in}}, F_{\text{ass}}, F_{\text{com}}$ and \bar{F}_{st} commute with flat base change.*

Proof. (a) Let $\tau: k \rightarrow R$ be a ring homomorphism, so R is a k -algebra. It is well known that $\eta: (\wedge^2 A) \otimes R \xrightarrow{\cong} \wedge^2(A_R)$, sending $(a \wedge b)_R \mapsto a_R \wedge b_R$, is an isomorphism of R -modules.

Let $M := \wedge^2 A$, considered as a Leibniz algebra over k as in 3.6 and 3.9, and let $M' = \wedge^2(A_R)$ be the analogously defined Leibniz algebra over R for A_R . The multiplication on A_R is just the R -linear extension of the multiplication on A . Hence $\eta: M_R \rightarrow M'$ is an isomorphism of Leibniz algebras over R .

Let $q: M \rightarrow M$ be the quadratic map $x \mapsto \{x, x\}$ and define $q': M' \rightarrow M'$ in the same way. Then by definition of $\mathfrak{d}_{\text{st}}(A)$ in 3.9, we have an exact sequence

$$0 \longrightarrow q(M) \xrightarrow{\iota} \wedge^2 A \xrightarrow{\pi} \mathfrak{d}_{\text{st}}(A) \longrightarrow 0,$$

which upon tensoring with R yields the first row of the following commutative diagram with exact rows:

$$\begin{array}{ccccccc} q(M) \otimes R & \xrightarrow{\iota_R} & M \otimes R & \xrightarrow{\pi_R} & \mathfrak{d}_{\text{st}}(A) \otimes R & \longrightarrow & 0 \\ \downarrow \varphi & & \cong \downarrow \eta & & \downarrow \psi & & \\ 0 \longrightarrow & q'(M') & \xrightarrow{\iota'} & M' & \xrightarrow{\pi'} & \mathfrak{d}_{\text{st}}(A_R) & \longrightarrow 0 \end{array}$$

In the second row, $q'(M')$ corresponds under η to the image of the quadratic map q_R as in Lemma 3.11. This yields the homomorphism φ and the commutativity of the left hand square. Exactness of the second row is clear from the definition of $\mathfrak{d}_{\text{st}}(A_R)$. Finally, ψ is the unique map making the right hand square commutative. We complete this diagram by adding the kernels and co-kernels of the vertical maps and obtain:

$$\begin{array}{ccccccc}
 \text{Ker}(\varphi) & \longrightarrow & 0 & \longrightarrow & \text{Ker}(\psi) & & \\
 \downarrow & & \downarrow & & \downarrow & & \\
 q(M) \otimes R & \longrightarrow & M \otimes R & \longrightarrow & \mathfrak{d}_{\text{st}}(A) \otimes R & \longrightarrow & 0 \\
 \downarrow \varphi & & \cong \downarrow \eta & & \downarrow \psi & & \\
 0 \longrightarrow & q'(M') & \longrightarrow & M' & \longrightarrow & \mathfrak{d}_{\text{st}}(A_R) & \longrightarrow 0 \\
 \downarrow & & \downarrow & & \downarrow & & \\
 \text{Coker}(\varphi) & \longrightarrow & 0 & \longrightarrow & \text{Coker}(\psi) & &
 \end{array}$$

Now the Snake Lemma [5, §1, No. 2, Prop. 2] yields an isomorphism $\text{Ker}(\psi) \cong \text{Coker}(\varphi)$. Lemma 3.11 implies that φ is surjective, so ψ is injective. But ψ is surjective as well, because η and π' are surjective. This establishes the isomorphisms $g(\tau) = \psi$ of (3-7), and (3-8) is easily verified.

(b) Suppose R is a flat k -algebra. We consider first the inner derivation functor F_{in} defined in 3.7. Here $M = W_{\text{in}}(A)$ is the kernel of the map $s : W(A) \rightarrow A$, and $M' = W_{\text{in}}(A_R)$ is similarly defined. Since R is flat over k , Lemma 1.7(a) yields an isomorphism $\eta : M \otimes R \rightarrow M'$ of Leibniz algebras over R . Now the argument in the proof of (a) can be repeated with \mathfrak{d}_{in} in place of \mathfrak{d}_{st} and yields an isomorphism $\mathfrak{d}_{\text{in}}(A) \otimes R \cong \mathfrak{d}_{\text{in}}(A_R)$. The proof of the other cases follows the same pattern. The details are left to the reader.

3.13. Remarks. The argumentation in the proof of Proposition 3.12(b) made use of flatness only to ensure that there is an isomorphism $\eta : M \otimes R \rightarrow M'$. This can be used to prove base change results for arbitrary R , by restricting the category \mathcal{A} . For example, let \mathcal{A} be the category octonion algebras (see 4.1 below for the definition), with morphisms unital homomorphisms of algebras, and consider associator derivations. Here $M = W_{\text{ass}}(A)$ is the kernel of the commutator map $s : \bigwedge^2 A \rightarrow A$. The linear span of all commutators $[a, b]$ in an octonion algebra A is precisely the kernel of the trace $t_A : A \rightarrow k$, and the trace is surjective. Hence $\text{Ker}(t_A) = \text{Im}(s)$ is a finitely generated and projective module (of rank 7). It follows

that the exact sequence

$$0 \longrightarrow M \longrightarrow \bigwedge^2 A \xrightarrow{s} [A, A] \longrightarrow 0$$

splits and therefore remains exact (and split) upon tensoring with an arbitrary $R \in k\text{-alg}$. Hence the natural map $\eta: M \otimes R \rightarrow M' = \text{Ker}(s_R)$ is an isomorphism. It follows that the functor F_{ass} commutes with arbitrary base change for octonion algebras.

4. Octonion algebras: basic properties

In this section, we prepare the ground for describing derivations of octonion algebras over arbitrary commutative rings.

4.1. The concept of an octonion algebra. Following [17, 1.8], a non-associative algebra C over k is called an *octonion algebra* if it is finitely generated projective of rank 8 as a k -module, contains an identity element and admits a *norm*, i.e., a quadratic form $n_C: C \rightarrow k$ uniquely determined by the following two conditions:

- (i) n_C is *non-singular*, so its induced symmetric bilinear form

$$n_C(x, y) = n_C(x + y) - n_C(x) - n_C(y)$$

defines a linear isomorphism from the k -module C onto its dual C^* by the assignment $x \mapsto n_C(x, -)$.

- (ii) n_C *permits composition*, i.e., the relation

$$n_C(xy) = n_C(x)n_C(y) \tag{4-1}$$

holds for all $x, y \in C$.

We then call $t_C = n_C(1_C, -)$ the *trace* of C . Since the rank of C is everywhere positive, $1_C \in C$ is a *unimodular vector* [14, 0.3], i.e., $k1_C$ is a free k -module of rank 1 and a direct summand of C (as a k -module).

Octonion algebras are alternative (but not associative) and invariant under base change. They also descend from faithfully flat base change: If $R \in k\text{-alg}$ is faithfully flat over k and C is a k -algebra such that C_R is an octonion algebra over R then C is an octonion algebra over k . This follows from faithfully flat descent and the fact that the norm and the unit element of an octonion algebra are uniquely determined.

By [16], given an octonion algebra C over k , the relations

$$n_C(1_C) = 1, \quad t_C(1_C) = 2, \tag{4-2}$$

$$x^2 - t_C(x)x + n_C(x)1_C = 0, \tag{4-3}$$

$$t_C(xy) = t_C(x)t_C(y) - n_C(x, y) \tag{4-4}$$

hold for all $x, y \in C$, and t_C is an *associative* linear form in the sense that it vanishes on all commutators and associators of the algebra. Moreover, the *conjugation* of C , i.e., the linear map $\iota_C: C \rightarrow C, x \mapsto \bar{x} := t_C(x)1_C - x$, is an algebra involution satisfying $x\bar{x} = n_C(x)1_C, x + \bar{x} = t_C(x)1_C$, and

$$xyx = n_C(x, \bar{y})x - n_C(x)\bar{y} \tag{4-5}$$

for all $x, y \in C$. In particular, x is invertible in C if and only $n_C(x)$ is a unit in k , in which case $x^{-1} = n_C(x)^{-1}\bar{x}$. Recall that octonion algebras over fields are simple [23, Chap. 2, Lemma 3]. As a consequence, octonion algebras over rings share with Azumaya algebras the property that *a unital homomorphism $f: C \rightarrow C'$ of octonion algebras is an isomorphism*. Indeed, localizing if necessary, we may assume that k is a local ring, with residue field K . Then the kernel of the induced homomorphism $f_K: C_K \rightarrow C'_K$ is an ideal $\neq C_K$, hence $\{0\}$. Thus f_K is injective and therefore bijective, because both algebras have dimension 8. It follows that f is an isomorphism by [3, II, §3.2, Cor. of Prop. 6].

Remark. The same argument leads to the same conclusion in the more general setting of arbitrary unital non-associative k -algebras C, C' that are finitely generated projective of the same rank as k -modules and have C_K simple for all fields $K \in k\text{-alg}$.

We now proceed to describe particularly simple and useful examples of octonion algebras.

4.2. Zorn vector matrices and split octonions. There are various formally different but equivalent ways of defining an octonion algebra structure on the k -module

$$Z := \text{Zor}(k) := \begin{bmatrix} k & k^3 \\ k^3 & k \end{bmatrix}$$

of *Zorn vector matrices* over k , i.e., of 2×2 matrices with diagonal entries in k and off-diagonal ones in column space k^3 over k . The normalization chosen here is due to Zorn [24] and turns out to be the most convenient for our subsequent computations. Accordingly, we define

$$\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & v \\ y & \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_1\beta_1 - u^t y & \alpha_1 v + \beta_2 u + x \times y \\ \beta_1 x + \alpha_2 y + u \times v & -x^t v + \alpha_2\beta_2 \end{bmatrix} \tag{4-6}$$

for $\alpha_i, \beta_i \in k, (i = 1, 2), u, v, x, y \in k^3$, where $u^t v$ and $u \times v$ stand for the ordinary scalar and vector product, respectively, of $u, v \in k^3$. Then $\text{Zor}(k)$ becomes an octonion algebra under the multiplication (4-6). Its unit element, norm, and trace

are given by the formulas

$$1_Z = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad n_Z(a) = \alpha_1\alpha_2 + u^t x, \quad t_Z(a) = \alpha_1 + \alpha_2$$

for $a = \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \in Z$. Note that this definition has the advantage of yielding complete symmetry in the indices 1, 2 but is not consistent with the usual definition of matrix multiplication of 2×2 -matrices in the following sense: Let $u, x \in k^3$ such that $u^t x = 1$. Then $\begin{bmatrix} k & k \cdot u \\ k \cdot x & k \end{bmatrix}$ is a subalgebra of $\text{Zor}(k)$ isomorphic to $\text{Mat}_2(k)$ under the map $\begin{bmatrix} \alpha & \beta \cdot u \\ \gamma \cdot x & \delta \end{bmatrix} \mapsto \begin{pmatrix} \alpha & \beta \\ -\gamma & \delta \end{pmatrix}$, and $n_Z \begin{bmatrix} \alpha & \beta \cdot u \\ \gamma \cdot x & \delta \end{bmatrix} = \det \begin{pmatrix} \alpha & \beta \\ -\gamma & \delta \end{pmatrix}$. The square brackets (instead of the usual round brackets) serve to indicate this fact.

Let u_i ($i = 1, 2, 3$) be the standard basis of k^3 . It is evident that Z is free of rank eight as a k -module with basis

$$\mathbf{b}_s = (E_1, X_1, X_2, X_3; E_2, Y_1, Y_2, Y_3)$$

given by

$$E_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad X_i = \begin{bmatrix} 0 & u_i \\ 0 & 0 \end{bmatrix}, \quad Y_i = \begin{bmatrix} 0 & 0 \\ u_i & 0 \end{bmatrix}.$$

We call \mathbf{b}_s the *standard basis* of Z . It satisfies the following relations:

$$E_1 E_2 = E_2 E_1 = 0, \quad E_1^2 = E_1, \quad E_2^2 = E_2, \quad (4-7)$$

$$E_1 X_i = X_i = X_i E_2, \quad E_2 X_i = 0 = X_i E_1, \quad (4-8)$$

$$E_2 Y_i = Y_i = Y_i E_1, \quad E_1 Y_i = 0 = Y_i E_2, \quad (4-9)$$

$$X_i X_j = \text{sgn}(i, j) Y_l, \quad Y_i Y_j = \text{sgn}(i, j) X_l, \quad (4-10)$$

$$X_i Y_j = -\delta_{ij} E_1, \quad Y_i X_j = -\delta_{ij} E_2, \quad (4-11)$$

$$E_1 + E_2 = 1_Z. \quad (4-12)$$

Here $\text{sgn}(i, j)$ is zero for $i = j$ and equals the sign of the permutation (i, j, l) (sending 1 to i , 2 to j , 3 to l) if $i \neq j$ and l is the missing index.

Our next step consists in introducing twisted versions of Zorn vector matrices.

4.3. Reduced octonion algebras. An octonion algebra over k is said to be *reduced* if it is isomorphic to an algebra $\text{Zor}(M, \theta)$, defined as follows ([17, 3.2, 3.3])¹: Let M be a finitely generated projective module of rank 3 over k . Writing $M^* = \text{Hom}_k(M, k)$ for the dual of M and $\langle \cdot, \cdot \rangle: M^* \times M \rightarrow k$ for the natural pairing, we identify $\bigwedge^3 M^* = (\bigwedge^3 M)^*$ canonically by means of the formula

$$\langle \alpha_1 \wedge \alpha_2 \wedge \alpha_3, x_1 \wedge x_2 \wedge x_3 \rangle = \det \langle (\alpha_i, x_j) \rangle$$

¹We deviate from the terminology in [17], where these algebras are called split.

for $\alpha_i \in M^*, x_j \in M, 1 \leq i, j \leq 3$. Now suppose we are given a *volume element* of M , i.e., an isomorphism $\theta: \bigwedge^3 M \xrightarrow{\sim} k$ of k -modules (which may not exist but if it does is unique up to an invertible factor in k). Then $\theta^*: k = k^* \xrightarrow{\sim} \bigwedge^3 M^*$, the dual of θ , gives rise to the volume element θ^{*-1} of M^* , and we obtain two associated vector products

$$\times_\theta : M \times M \longrightarrow M^*, \quad \times_\theta : M^* \times M^* \longrightarrow M$$

by means of the formulas

$$\langle x \times_\theta y, z \rangle = \theta(x \wedge y \wedge z), \quad \langle \zeta, \xi \times_\theta \eta \rangle = \theta^{*-1}(\zeta \wedge \xi \wedge \eta) \tag{4-13}$$

for all $x, y, z \in M, \xi, \eta, \zeta \in M^*$. Note that both vector products are alternating and induce isomorphisms $\bigwedge^2 M \xrightarrow{\sim} M^*, \bigwedge^2 M^* \xrightarrow{\sim} M$. To simplify notations, we write \times instead of \times_θ whenever the context is clear. Furthermore, to make matters more symmetric, we identify $M \cong M^{**}$ canonically, put $M^+ := M, M^- := M^*$ and then have two dualizing bilinear forms $\langle \cdot, \cdot \rangle : M^\pm \times M^\mp \rightarrow k$ satisfying the relation $\langle x^+, x^- \rangle = \langle x^-, x^+ \rangle$ for all $x^\pm \in M^\pm$. Now the k -module

$$C := \text{Zor}(M, \theta) = \begin{bmatrix} k & M^+ \\ M^- & k \end{bmatrix}$$

becomes an octonion algebra over k under the multiplication

$$\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \begin{bmatrix} \beta_1 & v \\ y & \beta_2 \end{bmatrix} = \begin{bmatrix} \alpha_1\beta_1 - \langle u, y \rangle & \alpha_1v + \beta_2u + x \times y \\ \beta_1x + \alpha_2y + u \times v & -\langle x, v \rangle + \alpha_2\beta_2 \end{bmatrix} \tag{4-14}$$

for $\alpha_i, \beta_i \in k (i = 1, 2), u, v \in M^+, x, y \in M^-,$ whose unit element, norm, trace are given by

$$1_C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad n_C(a) = \alpha_1\alpha_2 + \langle u, x \rangle, \quad t_C(a) = \alpha_1 + \alpha_2 \tag{4-15}$$

for $a = \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \in C$. If $M = k^3$ is free, then $C = \text{Zor}(k)$ is the *split* octonion algebra of ordinary Zorn vector matrices over k .

4.4. Elementary idempotents. We claim: *For an element e in an octonion algebra C over k to be an idempotent different from $0, 1$ in all scalar extensions (so $e^2 = e$ and $e_R \neq 0, 1_{C_R}$ for all $R \in k\text{-alg}, R \neq \{0\}$) it is necessary and sufficient that $t_C(e) = 1$ and $n_C(e) = 0$.* The condition is clearly sufficient, by (4-2), (4-3). To prove necessity, we may assume that k is a local ring, hence, in particular, connected. Then $n_C(e)$, being an idempotent in k by (4-1), satisfies $n_C(e) = 0$ or $n_C(e) = 1$. In the latter case, e would be invertible, forcing the contradiction $e = 1_C$. Hence $n_C(e) = 0$, and (4-3) yields $e = e^2 = t_C(e)e$. Taking traces, we

conclude that $t_C(e) \in k$ is an idempotent which cannot be zero since $e \neq 0$. Thus $t_C(e) = 1$.

Elements of C satisfying the equivalent conditions above are called *elementary idempotents*. If e is such and $e_1 := e, e_2 := 1_C - e = \bar{e}_1$, then e_2 is an elementary idempotent as well and (e_1, e_2) is a hyperbolic pair of the quadratic space (C, n_C) . Moreover, using (4-5) as well as (4-1) and its bi-linearizations, the Peirce components $C_{ij} := C_{ij}(e)$ ($i, j = 1, 2$) of C relative to e [20, III, §2] are easily seen to satisfy the relations ($i, j = 1, 2, i \neq j$)

$$C_{ii} = ke_i, \quad n_C(C_{ij}) = n_C(e_i, C_{12} + C_{21}) = t_C(C_{12} + C_{21}) = \{0\}. \quad (4-16)$$

Since e_i is a unimodular vector, $C_{ii} \cong k$ as k -algebras. Also, (4-16) implies that the k -modules C_{12} and C_{21} are dually paired by $n_C(-, -)$, so the decomposition $C = \bigoplus_{i,j \in \{1,2\}} C_{ij}$ together with $\text{rk } C = 8$ shows that they are both finitely generated projective of rank 3.

4.5. Schemes. In a slightly more general vein than 1.11, we follow [7] and view schemes over k as special covariant set-valued functors on $k\text{-alg}$. Then the affine scheme \mathbf{X} defined by a *fixed* k -algebra A is the functor $\mathbf{X}(R) = \text{Hom}_{k\text{-alg}}(A, R)$ ($R \in k\text{-alg}$); i.e., the affine schemes are precisely the representable functors. For example, if M is a k -module, we denote by $M_{\mathbf{a}}$ the functor defined by $M_{\mathbf{a}}(R) := M \otimes R$ for all $R \in k\text{-alg}$. If M is finitely generated and projective then $M_{\mathbf{a}}$ is an affine k -scheme represented by the symmetric algebra over the dual M^* of M . We say that a k -scheme \mathbf{X} is *faithful* if it has non-empty geometric fibres: $\mathbf{X}(K) \neq \emptyset$ for all algebraically closed fields $K \in k\text{-alg}$. In case \mathbf{X} is affine and represented by a finitely presented k -algebra A , this is equivalent to the canonical map $\text{Spec}(A) \rightarrow \text{Spec}(k)$ of the prime spectra being surjective.

Given any k -scheme \mathbf{X} , we will make use of the following facts:

- (i) ([8, (17.16.2)]) If \mathbf{X} is fppf (=flat, faithful and finitely presented), there exists an fppf extension R of k such that $\mathbf{X}(R) \neq \emptyset$.
- (ii) ([8, (17.16.3)]) If \mathbf{X} is smooth and faithful, we may choose R as in (i) to be even étale.
- (iii) ([8, (17.1.1), (17.3.1)], [7, I, §4, 4.6]) \mathbf{X} is smooth if and only if it is finitely presented and, for all $R \in k\text{-alg}$ and all ideals $I \subseteq R$ satisfying $I^2 = \{0\}$, the natural map $\mathbf{X}(R) \rightarrow \mathbf{X}(R/I)$ is surjective. A smooth scheme is flat.
- (iv) ([8, (17.7.3)]) If $R \in k\text{-alg}$ is faithfully flat, then for \mathbf{X} to be smooth over k it is necessary and sufficient that its base change \mathbf{X}_R from k to R be smooth over R .

4.6. Splittings and splitting bases. A *splitting* of an octonion algebra C over k is an isomorphism $f: Z = \text{Zor}(k) \rightarrow C$. We denote by $\text{Isom}(Z, C)$ the (possibly

empty) set of splittings of C and define a functor $\mathbf{X} = \mathbf{Isom}(Z, C): k\text{-alg} \rightarrow \mathbf{set}$ by

$$\mathbf{X}(R) = \text{Isom}(Z_{\text{or}}(R), C_R) \quad (R \in k\text{-alg}).$$

Let $\mathbf{G} = \mathbf{Aut}(Z)$ be the automorphism group scheme of Z . If $\mathbf{X}(R) \neq \emptyset$ then it is immediately seen that the group $\mathbf{G}(R)$ acts simply transitively on the right on $\mathbf{X}(R)$ by composition.

A *splitting basis* of C is an octuple $\mathbf{b} = (e_1, x_1, x_2, x_3; e_2, y_1, y_2, y_3) \in C^8$ satisfying the relations (4-7)–(4-12), with upper case letters replaced by lower case ones. Thus by its definition, a splitting basis is not required to be a basis of the k -module C but in fact is, as will be seen now.

Given a splitting $f: Z \xrightarrow{\cong} C$ of C , it is clear that the image $f(\mathbf{b}_s)$ of the standard basis of Z is a splitting basis of C . We claim that this establishes a bijection between $\text{Isom}(Z, C)$ and the set of splitting bases of C . Indeed, since f is linear and \mathbf{b}_s is in particular a basis of Z as a k -module, f is uniquely determined by its values on \mathbf{b}_s so the map $f \mapsto f(\mathbf{b}_s)$ is injective. To prove surjectivity, let \mathbf{b} be a splitting basis of C . The defining relations (4-7)–(4-12) of a splitting basis say precisely that the linear map $f: Z \rightarrow C$ defined by $f(\mathbf{b}_s) = \mathbf{b}$ is a unital homomorphism of octonion algebras and therefore an isomorphism, as remarked in 4.1. In particular, \mathbf{b} is a basis of C as a k -module.

An essential step in the proof of the main result of this section is to show that \mathbf{X} is a smooth k -scheme. The proof will be facilitated by introducing the following concept.

4.7. Splitting data. Let C be an octonion algebra over k . A *splitting datum* for C is a quadruple $\mathbf{d} = (e, x_1, x_2, x_3) \in C^4$ satisfying the following conditions:

$$e \text{ is an elementary idempotent,} \tag{4-17}$$

$$\text{the } x_i \text{ belong to the Peirce space } C_{12}(e), \tag{4-18}$$

$$x_1(x_2x_3) = -e. \tag{4-19}$$

Let \mathbf{b}_s be the standard basis of $Z = Z_{\text{or}}(k)$ as in 4.2. It is clear from 4.2 that $\mathbf{d}_s = (E_1, X_1, X_2, X_3)$ is a splitting datum of the split algebra Z , called the *standard splitting datum*.

4.8. Lemma. *Let C be an octonion algebra over k . Then the map $\phi: f \mapsto f(\mathbf{d}_s)$ is a bijection between $\text{Isom}(Z, C)$ and the set of splitting data of C .*

Proof. If $f: Z \rightarrow C$ is an isomorphism then it is clear that $f(\mathbf{d}_s)$ is a splitting datum of C , so the map ϕ is well-defined. To prove ϕ injective we have to show that an $f \in \text{Isom}(Z, C)$ is uniquely determined by its values on \mathbf{d}_s . Since f is a homomorphism of unital algebras, the relations (4-10) and (4-12) show

that $f(Y_i) = f(X_j X_l) = f(X_j) f(X_l)$ (where (i, j, l) is a cyclic permutation of $(1, 2, 3)$), and $f(E_2) = f(1 - E_1) = 1 - f(E_1)$. Hence $f(\mathbf{d}_s)$ determines the values of f on \mathbf{b}_s and therefore f , by 4.6.

Again by 4.6, ϕ surjective means every splitting datum \mathbf{d} of C extends to a splitting basis \mathbf{b} . Thus let $\mathbf{d} = (e, x_1, x_2, x_3)$ be a splitting datum of C and define $\mathbf{b} = (e_1, x_1, x_2, x_3; e_2, y_1, y_2, y_3)$ by

$$e_1 := e, \quad e_2 := 1_C - e_1, \quad y_1 := x_2 x_3, \quad y_2 := x_3 x_1, \quad y_3 := x_1 x_2.$$

We verify the relations (4-7)–(4-12) for \mathbf{b} . Here (4-7) and (4-12) are clear, and (4-8) holds by (4-18). The equations (4-9) just say $y_i \in C_{21}$, which follows from $C_{12}^2 \subseteq C_{21}$ (by the Peirce rules) and the definition of y_i .

Before continuing, we make the following remarks. Since e is an elementary idempotent by (4-17), we have $t_C(x) = n_C(x) = 0$ for $x \in C_{12} \cup C_{21}$ by (4-16). This implies $x^2 = t_C(x)x - n_C(x)1 = 0$, so the multiplication of C restricted to C_{12} and to C_{21} is alternating. Moreover:

The trilinear expressions $x(yz)$ and $(xy)z$ (where $x, y, z \in C_{12}$) are alternating. (4-20)

Indeed, their difference is the associator which is alternating, so it suffices to prove that $x(yz)$ is alternating. But this follows immediately from $xy^2 = 0 = x^2 y = x(xy)$ and the fact that a multilinear map is alternating as soon as it vanishes when two adjacent arguments are equal.

The equations $x_i x_j = \text{sgn}(i, j) y_l$ of (4-10) now hold by definition of the y_i and the alternating character of the product. The latter also allows us to assume, in proving the second group of equations $y_i y_j = \text{sgn}(i, j) x_l$, that (i, j, l) is a cyclic permutation. Then the middle Moufang identity (cf. (2-4)), the alternating nature of the product together with (4-20) and (4-19) imply

$$y_i y_j = (-x_l x_j)(-x_i x_l) = (x_l(x_j x_i)) x_l = -(x_1(x_2 x_3)) x_l = e_1 x_l = x_l.$$

To prove the first group of relations $x_i y_j = -\delta_{ij} e_1$ of (4-11), write $y_j = x_l x_m$ where (j, l, m) is cyclic, so that $x_i y_j = x_i(x_l x_m)$. If $i = j$ this is $x_j(x_l x_m) = x_1(x_2 x_3) = -e_1$ by (4-19) and (4-20). If $i \neq j$ then either $i = l$ or $i = m$, and hence $x_i(x_l x_m) = 0$, again by (4-20).

The remaining equations $y_i x_j = -\delta_{ij} e_2$ follow by applying the involution and observing that $t_C(x) = 0$ implies $\bar{x} = -x$ for $x \in C_{12} + C_{21}$.

4.9. Torsors. Let \mathbf{X} be a k -scheme and \mathbf{G} a k -group scheme acting on \mathbf{X} on the right in a simply transitive manner; i.e., for all $R \in k\text{-alg}$ and all $x, y \in \mathbf{X}(R)$ there exists exactly one $g \in \mathbf{G}(R)$ such that $y = xg$. Note that $\mathbf{X}(R)$ may well be empty. Then \mathbf{X} is said to be a *torsor in the flat topology with structure group \mathbf{G}* if there exists a fppf $S \in k\text{-alg}$ such that $\mathbf{X}(S) \neq \emptyset$ [7, III, §4]. If S can be chosen in addition

étale then \mathbf{X} is called a torsor in the étale topology. Fixing an element $x_0 \in \mathbf{X}(S)$, we have an isomorphism $\mathbf{G}_S \xrightarrow{\sim} \mathbf{X}_S$ by $g \mapsto x_0g$. Consequently, by faithfully flat descent, properties of \mathbf{X} and \mathbf{G} correspond to each other. In particular, \mathbf{X} is smooth if and only if \mathbf{G} is smooth, cf. 4.5(iv).

4.10. Theorem. *Let C be an octonion algebra over k and $Z = \text{Zor}(k)$ the algebra of Zorn vector matrices. Then $\mathbf{X} = \mathbf{Isom}(Z, C)$ is an affine smooth torsor in the étale topology with structure group $\mathbf{G} = \mathbf{Aut}(Z)$.*

Proof. Let \mathbf{Y} be the functor assigning to $R \in k\text{-alg}$ the set of splitting data of C_R . The map ϕ of Lemma 4.8 is compatible with arbitrary base changes and thus induces an isomorphism $\phi: \mathbf{X} \rightarrow \mathbf{Y}$ of functors. The conditions (4-17)–(4-19) show that $\mathbf{Y} \subset C_{\mathbf{a}}^4$ is defined by finitely many polynomial equations, so \mathbf{Y} and therefore \mathbf{X} is an affine finitely presented k -scheme. Hence to prove smoothness of \mathbf{X} , we may use 4.5(iii), and have to show: If $R \in k\text{-alg}$ and $I \subset R$ is an ideal of square zero then every splitting datum of $C_{R/I}$ over R/I can be lifted to a splitting datum of C_R over R .

We may assume $R = k$, replacing C by C_R if necessary. Write

$$\pi : C \rightarrow C' := C/IC = C \otimes (k/I)$$

for the canonical map and let $\mathbf{d}' = (e', x'_1, x'_2, x'_3)$ be a splitting datum of C' . Denote norm and trace of C and C' by n, t and n', t' , respectively. As $IC \subseteq C$ is a nil ideal, it is a standard fact that e can be lifted to an idempotent e of C . We have $n'(e') = 0, t'(e') = 1$ by 4.4, so we conclude from (4-1) that $n(e)$ is a nilpotent idempotent in k . Hence $n(e) = 0$ and $e = t(e)e$ by (4-3). Applying t , we obtain $t(e)^2 = t(e)$, and $t'(e') = 1$ shows $t(e) \equiv 1 \pmod I$, whence $t(e)$ is an invertible idempotent in k . Thus $t(e) = 1$, and we have shown that e is elementary.

Let C_{ij} and C'_{ij} denote the Peirce spaces of C and C' relative to e and e' , respectively. Since π is a surjective algebra homomorphism mapping e to e' , we have $\pi(C_{12}) = C'_{12}$. Hence the elements $x'_i \in C'_{12}$ can be lifted to elements $x_i \in C_{12}$ ($1 \leq i \leq 3$). Now $x'_1(x'_2x'_3) = -e'_1$ implies $x_1(x_2x_3) = (-1 + \alpha)e_1$ for some $\alpha \in I$. As $(1 - \alpha)^{-1} = 1 + \alpha$ (recall that $\alpha^2 = 0$ since I squares to zero), we see that (4-19) holds for $\mathbf{d} = (e, x_1, x_2, (1 + \alpha)x_3)$, so \mathbf{d} is the desired lift of \mathbf{d}' to a splitting datum of C .

As noted in 4.6, \mathbf{G} acts simply transitively on \mathbf{X} . Thus to prove that \mathbf{X} is a torsor in the étale topology, it remains to show that $\mathbf{X}(S) \neq \emptyset$ for some étale faithfully flat $S \in k\text{-alg}$. It is a standard fact that an octonion algebra over an algebraically closed field K is split [21, Thm. 1.8.1, 1.10(i)]. Hence $\mathbf{X}(K) \neq \emptyset$, so \mathbf{X} is a faithful k -scheme. Now the existence of S follows from (ii) of 4.5. □

The following corollaries are now an immediate consequence of the theorem and 4.5.

4.11. Corollary. *For C to be an octonion algebra over k it is necessary and sufficient that C be a k -algebra and there exist a faithfully flat étale k -algebra R such that $C_R \cong \text{Zor}(R)$ is a split octonion algebra over R .*

4.12. Corollary. *Let C be an octonion algebra over k . Then $\mathbf{Aut}(C)$ is a smooth group scheme.*

5. Octonion algebras: derivations

We are now ready for our main result on derivations of octonion algebras.

5.1. Theorem. *Every derivation of an octonion algebra C over k is an associator derivation: $\text{AssDer}(C) = \text{Der}(C)$, so every derivation of C has the form*

$$\sum_{i=1}^m [L_{a_i}, R_{b_i}]$$

where $m \in \mathbb{N}$ and $a_i, b_i \in C$ ($1 \leq i \leq m$) satisfy the relation $\sum [a_i, b_i] = 0$.

If the base ring contains $\frac{1}{3}$, every inner derivation of C is standard (Prop. 2.7(f)), so Thm. 5.1 yields the following extension of Schafer’s theorem ([19, Thm. 6], [20, Cor. 3.29] or, more generally, [1, Prop. 1]) to commutative rings.

5.2. Corollary. *If $\frac{1}{3} \in k$, then every derivation of an octonion algebra over k is standard, i.e., it is a sum of derivations $D_{u,v}$, $u, v \in C$. □*

Remarks. (a) Without the hypothesis on k , e.g., over fields of characteristic 3, Cor. 5.2 is false; see [1] for details.

(b) Derivations of Azumaya algebras are always inner [12, III, Thm. 1.4, Thm. 5.1]. Thm. 5.1 may be regarded as an analogue of this result for octonion algebras.

Proof of Thm. 5.1, step 1. Our proof of Thm. 5.1 proceeds in two steps, the first one combining Cor. 4.11 with the fact that the full derivation algebra of C by 1.11 and Cor. 4.12 (resp. the ideal of associator derivations by Prop. 2.9) commutes with arbitrary (resp. flat) base change. Hence we may assume that C is split.

In the sequel, we will work under the less restrictive assumption that C be reduced, and we do so for two reasons. For one, the proof will sometimes become more natural in this slightly more general setting. For another, we will be able to derive a number of intermediate results of independent interest that retain their validity for reduced rather than just split octonions.

The second step of the proof will be preceded by a digression into graded modules and algebras.

5.3. Graded modules. Let Γ be a finite additive abelian group and $M = \bigoplus_{\gamma \in \Gamma} M_\gamma$ a Γ -graded k -module [4, II, §11, n° 2]. Since Γ is finite, $\text{End}_k(M)$ becomes a Γ -graded k -algebra whose γ -homogeneous component, $\text{End}_k(M)_\gamma$, $\gamma \in \Gamma$, consists of all graded homomorphisms $f : M \rightarrow M$ of degree γ , so f is k -linear and satisfies $f(M_\delta) \subseteq M_{\gamma+\delta}$ for all $\delta \in \Gamma$. Clearly, since $\text{End}_k(M)$ is a Γ -graded k -algebra, so is $\mathfrak{gl}(M)$. Moreover, if we are given a non-associative Γ -graded k -algebra structure A on M , one checks easily that $\text{Der}(A)$ is a graded subalgebra of $\mathfrak{gl}(A)$.

5.4. Example. Let C be a unital alternative k -algebra and $e \in C$ an idempotent. Writing $e_1 := e$, $e_2 := 1_C - e$, the multiplication rules for the Peirce components $C_{ij} = C_{ij}(e)$ ($i, j = 1, 2$) [20, III, §2] imply that

$$C = C_0 \oplus C_1 \oplus C_2, \quad C_0 := C_{00} \oplus C_{11}, \quad C_1 := C_{12}, \quad C_2 := C_{21}$$

gives a $\mathbb{Z}/3\mathbb{Z}$ -grading of C as a k -algebra, called the e -grading of C . We write

$$\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2$$

for the corresponding $\mathbb{Z}/3\mathbb{Z}$ -grading of the derivation algebra $\mathfrak{g} := \text{Der}(C)$ in the sense of 5.3 and call this the e -grading of \mathfrak{g} . Fixing $i, j \in \{1, 2\}$, $i \neq j$ and $u_{ij} \in C_{ij} = C_i$, it is straightforward to check, using (2-28), that the derivation

$$D_i(u_{ij}) := -D_{e_i, u_{ij}} = D_{e_j, u_{ij}} \in \text{StanDer}(C)$$

satisfies the relations

$$D_i(u_{ij})x_{ii} = x_{ii}u_{ij}, \tag{5-1}$$

$$D_i(u_{ij})x_{jj} = -u_{ij}x_{jj}, \tag{5-2}$$

$$D_i(u_{ij})x_{ij} = u_{ij}x_{ij}, \tag{5-3}$$

$$D_i(u_{ij})x_{ji} = -[u_{ij}, x_{ji}] \tag{5-4}$$

for all $x_{\lambda, \mu} \in C_{\lambda, \mu}$, $\lambda, \mu = 1, 2$.

5.5. Proposition. *Assumptions and notations being as in 5.4, the e -grading of $\mathfrak{g} = \text{Der}(C)$ is given by $\mathfrak{g} = \mathfrak{g}_0 \oplus \mathfrak{g}_1 \oplus \mathfrak{g}_2$, where*

$$\mathfrak{g}_0 = \{D \in \mathfrak{g} \mid De = 0\}, \quad \mathfrak{g}_i = \{D_i(u_{ij}) \mid u_{ij} \in C_{ij}\} \quad (\{i, j\} = \{1, 2\}).$$

Moreover, the maps $u_{ij} \mapsto D_i(u_{ij})$ are k -module isomorphisms $C_{ij} \cong \mathfrak{g}_i$ for $i \neq j$.

Proof. Setting $\mathfrak{g}'_0 := \{D \in \mathfrak{g} \mid De = 0\}$ and $\mathfrak{g}'_i := \{D_i(u_{ij}) \mid u_{ij} \in C_{ij}\}$ ($\{i, j\} = \{1, 2\}$), we first claim

$$\mathfrak{g} = \mathfrak{g}'_0 + \mathfrak{g}'_1 + \mathfrak{g}'_2. \tag{5-5}$$

Given $D \in \mathfrak{g}$, let $D(e) = u_{11} + u_{12} + u_{21} + u_{22}$ be the Peirce decomposition of $D(e)$. Then $D(e) = D(e^2) = D(e) \cdot e + e \cdot D(e)$, and comparing Peirce components yields

$D(e) = u_{12} + u_{21}$. This implies $D_0 := D - D_1(u_{12}) + D_2(u_{21}) \in \mathfrak{g}'_0$ by (5-1), (5-2), which completes the proof of (5-5). But now, consulting (5-1)–(5-4) and the Peirce rules, we conclude $\mathfrak{g}'_i \subseteq \mathfrak{g}_i$ for $i = 0, 1, 2$, hence obtain equality by comparing (5-5) with the e -grading of \mathfrak{g} . Finally, $D_i(u_{ij}) = 0$ implies $D_i(u_{ij}) \cdot e_i = u_{ij} = 0$ by (5-1), proving the last statement. \square

The elements of $\mathfrak{g}_1, \mathfrak{g}_2$ are obviously standard derivations. But one can do better than that by deriving the following result, which is due to the referee.

5.6. Proposition. *Notations and assumptions being as in 5.4, let $i, j \in \{1, 2\}$ be distinct and $u_{ji}, v_{ji} \in C_{ji}$. Using the formalism of 2.4, the element*

$$u := 2e_i \wedge (u_{ji}v_{ji}) - u_{ji} \wedge v_{ji} \in \bigwedge^2 C \subseteq W(C) = C \oplus \bigwedge^2 C$$

satisfies $s(u) = 0$ and

$$D_i(u_{ji}v_{ji}) = \Delta_u \in \text{AssDer}(C). \tag{5-6}$$

Proof. From 2.4 and the Peirce rules we conclude $s(u) = 2e_i(u_{ji}v_{ji}) - 2(u_{ji}v_{ji})e_i - u_{ji}v_{ji} + v_{ji}u_{ji} = 2u_{ji}v_{ji} - 2u_{ji}v_{ji} = 0$. Hence the proposition will follow once we have shown

$$D_i(u_{ji}v_{ji})a = Da, \quad D := \Delta_u = 2[L_{e_i}, R_{u_{ji}v_{ji}}] - [L_{u_{ji}}, R_{v_{ji}}] \tag{5-7}$$

for $a = x_{lm} \in C_{lm}$, $l, m \in \{i, j\}$. The Peirce rules combine with (5-1) and the linearization of left alternativity (2-1) to yield

$$\begin{aligned} Dx_{ii} &= 2e_i[x_{ii}(u_{ji}v_{ji})] - 2(e_ix_{ii})(u_{ji}v_{ji}) - u_{ji}(x_{ii}v_{ji}) + (u_{ji}x_{ii})v_{ji} \\ &= (u_{ji}x_{ii} + x_{ii}u_{ji})v_{ji} = u_{ji}(x_{ii}v_{ji}) + x_{ii}(u_{ji}v_{ji}) \\ &= D_i(u_{ji}v_{ji})x_{ii}; \end{aligned}$$

hence (5-7) holds for $a = x_{ii}$. Similarly, invoking (2-2) and (5-2), we obtain

$$\begin{aligned} Dx_{jj} &= 2e_i[x_{jj}(u_{ji}v_{ji})] - 2(e_ix_{jj})(u_{ji}v_{ji}) - u_{ji}(x_{jj}v_{ji}) + (u_{ji}x_{jj})v_{ji} \\ &= -u_{ji}(x_{jj}v_{ji} + v_{ji}x_{jj}) = -(u_{ji}x_{jj})v_{ji} - (u_{ji}v_{ji})x_{jj} \\ &= D_i(u_{ji}v_{ji})x_{jj}; \end{aligned}$$

hence (5-7) holds for $a = x_{jj}$. But the Peirce rules also combine with (5-4) and the linearization of $s_{ji}(s_{ji}t_{ji}) = s_{ji}^2t_{ji} = 0 = t_{ji}s_{ji}^2 = (t_{ji}s_{ji})s_{ji}$ for $s_{ji}, t_{ji} \in C_{ji}$ to yield $Dx_{ji} = 2e_i[x_{ji}(u_{ji}v_{ji})] - 2(e_ix_{ji})(u_{ji}v_{ji}) - u_{ji}(x_{ji}v_{ji}) + (u_{ji}x_{ji})v_{ji} = x_{ji}(u_{ji}v_{ji}) - (u_{ji}v_{ji})x_{ji} = -[u_{ji}v_{ji}, x_{ji}] = D_i(u_{ji}v_{ji})x_{ji}$, hence (5-7) for $a = x_{ji}$

as well. Finally, for $a = x_{ij}$, we linearize (2-1), (2-2) and obtain

$$\begin{aligned} Dx_{ij} &= 2e_i[x_{ij}(u_{ji}v_{ji})] - 2(e_ix_{ij})(u_{ji}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= -2x_{ij}(u_{ji}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= -2(x_{ij}u_{ji})v_{ji} - 2(u_{ji}x_{ij})v_{ji} + 2u_{ji}(x_{ij}v_{ji}) - u_{ji}(x_{ij}v_{ji}) + (u_{ji}x_{ij})v_{ji} \\ &= u_{ji}(x_{ij}v_{ji}) - (u_{ji}x_{ij})v_{ji} = u_{ji}(v_{ji}x_{ij} + x_{ij}v_{ji}) - (u_{ji}x_{ij})v_{ji} = (u_{ji}v_{ji})x_{ij} \\ &= D_i(u_{ji}v_{ji})x_{ij} \end{aligned}$$

by (5-3), as desired. □

5.7. Corollary. *If $C_{ij} = C_{ji}^2$, then $\mathfrak{g}_i \subseteq \text{AssDer}(C)$.* □

5.8. Setting the stage. In order to continue with the proof of Thm. 5.1, we fix once and for all a reduced octonion algebra $C = \text{Zor}(M, \theta)$ over k as in 4.3. Our aim will be to describe in more detail the zero component of the derivation algebra $\mathfrak{g} = \text{Der}(C)$ relative to its e -grading, where e is one of the two standard idempotents

$$e = e_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \in C.$$

To this end, we first note

$$C_0 = \begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix} = ke_1 \oplus ke_2, \quad C_1 = C_{12} = \begin{bmatrix} 0 & M^+ \\ 0 & 0 \end{bmatrix}, \quad C_2 = C_{21} = \begin{bmatrix} 0 & 0 \\ M^- & 0 \end{bmatrix} \quad (5-8)$$

by 5.4. The properties of the vector product assembled in 4.3 ensure the relations

$$C_{ji}^2 = C_{ij} \quad (i, j \in \{1, 2\} \text{ distinct}). \quad (5-9)$$

As usual, we will write $\text{SL}(M)$ for the special linear group of $M = M^+$, and $\mathbf{SL}(M)$ for the affine group scheme over k given by

$$\mathbf{SL}(M)(R) := \text{SL}(M_R) \quad (R \in k\text{-alg}),$$

whose Lie algebra is

$$\mathfrak{sl}(M) := \{g \in \mathfrak{gl}(M) \mid \text{tr}(g) = 0\}.$$

The centralizer in the sense of [7, II, §1, 3.4] of e in $\mathbf{G} := \mathbf{Aut}(C)$ will be denoted by \mathbf{G}_0 . Then \mathbf{G}_0 acts like the identity on C_0 in all extensions since $e_2 = 1 - e_1$ and \mathbf{G} fixes the unit element. \mathbf{G}_0 is a subgroup scheme of \mathbf{G} whose Lie algebra is the subalgebra \mathfrak{g}_0 of \mathfrak{g} described in Prop. 5.5.

For $g \in \text{End}(M)$ let $g^* \in \text{End}(M^*)$ be defined by $\langle g^*(\zeta), x \rangle = \langle \zeta, g(x) \rangle$, for all $x \in M, \zeta \in M^*$. Then formula (4-13) implies

$$\langle g(x) \times g(y), g(z) \rangle = \theta(g(x) \wedge g(y) \wedge g(z)) = \det(g) \cdot \theta(x \wedge y \wedge z) = \det(g) \cdot \langle x \times y, z \rangle$$

for all $x, y, z \in M$. Now let $g \in \text{GL}(M)$ and replace z by $g^{-1}(z)$. Then

$$\langle g(x) \times g(y), z \rangle = \det(g) \cdot \langle x \times y, g^{-1}(z) \rangle = \det(g) \cdot \langle g^{*-1}(x \times y), z \rangle$$

holds for all z , so

$$g(x) \times g(y) = \det(g) \cdot g^{*-1}(x \times y) \quad (x, y \in M, g \in \text{GL}(M)). \quad (5-10)$$

The following theorem generalizes [10, Thm. 4] from fields to commutative rings. It says that the group scheme \mathbf{G}_0 defined above is of type A_2 , i.e., a twisted form of SL_3 .

5.9. Theorem. *With the notations of 5.8, there is an isomorphism*

$$\Phi : \text{SL}(M) \xrightarrow{\sim} \mathbf{G}_0$$

of group schemes where $\Phi(g)$ is given by

$$\Phi(g) \cdot \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} := \begin{bmatrix} \alpha_1 & g(u) \\ g^{*-1}(x) & \alpha_2 \end{bmatrix}$$

for all $g \in \text{SL}(M_R)$, $\alpha_1, \alpha_2 \in R$, $u \in M_R^+$, $x \in M_R^-$, $R \in k\text{-alg}$.

Proof. After a suitable base extension, it suffices to prove this for $R = k$. Clearly, Φ is a group monomorphism from $\text{SL}(M)$ to $\text{GL}(C)$. Given $g \in \text{SL}(M)$, a straightforward verification shows, using (5-10) and (4-14), that $\Phi(g)$ as defined above is an automorphism of C fixing e_1 . It therefore remains to show that $\Phi : \text{SL}(M) \rightarrow \mathbf{G}_0(k)$ is surjective, so let $\varphi \in \mathbf{G}_0(k)$ be an automorphism of C fixing e_1 . Then φ also fixes $e_2 = 1_C - e_1$ and stabilizes the C_{ij} . Hence (5-8) yields elements $g \in \text{GL}(M^+)$, $h \in \text{GL}(M^-)$ such that

$$\varphi \left(\begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} \right) = \begin{bmatrix} \alpha_1 & g(u) \\ h(x) & \alpha_2 \end{bmatrix} \quad (\alpha_i \in k, u \in M^+, x \in M^-).$$

Since φ leaves the norm of C invariant, (4-15) implies $h = g^{*-1}$, so it remains to prove $\det(g) = 1$. Now $\begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & v \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ u \times v & 0 \end{bmatrix}$ by (4-14), so φ multiplicative implies $g^{*-1}(u \times v) = g(u) \times g(v)$. But the linear map $\wedge^2 M^+ \rightarrow M^-$ determined by the vector product is surjective, forcing $\det(g) = 1$ by comparison with (5-10), as desired. \square

Passing to the level of Lie algebras in Thm. 5.9, we obtain

5.10. Corollary. *Notations being as in 5.8, the map*

$$\mathfrak{sl}(M) \xrightarrow{\sim} \mathfrak{g}_0, \quad f \mapsto D_0(f)$$

determined by

$$D_0(f) \begin{bmatrix} \alpha_1 & u \\ x & \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 & f(u) \\ f^\vee(x) & 0 \end{bmatrix}$$

for $f \in \mathfrak{sl}(M)$, $\alpha_i \in k$ ($i = 1, 2$), $u \in M^+$, $x \in M^-$, where $f^\vee := -f^*$, is an isomorphism of Lie algebras. □

Remark. Combining Cor. 5.10 with Prop. 5.5, we see that $\text{Der}(C)$ is a finitely generated projective module of rank 14, as it should be.

Proof of Thm. 5.1, step 2. We are now in a position to finish the proof of Thm. 5.1. By Prop. 5.5 and Cor. 5.7 combined with (5-9), it suffices to show that \mathfrak{g}_0 consists entirely of associator derivations, i.e.,

$$\mathfrak{g}_0 \subseteq \text{AssDer}(C). \tag{5-11}$$

Inspired by the proof of [1, Prop. 1], we do so by identifying $M^+ \otimes M^- = \text{End}_k(M^+)$ canonically via

$$(u \otimes x)(v) = \langle x, v \rangle u \quad (u, v \in M^+, x \in M^-). \tag{5-12}$$

Observe

$$\text{tr}(u \otimes x) = \langle u, x \rangle \quad (u \in M^+, x \in M^-), \tag{5-13}$$

let $u, v \in M^+$, $x \in M^-$ and put

$$a := \begin{bmatrix} 0 & u \\ 0 & 0 \end{bmatrix}, \quad b := \begin{bmatrix} 0 & 0 \\ x & 0 \end{bmatrix}, \quad c := \begin{bmatrix} 0 & v \\ 0 & 0 \end{bmatrix} \in C. \tag{5-14}$$

A direct computation, involving (2-12), (4-14) and (5-12) yields

$$[a, b] = \langle u, x \rangle (e_2 - e_1), \quad [L_a, R_b]e_1 = 0, \quad [L_a, R_b]c = \begin{bmatrix} 0 & h(v) \\ 0 & 0 \end{bmatrix}, \tag{5-15}$$

where

$$h = u \otimes x - \langle u, x \rangle \text{Id}_{M^+}. \tag{5-16}$$

By Cor. 5.10, every element of \mathfrak{g}_0 has the form $D_0(f)$ for some $f \in \mathfrak{sl}(M^+)$. Write $f = \sum u_i \otimes x_i$, $u_i \in M^+$, $x_i \in M^-$, and put

$$a_i := \begin{bmatrix} 0 & u_i \\ 0 & 0 \end{bmatrix}, \quad b_i := \begin{bmatrix} 0 & 0 \\ x_i & 0 \end{bmatrix} \in C.$$

Then (5-13)–(5-16) combine with Cor. 5.10 to show $\sum [a_i, b_i] = 0$ and $D_0(f) = \sum [L_{a_i}, R_{b_i}] \in \text{AssDer}(C)$. This completes the proof of (5-11). □

Acknowledgments

We thank Stefan Gille and Angelika Welte for useful comments. We are particularly grateful to Erhard Neher, who made important suggestions for improving an earlier version of the paper. He also drew our attention to McCrimmon's unpublished monograph [15] and was instrumental in making it accessible online to the international mathematical community. Kevin McCrimmon's kind permission to include his results on derivations of alternative algebras [15, A5.2] in the present investigation is also gratefully acknowledged. The first-named author is indebted to the Department of Theoretical Physics at the University of Innsbruck for its hospitality during the preparation of this paper.

Finally, we are especially grateful to the referee, who made numerous useful suggestions for improving an earlier version of the paper. In particular, Prop. 5.6 below, which greatly simplifies our original proof of Thm. 5.1, is due to him. More important still, the referee's penetrating analysis illuminating the intuitive background of inner derivations induced us to re-examine our own preconceptions of the subject, leading to the notion of derivation functor expounded in Section 3.

References

- [1] P. Alberca, A. Elduque, C. Martín, and F. J. Navarro, *On the Cartan-Jacobson theorem*, *J. Algebra* **250** (2002), 397–407.
- [2] E. Bannow, *Die Automorphismengruppen der Cayley-Zahlen*, *Abh. Math. Sem. Univ. Hamburg* **13** (1940), 240–256.
- [3] N. Bourbaki, *Elements of mathematics. Commutative algebra*, Hermann, Paris, 1972, Translated from the French.
- [4] ———, *Elements of mathematics. Algebra, Part I: Chapters 1-3*, Hermann, Paris, 1974, Translated from the French.
- [5] ———, *Algèbre homologique*, Masson, Paris, 1980.
- [6] É. Cartan, *Les groupes réels simples, finis et continus*, *Ann. Sci. École Norm. Sup.* **31** (1914), 263–355.
- [7] M. Demazure and P. Gabriel, *Groupes algébriques. Tome I*, Masson & Cie, Paris, 1970.
- [8] A. Grothendieck, *Éléments de géométrie algébrique. IV. Étude locale des schémas et des morphismes de schémas IV*, *Inst. Hautes Études Sci. Publ. Math.* (1967), no. 32, 361.
- [9] N. Jacobson, *Cayley numbers and normal simple Lie algebras of type G*, *Duke Math. J.* **5** (1939), 775–783.
- [10] ———, *Composition algebras and their automorphisms*, *Rend. Circ. Mat. Palermo* (2) **7** (1958), 55–80.
- [11] ———, *Exceptional Lie algebras*, *Lecture Notes in Pure and Applied Mathematics*, vol. 1, Marcel Dekker Inc., New York, 1971.
- [12] M.-A. Knus and M. Ojanguren, *Théorie de la descente et algèbres d'Azumaya*, Springer-Verlag, Berlin, 1974, *Lecture Notes in Mathematics*, Vol. 389.
- [13] J.-L. Loday, *Une version non commutative des algèbres de Lie: les algèbres de Leibniz*, *Enseign. Math.* (2) **39** (1993), no. 3-4, 269–293.

- [14] O. Loos, *Generically algebraic Jordan algebras over commutative rings*, J. Algebra **297** (2006), 474–529.
- [15] K. McCrimmon, *Alternative algebras*, <http://www.mathstat.uottawa.ca/~neher/Papers/alternative/>, 1980.
- [16] ———, *Nonassociative algebras with scalar involution*, Pacific J. Math. **116** (1985), 85–109.
- [17] H. P. Petersson, *Composition algebras over algebraic curves of genus zero*, Trans. Amer. Math. Soc. **337** (1993), 473–493.
- [18] N. Roby, *Lois polynomes et lois formelles en théorie des modules*, Ann. Sci. École Norm. Sup. (3) **80** (1963), 213–348.
- [19] R. D. Schafer, *Inner derivations of non-associative algebras*, Bull. Amer. Math. Soc. **55** (1949), 769–776.
- [20] ———, *An introduction to nonassociative algebras*, Pure and Applied Mathematics, Vol. 22, Academic Press, New York, 1966.
- [21] T. A. Springer and F. D. Veldkamp, *Octonions, Jordan algebras and exceptional groups*, Springer Monographs in Mathematics, Springer-Verlag, Berlin, 2000.
- [22] M. L. Thakur, *Cayley algebra bundles on \mathbf{A}_K^2 revisited*, Comm. Algebra **23** (1995), 5119–5130.
- [23] K. A. Zhevlakov, A. M. Slin'ko, I. P. Shestakov, and A. I. Shirshov, *Rings that are nearly associative*, Pure and Applied Mathematics, Vol. 104, Academic Press, New York, 1982.
- [24] M. Zorn, *Alternativkörper und quadratische Systeme*, Abh. Math. Sem. Univ. Hamburg **9** (1933), 395–402.

Communicated by Georgia Benkart

Received 2008-04-06 Revised 2008-09-26 Accepted 2008-10-26

ottmar.loos@uibk.ac.at

*Fakultät für Mathematik und Informatik,
FernUniversität in Hagen, D-58084 Hagen, Germany*

holger.petersson@fernuni-hagen.de

*Fakultät für Mathematik und Informatik,
FernUniversität in Hagen, D-58084 Hagen, Germany*

mracine@science.uottawa.ca

*Department of Mathematics and Statistics,
University of Ottawa, Ottawa, Ontario K1N 6N5, Canada*

On Oliver's p -group conjecture

David J. Green, László Héthelyi and Markus Lilienthal

Let S be a p -group for an odd prime p . B. Oliver conjectures that a certain characteristic subgroup $\mathfrak{X}(S)$ always contains the Thompson subgroup $J(S)$. We obtain a reformulation of the conjecture as a statement about modular representations of p -groups. Using this we verify Oliver's conjecture for groups where $S/\mathfrak{X}(S)$ has nilpotence class at most two.

1. Introduction

The recently introduced concept of a p -local finite group seeks to provide a treatment of the p -local structure of a finite group G which does not refer directly to the group G itself and yet retains enough information to construct the p -localisation of the classifying space BG . Ideally one could then associate a p -local classifying space to a p -block of G , and to certain exotic fusion systems. See the survey article by Broto, Levi and Oliver [2004] for an introduction to this area.

A key open question about p -local finite groups is whether or not there is a unique centric linking system associated to each saturated fusion system. Oliver showed that this would follow from a conjecture about higher limits (see [Oliver 2004, Conjecture 2.2]) and that for odd primes this higher limits conjecture would in turn follow from a purely group-theoretic conjecture:

Conjecture [Oliver 2004, Conjecture 3.9]. *Let S be a p -group for an odd prime p . Then*

$$J(S) \leq \mathfrak{X}(S),$$

where $J(S)$ is the Thompson subgroup generated by all elementary abelian p -subgroups whose rank is the p -rank of S , and $\mathfrak{X}(S)$ is the Oliver subgroup described in Section 2.

Our main result on Oliver's conjecture is:

Theorem 1.1. *Let S be a p -group for an odd prime p . If $S/\mathfrak{X}(S)$ has nilpotency class at most two, then S satisfies Oliver's conjecture.*

MSC2000: 20D15.

Keywords: p -group, characteristic subgroup, Thompson subgroup, p -local finite group, Replacement Theorem.

Héthelyi was supported by the Hungarian Scientific Research Fund, OTKA grant T049 841.

Remark. This subsumes all three cases of Oliver’s Proposition 3.7 in the first case $\mathfrak{X}(S) \geq J(S)$.

The proof of Theorem 1.1 depends on a reformulation of Oliver’s conjecture, for which we need to recall the terms F -module and offender. See for example [Meierfrankenfeld and Stellmacher 2006] for recent results about offenders.

Definition [Gorenstein et al. 1994, Definition 26.5]. Let G be a finite group and V a faithful $\mathbb{F}_p G$ -module. If there exists a nonidentity elementary abelian p -subgroup $E \leq G$ which satisfies the inequality $|E||C_V(E)| \geq |V|$, then V is called an F -module for G , and E an *offending subgroup*.

Remark. F -module is short for “failure of (Thompson) factorization module”. Another way to phrase the inequality is $\dim(V) - \dim(V^E) \leq \text{rank}(E)$.

We will always take G to be a nontrivial p -group. Hence the $\mathbb{F}_p G$ -module V is faithful if and only if it is faithful as a module for $\Omega_1(Z(G))$. We shall be interested in the stronger condition:

(PS) The restriction of V to each central order p subgroup has a nontrivial projective summand.

Remark. Projective and free are equivalent here. We are grateful to the referee for suggesting this formulation of the property. Another formulation is that every central order p element operates with minimal polynomial $(X - 1)^p$: equivalence follows from the standard properties of the Jordan normal form.

Theorem 1.2. *Let $G \neq 1$ be a finite p -group. Then Oliver’s conjecture holds for every finite p -group S with $S/\mathfrak{X}(S) \cong G$ if and only if G has no F -modules satisfying **(PS)**.*

Conjecture 1.3. *Let p be an odd prime and $G \neq 1$ a finite p -group. Then G has no F -modules which satisfy **(PS)**.*

Corollary 1.4. *Conjecture 1.3 is equivalent to Oliver’s Conjecture 3.9.*

We prove Theorem 1.1 by verifying Conjecture 1.3 for groups of class at most two. For this we need this result:

Definition (See [Glauberman 1972]). Let V be a faithful $\mathbb{F}_p G$ -module. A non-identity element $g \in G$ is called *quadratic* if $(g - 1)^2 V = 0$.

Theorem 1.5. *Suppose that p is an odd prime, G is a p -group of nilpotence class at most two, and V is a faithful $\mathbb{F}_p G$ -module. If G contains a quadratic element, then so does $\Omega_1(Z(G))$.*

Structure of the paper. We prove Theorem 1.2 and Corollary 1.4 in Section 2. In Section 3 we derive a consequence of the Replacement Theorem, Theorem 3.3. Then in Section 4 we prove Theorems 1.5 and 1.1. Finally in Section 5 we discuss a class three example which cannot be handled using Theorem 3.3.

2. The reformulation of Oliver's conjecture

For the convenience of the reader we start by recapping the definition and elementary properties of $\mathfrak{X}(S)$, as given in [Oliver 2004, §3].

Definition [Oliver 2004, Definition 3.1]. Let S be a p -group and $K \triangleleft S$ a normal subgroup. A Q -series leading up to K consists of a series of subgroups

$$1 = Q_0 \leq Q_1 \leq \dots \leq Q_n = K$$

such that each Q_i is normal in S , and such that

$$[\Omega_1(C_S(Q_{i-1})), Q_i; p - 1] = 1$$

holds for each $1 \leq i \leq n$. The unique largest normal subgroup of S which admits such a Q -series is called $\mathfrak{X}(S)$, the Oliver subgroup of S .

Lemma 2.1 (Oliver). *If $1 = Q_0 \leq Q_1 \leq \dots \leq Q_n = K$ is such a Q -series and $H \triangleleft G$ also admits a Q -series, then there is a Q -series leading up to HK which starts with Q_0, \dots, Q_n .*

Hence there is indeed a unique largest subgroup admitting a Q -series, and this subgroup $\mathfrak{X}(S)$ is characteristic in S . In addition, $\mathfrak{X}(S)$ is centric in S : recall that $P \leq S$ is centric if $C_S(P) = Z(P)$.

Proof. See [Oliver 2004, pp. 334–5]. □

Now we can start to derive the reformulation of Oliver's conjecture.

Lemma 2.2. *Let S be a finite p -group with $\mathfrak{X}(S) < S$. Then the induced action of $G := S/\mathfrak{X}(S)$ on $V := \Omega_1(Z(\mathfrak{X}(S)))$ satisfies (PS).*

Proof. Pick $g \in S$ such that $1 \neq g\mathfrak{X}(S) \in \Omega_1(Z(G))$. Then $\langle \mathfrak{X}(S), g \rangle \triangleleft S$ and so $[V, g; p - 1] \neq 1$, by maximality of $\mathfrak{X}(S)$. So the minimal polynomial of the action of g does not divide $(X - 1)^{p-1}$. But it has to divide $(X - 1)^p = X^p - 1$. So $(X - 1)^p$ is the minimal polynomial. This is the reformulation of (PS). □

Proof of Theorem 1.2. Suppose first that no F -module for G satisfies (PS), and that $S/\mathfrak{X}(S) \cong G$. Let us prove Oliver's Conjecture for G . By Lemma 2.2 the induced action of G on $V := \Omega_1(Z(\mathfrak{X}(S)))$ satisfies (PS), so by assumption there are no offending subgroups.

Let $E \leq S$ be an elementary abelian subgroup not contained in $\mathfrak{X}(S)$. It suffices for us to show that $\mathfrak{X}(S)$ contains an elementary abelian subgroup of greater rank

than E . We can split E up as $E = E_1 \times E_2 \times E_3$, with $E_1 = E \cap V \leq V^E$ and $E_1 \times E_2 = E \cap \mathfrak{X}(S)$. By assumption, $1 \neq E_3$ embeds in $S/\mathfrak{X}(S) \cong G$. As there are no offenders, we have $\dim(V) - \dim(V^{E_3}) > \text{rank}(E_3)$. But $V^{E_3} = V^E$. So $V \times E_2$ lies in $\mathfrak{X}(S)$ and has greater rank than E .

Conversely suppose that the $\mathbb{F}_p G$ -module V is an F -module and satisfies **(PS)**. Set S to be the semidirect product $S = V \rtimes G$ defined by this action. From Lemma 2.3 below we see that $V = \mathfrak{X}(S)$. As V is an F -module, there is an offender: an elementary abelian subgroup $1 \neq E \leq G$ with $\dim(V) - \dim(V^E) \leq \text{rank}(E)$. This means that $W := V^E \times E$ is an elementary abelian subgroup which does not lie in $V = \mathfrak{X}(S)$ but does have rank at least as great as that of $\mathfrak{X}(S)$. So $W \leq J(S)$ and therefore $J(S) \not\leq \mathfrak{X}(S)$. □

Lemma 2.3. *Suppose that V is an $\mathbb{F}_p G$ -module which satisfies **(PS)**. Let S be the semidirect product $S = V \rtimes G$ defined by this action. Then $V = \mathfrak{X}(S)$.*

Proof. First we prove that V is a maximal normal abelian subgroup of S : clearly it is abelian and normal. If A is a normal abelian subgroup strictly containing V , then $A = V \rtimes H$ for some nontrivial abelian $H \triangleleft G$. As H is nontrivial and normal it contains an order p element g of $Z(G)$. Since V satisfies **(PS)**, it follows that g acts on V with minimal polynomial $(X - 1)^p$. But that is a contradiction, as A is abelian. So V is indeed maximal normal abelian.

We now argue as in the proof of Oliver’s Lemma 3.2. Since V is maximal normal abelian, it is centric in S : for if not then $V < C_S(V) \triangleleft S$, and so $C_S(V)/V$ has nontrivial intersection with the centre of S/V . Picking an $x \in C_S(V)$ whose image in $C_S(V)/V$ is a nontrivial element of this intersection, we obtain a strictly larger normal abelian subgroup $\langle V, x \rangle$, a contradiction. Hence $\Omega_1(C_S(V)) = V$.

Moreover, since V is normal abelian and $p > 2$, there is a Q -series $1 < V$. So by Lemma 2.1 there is a Q -series leading up to $\mathfrak{X}(S)$ with $Q_1 = V$. If $V < \mathfrak{X}(S)$ then there is $Q_1 < Q_2 \triangleleft S$ with $[V, Q_2; p - 1] = 1$. But this cannot happen, because by the argument of the first paragraph of this proof there is a $g \in Q_2$ whose action on V has minimal polynomial $(X - 1)^p$. So $V = \mathfrak{X}(S)$. □

Proof of Corollary 1.4. Immediate from Theorem 1.2. If $\mathfrak{X}(S) = S$ then Oliver’s Conjecture holds automatically. □

3. The Replacement Theorem

We shall need the following lemma, which is a special case of the Replacement Theorem and its proof in [Huppert and Blackburn 1982, X, 3.3].

Lemma 3.1. *Suppose that $G \neq 1$ is elementary abelian, that V is a faithful $\mathbb{F}_p G$ -module, and that G contains no quadratic elements. Let us write*

$$T = \{(H, W) \mid H \leq G \text{ and } W \text{ is a subspace of } V^H\}.$$

Suppose that $(H, W) \in T$ with $H \neq 1$. Then there is $(K, U) \in T$ with $K < H$, $W \subsetneq U \subsetneq V$ and $|H \times W| = |K \times U|$.

Proof. Let us set

$$I = \{v \in V \mid (h - 1)v \in W \text{ for every } h \in H\},$$

$$J = \{v \in V \mid (h - 1)v \in I \text{ for every } h \in H\}.$$

If $1 \neq h \in H$ then $(h - 1)^2v \neq 0$ for some $v \in V$. Then $v \notin I$, for otherwise $(h - 1)v \in W$ and so $(h - 1)^2v = 0$. So $I \subsetneq V$, and therefore $W \subsetneq I \subsetneq J$ by the usual orbit length argument. Pick $v_0 \in J \setminus I$ and set U to be the subspace spanned by W and $\{(h - 1)v_0 \mid h \in H\}$. Set $K = \{h \in H \mid (h - 1)v_0 \in W\}$. So $U \supsetneq W$ by choice of v_0 . Also $U \subseteq I \subsetneq V$. If $h, h' \in H$ then

$$(hh' - 1)v_0 = (h - 1)v_0 + (h' - 1)v_0 + (h - 1)(h' - 1)v_0,$$

and so

$$(hh' - 1)v_0 \equiv (h - 1)v_0 + (h' - 1)v_0 \pmod{W}. \tag{3-1}$$

So $K \leq H$, and in fact $K < H$ by choice of v_0 . By (3-1) it also follows that $|H : K| = p^r$ for $r = \dim U - \dim W$. Finally $U \subseteq V^K$, for if $k \in K$ and $u \in U$, then

$$u = \sum_{h \in H} \lambda_h (h - 1)v_0 + w$$

for suitable $\lambda_h \in \mathbb{F}_p$, $w \in W$. So

$$(k - 1)u = \sum_{h \in H} \lambda_h (h - 1)(k - 1)v_0 = 0,$$

since $(k - 1)v_0 \in W \subseteq V^H$. □

Corollary 3.2. *Suppose as in Lemma 3.1 that $(H, W) \in T$ and $H \neq 1$. Then $|H \times W| < |V|$.*

Proof. By induction on $|H|$. By the lemma we may reduce $|H|$ whilst keeping $|H \times W|$ constant. This process only stops when we arrive at (K, U) with $K = 1$. But $U \subsetneq V$ by the lemma. □

The following result is presumably well known to those familiar with Thompson factorization.

Theorem 3.3. *Suppose that p is an odd prime, G is a finite group, V is a faithful $\mathbb{F}_p G$ -module, and $E \leq G$ is a nonidentity elementary abelian p -subgroup. If E is an offender, then it must contain a quadratic element.*

Proof. Without loss of generality $E = G$. Apply Corollary 3.2 to the pair

$$(G, V^G) \in T. \tag{3-1} \quad \square$$

Remark. Pursuing this direction further, it might be worthwhile to investigate potential applications of the $P(G, V)$ -theorem in the theory of p -local finite groups. The properties of the Thompson subgroup $J(S)$ which Chermak describes in his comments on the motivation for the $P(G, V)$ -theorem [Chermak 1999, Remark 2] are the same properties which led to $J(S)$ featuring in Oliver’s conjecture. And Timmesfeld’s Replacement Theorem plays an important part in the proof of the $P(G, V)$ -theorem.

4. Nilpotence class at most two

We can now start work on the proof of Theorem 1.1.

Lemma 4.1. *Suppose that p is an odd prime, that $G \neq 1$ is a finite p -group, and that V is a faithful $\mathbb{F}_p G$ -module. Suppose that $A, B \in G$ are such that $C := [A, B]$ is a nontrivial element of $C_G(A, B)$. If C is nonquadratic, then so are A and B .*

Proof. By symmetry it suffices to prove that B is nonquadratic. So suppose that B is quadratic. Denote by α, β, γ the action matrices on V of $A - 1, B - 1$ and $C - 1$ respectively.

By assumption we have $\gamma^2 \neq 0$ and $\beta^2 = 0$. As C commutes with A and B , we have $\alpha\gamma = \gamma\alpha$ and $\beta\gamma = \gamma\beta$. Since $[A, B] = C$, we have $AB = BAC$ and therefore

$$\alpha\beta - \beta\alpha = \gamma(1 + \beta + \alpha + \beta\alpha). \tag{4-1}$$

Evaluating $\beta \cdot (4-1) \cdot \beta$, we deduce that $\gamma\beta\alpha\beta = 0$. So when we evaluate $\beta \cdot (4-1) + (4-1) \cdot \beta$, we find that $\gamma(2\beta + \beta\alpha + \alpha\beta) = 0$. Let us write $\lambda = -\frac{1}{2}$ and $\delta = \gamma\beta$. Then we have

$$\delta = \lambda(\delta\alpha + \alpha\delta).$$

From this one sees by induction upon $r \geq 1$ that

$$\delta = \lambda^r \sum_{s=0}^r \binom{r}{s} \alpha^s \delta \alpha^{r-s}.$$

Since the order of A is a power of p , it follows that $(A - 1)$ and its action matrix α are nilpotent. From this we deduce that $\delta = 0$, that is $\gamma\beta = 0$. Applying this to $\gamma \cdot (4-1)$ we see that $\gamma^2(1 + \alpha) = 0$. As α is nilpotent it follows that $\gamma^2 = 0$, a contradiction. So $\beta^2 \neq 0$ after all. □

Proof of Theorem 1.5. We suppose that $\Omega_1(Z(G))$ has no quadratic elements, and show that G has none either. Suppose $1 \neq B \in Z(G)$. Then there is an $r \geq 0$ with $1 \neq B^{p^r} \in \Omega_1(Z(G))$. So B^{p^r} is not quadratic. Hence $(B - 1)^{2p^r} = (B^{p^r} - 1)^2$ has nonzero action. So $(B - 1)^2$ has nonzero action, and $Z(G)$ contains no quadratic elements.

If $B \notin Z(G)$ then the nilpotency class is two and there is an element $A \in G$ with $1 \neq [A, B] \in Z(G)$. So $(B - 1)^2$ has nonzero action by Lemma 4.1. \square

Corollary 4.2. *Suppose that p is an odd prime, $G \neq 1$ a finite p -group and V an $\mathbb{F}_p G$ -module which satisfies (PS). If the nilpotence class of G is at most two then V cannot be an F -module.*

Proof. As p is odd, condition (PS) means that there are no quadratic elements in $\Omega_1(Z(G))$. Then Theorem 1.5 says that there are no quadratic elements in G . So by Theorem 3.3 there are no offenders. \square

Proof of Theorem 1.1. Follows from Corollary 4.2 and Theorem 1.2 if $\mathfrak{X}(S) < S$. If $\mathfrak{X}(S) = S$ then there is nothing to prove. \square

5. A class 3 example

Theorem 1.5 was a key step in the proof of Theorem 1.1. We now give an example which shows that Theorem 1.5 does not apply to groups of nilpotence class three.

Let G be the semidirect product $G = K \rtimes L$, where the $K = \mathbb{F}_3^3$ is elementary abelian of order 3^3 , $L = \langle A \rangle$ is cyclic of order 3, and the action of L on $v \in K$ is given by

$$AvA^{-1} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot v.$$

Observe that G is isomorphic to the wreath product $C_3 \wr C_3$, as the action of A permutes the following basis of K cyclically: $(0, 0, 1), (0, 1, 1), (1, 2, 1)$.

Setting $B = (0, 0, 1), C = (0, 1, 0)$ and $D = (1, 0, 0)$ we obtain the following presentation of G , where we take $[A, B]$ to mean $ABA^{-1}B^{-1}$.

$$G = \left\langle A, B, C, D \mid \begin{array}{l} A^3 = B^3 = C^3 = D^3 = 1, \quad D \text{ central,} \\ [B, C] = 1, \quad [A, B] = C, \quad [A, C] = D \end{array} \right\rangle,$$

From this we deduce that matrices $\alpha, \beta, \gamma, \delta \in M_n(\mathbb{F}_3)$ induce a representation $\rho: G \rightarrow GL_n(\mathbb{F}_3)$ with

$$\rho(A) = 1 + \alpha, \quad \rho(B) = 1 + \beta, \quad \rho(C) = 1 + \gamma, \quad \rho(D) = 1 + \delta,$$

if and only if the following relations are satisfied, where $[\alpha, \beta]$ now of course means $\alpha\beta - \beta\alpha$:

$$\begin{aligned} \alpha^3 &= \beta^3 = \gamma^3 = \delta^3 = 0, \\ [\alpha, \delta] &= [\beta, \delta] = [\gamma, \delta] = [\beta, \gamma] = 0, \\ [\alpha, \beta] &= \gamma(1 + \beta)(1 + \alpha), \quad [\alpha, \gamma] = \delta(1 + \gamma)(1 + \alpha). \end{aligned} \tag{5-1}$$

Now we consider what it means for such a representation to satisfy **(PS)**. Here,

$$Z(G) = \langle D \rangle$$

is cyclic of order 3. So we need both $(\rho(D) - 1)^2$ and $(\rho(D^2) - 1)^2$ to be nonzero. That is, δ^2 and $(\delta^2 + 2\delta)^2 = \delta^2(1 + \delta + \delta^2)$ should both be nonzero. But $1 + \delta + \delta^2$ is invertible, since δ is nilpotent.

We deduce therefore that matrices $\alpha, \beta, \gamma, \delta \in GL_n(\mathbb{F}_3)$ induce a representation of G satisfying **(PS)** if and only if they satisfy the inequality

$$\delta^2 \neq 0 \tag{5-2}$$

in addition to (5-1).

Using GAP [2007] we obtained the following matrices in $GL_8(\mathbb{F}_3)$. The reader is invited to check¹ that they satisfy the relations (5-1) and (5-2).

$$\delta = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \gamma = \begin{pmatrix} 0 & 0 & 1 & 1 & 2 & 2 & 2 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

$$\beta = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \alpha = \begin{pmatrix} 2 & 2 & 0 & 2 & 0 & 1 & 0 & 1 \\ 1 & 1 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Observe that $\beta^2 = 0$. So although this module satisfies **(PS)**, the elementary abelian subgroups $\langle B \rangle$ and $\langle B, C, D \rangle$ both contain B , a quadratic element. So we must find another way to show that they are not offenders: Theorem 3.3 does not apply.

Remark. More generally, we are not currently able to decide Conjecture 1.3 either way for the wreath product group $H \wr C_3$, where the group H on the bottom is an elementary abelian 3-group.

¹See <http://users.minet.uni-jena.de/~green/Documents/matTest.g> for a GAP script that performs these checks.

Acknowledgements

We are grateful to George Glauberman for generously sharing his background knowledge with us, and to the referee for advice concerning terminology.

References

- [Broto et al. 2004] C. Broto, R. Levi, and B. Oliver, “The theory of p -local groups: a survey”, pp. 51–84 in *Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic K-theory* (Evanston, IL, 2002), edited by P. Goerss and S. Priddy, Contemp. Math. **346**, Amer. Math. Soc., Providence, RI, 2004. MR 2005h:20040 Zbl 1063.55013
- [Chermak 1999] A. Chermak, “Quadratic action and the $P(G, V)$ -theorem in arbitrary characteristic”, *J. Group Theory* **2**:1 (1999), 1–13. MR 99m:20062 Zbl 0940.20016
- [GAP 2007] *GAP – Groups, Algorithms, and Programming*, Version 4.4.10, The GAP Group, 2007, Available at <http://www.gap-system.org>.
- [Glauberman 1972] G. Glauberman, “Quadratic elements in unipotent linear groups”, *J. Algebra* **20** (1972), 637–654. MR 49 #9079 Zbl 0231.20015
- [Gorenstein et al. 1994] D. Gorenstein, R. Lyons, and R. Solomon, *The classification of the finite simple groups*, Mathematical Surveys and Monographs **40**, American Mathematical Society, Providence, RI, 1994. MR 95m:20014 Zbl 0816.20016
- [Huppert and Blackburn 1982] B. Huppert and N. Blackburn, *Finite groups, III*, Grundlehren der Mathematischen Wissenschaften **243**, Springer, Berlin, 1982. MR 84i:20001b Zbl 0514.20002
- [Meierfrankenfeld and Stellmacher 2006] U. Meierfrankenfeld and B. Stellmacher, “The other $P(G, V)$ -theorem”, *Rend. Sem. Mat. Univ. Padova* **115** (2006), 41–50. MR 2007d:20056
- [Oliver 2004] B. Oliver, “Equivalences of classifying spaces completed at odd primes”, *Math. Proc. Cambridge Philos. Soc.* **137**:2 (2004), 321–347. MR 2006g:55010 Zbl 1077.55006

Communicated by Dave Benson

Received 2008-04-17 Revised 2008-08-14 Accepted 2008-09-19

David.Green@uni-jena.de *Mathematical Institute, Friedrich-Schiller-Universität Jena, 07737 Jena, Germany*
<http://users.minet.uni-jena.de/~green/index-en.php>

fobaba@t-online.hu *Department of Algebra, Budapest University of Technology and Economics, Budapest, Pf. 91, H-1521, Hungary*

mllf@gmx.de *FB Wirtschaftswissenschaften, Johann-Wolfgang-Goethe-Universität, House of Finance, Grüneburgplatz 1, 60325 Frankfurt (Main), Germany*

The number of 2×2 integer matrices having a prescribed integer eigenvalue

Greg Martin and Erick Wong

What is the probability that an integer matrix chosen at random has a particular integer as an eigenvalue, or an integer eigenvalue at all? For a random real matrix, what is the probability of there being a real eigenvalue in a particular interval? This paper solves these questions for 2×2 matrices, after specifying the probability distribution suitably.

1. Introduction

Random matrices arise in many mathematical contexts, and it is natural to ask about the properties that such matrices satisfy. If we choose a matrix with integer entries at random, for example, we would like to know the probability that it has a particular integer as an eigenvalue, or an integer eigenvalue at all. Similarly, if we choose a matrix with real entries at random, we would like to know the probability that it has a real eigenvalue in a particular interval. Certainly the answer depends on the probability distribution from which the matrix entries are drawn.

In this paper, we are primarily concerned with uniform distribution, so for both integer-valued and real-valued cases we must restrict the entries to a bounded interval. In [Martin and Wong 2008], we showed that random $n \times n$ matrices of integers almost never have integer eigenvalues. An explicit calculation by Hetzel, Liew and Morrison [Hetzel et al. 2007] shows that a 2×2 matrix with entries independently chosen uniformly from $[-1, 1]$ has real eigenvalues with probability $\frac{49}{72}$. This calculation gives hope that our more precise questions about eigenvalues of a particular size might be accessible in the 2×2 setting. The purpose of this paper is to resolve these questions, once we make them suitably precise.

For an integer $k \geq 1$, let $\mathcal{M}_2(k)$ denote the uniform probability space of 2×2 matrices of integers with absolute value at most k . Note that

$$|\mathcal{M}_2(k)| = (2k + 1)^4 = 16k^4 + O(k^3).$$

MSC2000: primary 15A36, 15A52; secondary 11C20, 15A18, 60C05.

Keywords: random matrix, eigenvalue, integer eigenvalue, integer matrix, distribution of eigenvalues.

We will obtain exact asymptotics for the number of matrices in $\mathcal{M}_2(k)$ having integer eigenvalues and, more precisely, for the number of matrices with a given integer eigenvalue λ .

For any integer λ , define

$$\mathcal{M}_2^\lambda(k) = \{M \in \mathcal{M}_2(k) : \lambda \text{ is an eigenvalue of } M\},$$

and let

$$\mathcal{M}_2^{\mathbb{Z}}(k) = \bigcup_{\lambda \in \mathbb{Z}} \mathcal{M}_2^\lambda(k).$$

Theorem 1. *Define the function $V : [-2, 2] \rightarrow \mathbb{R}$ by $V(-\delta) = V(\delta)$ and*

$$V(\delta) = \begin{cases} 4 - 2\delta - \delta^2 + \delta^2 \log(1 + \delta) - 2(1 - \delta) \log(1 - \delta) & \text{if } 0 \leq \delta < 1, \\ 1 + \log 2 & \text{if } \delta = 1, \\ 4 - 2\delta - \delta^2 + \delta^2 \log(\delta + 1) + 2(\delta - 1) \log(\delta - 1) & \text{if } 1 < \delta \leq \sqrt{2}, \\ \delta^2 - 2\delta - (\delta^2 - 2\delta + 2) \log(\delta - 1) & \text{if } \sqrt{2} < \delta \leq 2 \end{cases} \quad (1)$$

(where \log is the natural logarithm). Then for any integer λ between $-2k$ and $2k$,

$$|\mathcal{M}_2^\lambda(k)| = \frac{24V(\lambda/k)}{\pi^2} k^2 \log k + O(k^2), \quad (2)$$

where the implied constant is absolute. On the other hand, if $|\lambda| > 2k$ then $\mathcal{M}_2^\lambda(k)$ is empty.

We remark that the function $V(\delta)$ is continuous and, with the exception of the points of infinite slope at $\delta = \pm 1$, differentiable everywhere (even at $\delta = \pm 2$, if we imagine that $V(\delta)$ is defined to be 0 when $|\delta| > 2$). Notice that (2) is technically not an asymptotic formula when λ is extremely close to $\pm 2k$, because then the value of $V(\lambda/k)$ can have order of magnitude $1/\log k$ or smaller, making the “main term” no bigger than the error term. However, (2) is truly an asymptotic formula for $|\lambda| < 2k - \psi(k)k/(\log k)^{1/3}$, where $\psi(k)$ is any function tending to infinity (the exponent $\frac{1}{3}$ arises because $V(\delta)$ approaches 0 cubically as δ tends to 2 from below).

By summing the formula (2) over all possible values of λ , we obtain an asymptotic formula for $|\mathcal{M}_2^{\mathbb{Z}}(k)|$. We defer the details of the proof to Section 3.

Corollary 2. *Let*

$$C = \frac{7\sqrt{2} + 4 + 3 \log(\sqrt{2} + 1)}{3\pi^2} \approx 0.55873957.$$

The probability that a randomly chosen matrix in $\mathcal{M}_2(k)$ has integer eigenvalues is asymptotically $C(\log k)/k$. More precisely,

$$|\mathcal{M}_2^{\mathbb{Z}}(k)| = 16Ck^3 \log k + O(k^3).$$

If $M \in \mathcal{M}_2(k)$ has eigenvalue λ , then the scaled matrix $k^{-1}M$ has eigenvalue λ/k , which is the argument of V that appears on the right-hand side of (2). Thus one interpretation of Theorem 1 is that for large k , the rational eigenvalues of $k^{-1}M$ tend to be distributed like the function V .

Note that the entries of $k^{-1}M$ are sampled uniformly from a discrete, evenly-spaced subset of $[-1, 1]$. As $k \rightarrow \infty$ this probability distribution converges in law to the uniform distribution on the interval $[-1, 1]$. Let $\mathcal{M}_2([-1, 1])$ denote the probability space of all 2×2 matrices whose entries are independent random variables drawn from this distribution. One might ask whether the distribution given by Theorem 1 is just a discrete approximation to the distribution of eigenvalues in $\mathcal{M}_2([-1, 1])$; the answer, perhaps surprisingly, is no. The next theorem provides this latter distribution.

Theorem 3. *Define $W(\delta)$ to be the density function for real eigenvalues of matrices in $\mathcal{M}_2([-1, 1])$: if M is a randomly chosen matrix from $\mathcal{M}_2([-1, 1])$, then the expected number of eigenvalues of M in the interval $[s, t]$ is $\int_s^t W(\delta) d\delta$. Then $W(-\delta) = W(\delta)$ and*

$$W(\delta) = \begin{cases} \frac{80+20\delta+90\delta^2+52\delta^3-107\delta^4}{144(1+\delta)} - \frac{1}{4}\delta(1-\delta^2)\log(1+\delta) \\ \quad - \frac{1}{12}(5-7\delta+8\delta^2)(1-\delta)\log(1-\delta) & \text{if } 0 \leq \delta \leq 1, \\ \frac{\delta(20+10\delta-12\delta^2-3\delta^3)}{16(1+\delta)} + \frac{1}{4}\delta(\delta^2-1)\log(\delta+1) \\ \quad + \frac{1}{4}(3\delta-1)(\delta-1)\log(\delta-1) & \text{if } 1 \leq \delta \leq \sqrt{2}, \\ \frac{\delta(\delta-2)(2-6\delta+3\delta^2)}{16(\delta-1)} - \frac{1}{4}(\delta-1)^3\log(\delta-1) & \text{if } \sqrt{2} \leq \delta \leq 2, \\ 0 & \text{if } \delta \geq 2. \end{cases}$$

As in the case of $V(\delta)$, the function $W(\delta)$ is continuous and differentiable everywhere, with the exception of the points of infinite slope at $\delta = \pm 1$. (The value $W(1) = \frac{15}{32}$ makes the function continuous there, although the value of a density function at a single point is irrelevant to the probability distribution.) It also shares the same cubic decay as δ tends to 2 from below. However, there are obvious qualitative differences between the functions V and W . In Figure 1 we plot V and W on the same axes, normalized so that the area under each is 2 (these normalized versions are denoted $U^{\mathbb{Z}}$ and $U^{\mathbb{R}}$ in our earlier paper [Martin and Wong 2008]). In the case of $\mathcal{M}_2(k)$, this normalization corresponds to conditioning on having integer eigenvalues, that is, scaling by the probability $C(\log k)/k$ from Corollary 2. For $\mathcal{M}_2([-1, 1])$ we are conditioning on having real eigenvalues, which occurs with probability $\frac{49}{72}$ (this can be obtained by integrating $W(\delta)$, analogously to the proof of Corollary 2; the computation in [Hetzl et al. 2007] is more direct).

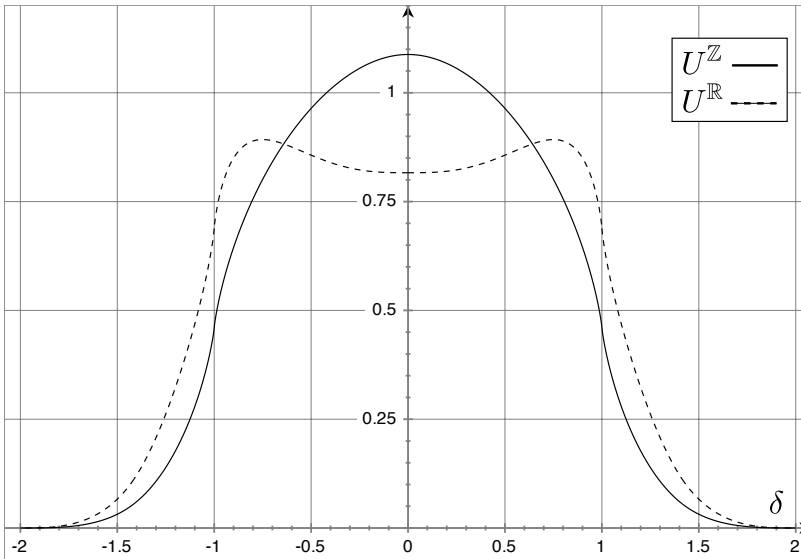


Figure 1. Graph of $U^{\mathbb{Z}}$ (V normalized) versus $U^{\mathbb{R}}$ (W normalized).

Note that the distribution $W(\delta)$ is bimodal, having its maxima at

$$\delta \approx \pm 0.75030751.$$

Thus, a random matrix in $\mathcal{M}_2([-1, 1])$ is more likely to have an eigenvalue of magnitude near $\frac{3}{4}$ than one of magnitude near 0. We expect this would still hold if we were to condition on matrices in $\mathcal{M}_2([-1, 1])$ having rational eigenvalues, since any matrix with real eigenvalues is a small perturbation from one with rational eigenvalues. That this is not true for $V(\delta)$ shows that the eigenvalue distribution of $\mathcal{M}_2(k)$ is not purely the result of magnitude considerations but also encodes some of the arithmetic structure of the integers up to k .

We remark that Theorem 1 can also be obtained from a powerful result of Katznelson [1993]. Let \mathcal{B} be a convex body containing the origin in \mathbb{R}^4 , and embed the set of 2×2 integer matrices

$$\begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

as lattice points $(a, b, c, d) \in \mathbb{Z}^4$. Then [Katznelson 1993, Theorem 1] gives an asymptotic formula for the number of singular integer matrices inside the dilate $k \cdot \mathcal{B}$. Taking

$$\mathcal{B} = [-1, 1]^4$$

then yields an asymptotic formula for $|M_2^0(k)|$, and more generally one can obtain $|M_2^\lambda(k)|$ by adding and subtracting appropriate shifts of \mathcal{B} . The asymptotic formula in [Katznelson 1993] is defined in terms of an unusual singular measure

supported on the Zariski-closed subset of \mathbb{R}^4 corresponding to singular matrices. The explicit computation of this measure is roughly analogous to our case-by-case considerations in Section 4, modulo the significant complications of carrying error terms. Our techniques are more elementary, but Katznelson’s results apply in theory to matrices of any size, whereas our methods become unwieldy even for 3×3 matrices.

In the case of $n \times n$ matrices with entries independently chosen from a Gaussian distribution, a great deal more is known. Edelman [1997] has computed the exact distributions of the real and complex eigenvalues for any n , as well as the number of real eigenvalues (for instance, the probability of having all n eigenvalues real is precisely $2^{-n(n-1)/4}$). As $n \rightarrow \infty$, the real eigenvalues, suitably rescaled by a factor of $1/\sqrt{n}$, converge to the uniform distribution on $[-1, 1]$. Similarly, the complex eigenvalues converge to the “circular law” predicted by Girko [1984], namely the uniform distribution on the unit disk centered at the origin. Very recently, Tao and Vu [2008] have shown that the circular law is universal: one can replace the Gaussian distribution by an arbitrary distribution with mean 0 and variance 1. Similar results have been established for random symmetric matrices with entries independently chosen from a Gaussian distribution (Wigner law) or from other distributions.

Those who are interested in the connections between analytic number theory and random matrix theory might wonder whether those connections are related to the present paper. The matrices in that context, however, are selected from classical matrix groups, such as the group of $n \times n$ Hermitian matrices, randomly according to the Haar measures on the groups. The relationship to our results is therefore minimal.

2. Preliminaries about matrices

We begin with some elementary observations about 2×2 matrices that will simplify our computations. The first lemma explains why the functions V and W are supported only on $[-2, 2]$.

Lemma 4. *Any eigenvalue of a matrix in $\mathcal{M}_2(k)$ is bounded in absolute value by $2k$. Any eigenvalue of a matrix in $\mathcal{M}_2([-1, 1])$ is bounded in absolute value by 2.*

Proof. We invoke Gershgorin’s circle theorem [1931], a standard result in spectral theory: let $M = (m_{ij})$ be an $n \times n$ matrix, and let $D(z, r)$ denote the disk of radius r around the complex number z . Then Gershgorin’s theorem says that all of the eigenvalues of M must lie in the union of the disks

$$D\left(m_{11}, \sum_{\substack{1 \leq j \leq n \\ j \neq 1}} |m_{1j}|\right), \quad D\left(m_{22}, \sum_{\substack{1 \leq j \leq n \\ j \neq 2}} |m_{2j}|\right), \quad \dots, \quad D\left(m_{nn}, \sum_{\substack{1 \leq j \leq n \\ j \neq n}} |m_{nj}|\right).$$

In particular, if all of the entries of M are bounded in absolute value by B , then all the eigenvalues are bounded in absolute value by nB . \square

The key to the precise enumeration of $\mathcal{M}_2^\lambda(k)$ is the simple structure of singular integer matrices:

Lemma 5. *For any singular matrix $M \in \mathcal{M}_2(k)$, either at least two entries of M equal zero, or else there exist nonzero integers a, b, c, d with $(a, b) = 1$ such that*

$$M = \begin{pmatrix} ac & bc \\ ad & bd \end{pmatrix}. \tag{3}$$

This representation of M is unique up to replacing each of a, b, c, d by its negative.

Proof. If one of the entries of M equals zero, then a second one must equal zero as well for the determinant to vanish. Otherwise, given

$$M = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$$

with none of the m_{ij} equal to zero, define $c = (m_{11}, m_{12})$, and set $a = m_{11}/c$ and $b = m_{12}/c$, so that $(a, b) = 1$. Since M is singular, the second row of m must be a multiple of the first row, that is, there exists a real number d such that $m_{21} = ad$ and $m_{22} = bd$. And since a and b are relatively prime, d must in fact be an integer.

This argument shows that every such matrix has one such representation. If

$$M = \begin{pmatrix} a'c' & b'c' \\ a'd' & b'd' \end{pmatrix}$$

is another such representation, then $(a', b') = 1$ implies $(a'c', b'c') = |c'|$, which shows that $|c'| = c$; the equalities $|a'| = |a|$, $|b'| = |b|$, and $|d'| = |d|$ follow quickly. \square

For a 2×2 matrix $M = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$, we define

$$\text{disc } M = (\text{tr } M)^2 - 4 \det M = (a - d)^2 + 4bc.$$

It is easily seen that this is the discriminant of the characteristic polynomial of M . We record the following elementary facts, which will be useful in the proofs of Lemma 7 and Proposition 13.

Lemma 6. *Let M be a 2×2 matrix with real entries*

- (a) *M has repeated eigenvalues if and only if $\text{disc } M = 0$.*
- (b) *M has real eigenvalues if and only if $\text{disc } M \geq 0$.*
- (c) *$\det M < 0$ if and only if M has two real eigenvalues of opposite sign.*
- (d) *If $\det M > 0$ and $\text{disc } M \geq 0$, then the eigenvalues of M have the same sign as $\text{tr } M$.*

Proof. Let λ_1, λ_2 denote the eigenvalues of M , so

$$\operatorname{tr} M = \lambda_1 + \lambda_2, \quad \det M = \lambda_1 \lambda_2 \quad \text{and} \quad \operatorname{disc} M = (\lambda_1 - \lambda_2)^2,$$

each of which is real. Parts (a), (b) and (d) follow immediately from these observations, and part (c) from the fact that $\lambda_2 = \overline{\lambda_1}$ if λ_1, λ_2 are complex. \square

The next lemma gives a bound for the probability of a matrix having repeated eigenvalues. It is natural to expect this probability to converge to 0 as k increases, and indeed such a result was obtained in [Hetzel et al. 2007] for matrices of arbitrary size. We give a simple proof of a stronger bound for the 2×2 case, as well as an analogous qualitative statement for real matrices which will be helpful in the proof of Theorem 3.

Lemma 7. *The number of matrices in $\mathcal{M}_2(k)$ with a repeated eigenvalue, for every $\varepsilon > 0$, is $\ll_\varepsilon k^{2+\varepsilon}$. The probability that a random matrix in $\mathcal{M}_2([-1, 1])$ has a repeated eigenvalue or is singular is 0.*

Proof. By Lemma 6 (a), the 2×2 matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ has a double eigenvalue if and only if $4bc = -(a - d)^2$. For matrices in $\mathcal{M}_2([-1, 1])$ this is easily seen to be a zero-probability event, as is the event that $\det M = ad - bc = 0$.

For matrices in $\mathcal{M}_2(k)$, we enumerate how many can satisfy $4bc = -(a - d)^2$. If $a = d$ then there are $4k + 1$ choices for $b, c \in \{-k, \dots, k\}$; otherwise there are at most $2\tau((a - d)^2/4)$ choices if $a \equiv d \pmod{2}$ and no choices otherwise. (Here $\tau(n)$ is the number-of-divisors function; the factor of 2 comes from the fact that b and c can be positive or negative, while the “at most” is due to the fact that not all factorizations of $-(a - d)^2$ result in two factors not exceeding k .) Therefore the number of matrices in $\mathcal{M}_2(k)$ with a repeated eigenvalue is at most

$$\sum_{|a| \leq k} (4k + 1) + 2 \sum_{|a| \leq k} \sum_{\substack{|d| \leq k, d \neq a \\ a \equiv d \pmod{2}}} \tau\left(\frac{(a-d)^2}{4}\right) \ll_\varepsilon k^{2+\varepsilon},$$

where the inequality follows from $(a - d)^2/4 \leq k^2$ and the well-known fact that $\tau(n) \ll_\varepsilon n^\varepsilon$ for any $\varepsilon > 0$ — see for instance [Montgomery and Vaughan 2007, p. 56]. \square

3. Enumeration theorems for integer eigenvalues

Let $\mu(n)$ be the Möbius function, characterized by the identity

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{if } n = 1, \\ 0 & \text{if } n > 1. \end{cases} \tag{4}$$

The well-known Dirichlet series identity $1/\zeta(s) = \sum_{n=1}^{\infty} \mu(n)n^{-s}$ is valid for $\text{Re } s > 1$ (see [Montgomery and Vaughan 2007, Corollary 1.10], for example). In particular,

$$\sum \frac{\mu(n)}{n^2} = \frac{6}{\pi^2},$$

and we can estimate the tail of this series (using $|\mu(n)| \leq 1$) to obtain the quantitative estimate

$$\sum_{d \leq k} \frac{\mu(d)}{d^2} = \frac{6}{\pi^2} + O\left(\frac{1}{k}\right). \tag{5}$$

Lemma 8. *For nonzero integers a, b and parameters k, λ , define the function*

$$N_{k,\lambda}(a, b) = \#\{(c, d) \in \mathbb{Z}^2, c \neq 0, d \neq 0: |ac + \lambda|, |bc|, |ad|, |bd + \lambda| \leq k\}. \tag{6}$$

Then

$$|\mathcal{M}_2^\lambda(k)| = 4 \sum_{d \leq k} \mu(d) \sum_{1 \leq \alpha < \beta \leq k/d} N_{k,\lambda}(d\alpha, d\beta) + O(k^2),$$

where the implied constant is independent of λ and k .

Proof. Fix an integer $0 \leq \lambda \leq 2k$, and let $M \in \mathcal{M}_2^\lambda(k)$, so that $M - \lambda I$ is singular. By Lemma 5, either at least two entries of $M - \lambda I$ equal zero, or else $M - \lambda I$ has exactly two representations of the form (3). In the former case, there are $2k + 1$ choices for each of the two potentially nonzero entries, hence $O(k^2)$ such matrices in total (even taking into account the several different choices of which two entries are nonzero). In the latter case, there are exactly two corresponding quadruples a, b, c, d of integers as in Lemma 5. Taking into account that each entry of M must be at most k in absolute value, we deduce that

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= \frac{1}{2} \#\{(a, b, c, d) \in \mathbb{Z}^4 : a, b, c, d \neq 0, (a, b) = 1, \\ &\quad |ac + \lambda|, |bc|, |ad|, |bd + \lambda| \leq k\} + O(k^2) \\ &= \frac{1}{2} \sum_{\substack{1 \leq |a|, |b| \leq k \\ (a, b) = 1}} N_{k,\lambda}(a, b) + O(k^2). \end{aligned}$$

Because of the symmetries $N_{k,\lambda}(a, b) = N_{k,\lambda}(-a, b) = N_{k,\lambda}(a, -b) = N_{k,\lambda}(-a, -b)$, we have

$$|\mathcal{M}_2^\lambda(k)| = 2 \sum_{\substack{1 \leq a, b \leq k \\ (a, b) = 1}} N_{k,\lambda}(a, b) + O(k^2).$$

The only term in the sum where $a = b$ is the term $a = b = 1$, and for all other terms we can invoke the additional symmetry $N_{k,\lambda}(a, b) = N_{k,\lambda}(b, a)$, seen to be valid by switching the roles of c and d in the definition (6) of $N_{k,\lambda}(a, b)$. We obtain

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4 \sum_{\substack{1 \leq a < b \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + 2N_{k,\lambda}(1, 1) + O(k^2) \\ &= 4 \sum_{\substack{1 \leq a < b \leq k \\ (a,b)=1}} N_{k,\lambda}(a, b) + O(k^2), \end{aligned}$$

where the last step used the fact that $N_{k,\lambda}(1, 1) \leq \#\{(c, d) \in \mathbb{Z}^2 : |c|, |d| \leq k\} \ll k^2$.

Using the characteristic property of the Möbius function (4), we can write the last expression as

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4 \sum_{1 \leq a < b \leq k} N_{k,\lambda}(a, b) \sum_{d|(a,b)} \mu(d) + O(k^2) \\ &= 4 \sum_{d \leq k} \mu(d) \sum_{\substack{1 \leq a < b \leq k \\ d|a, d|b}} N_{k,\lambda}(a, b) + O(k^2) \\ &= 4 \sum_{d \leq k} \mu(d) \sum_{1 \leq \alpha < \beta \leq k/d} N_{k,\lambda}(d\alpha, d\beta) + O(k^2). \quad \square \end{aligned}$$

Lemma 9. *Let k and λ be integers with $0 \leq \lambda \leq 2k$, and let x and y be integers with $1 \leq x \leq y \leq k$. Then*

$$N_{k,\lambda}(x, y) = k^2 C\left(\frac{\lambda}{k}; x, y\right) D\left(\frac{\lambda}{k}; x, y\right) + O\left(\frac{k}{y}\right),$$

where

$$C(\delta; x, y) = \max\left\{0, \min\left\{\frac{1-\delta}{x} + \frac{1}{y}, \frac{2}{y}\right\}\right\}, \quad D(\delta; x, y) = \min\left\{\frac{1-\delta}{y} + \frac{1}{x}, \frac{2}{y}\right\}. \quad (7)$$

Proof. We have

$$\begin{aligned} N_{k,\lambda}(x, y) &= \#\{(c, d) \in \mathbb{Z}^2, c \neq 0, d \neq 0 : |xc + \lambda|, |yc|, |xd|, |yd + \lambda| \leq k\} \\ &= \#\{c \in \mathbb{Z}, c \neq 0 : -k \leq xc + \lambda \leq k, -k \leq yc \leq k\} \\ &\quad \times \#\{d \in \mathbb{Z}, d \neq 0 : -k \leq xd \leq k, -k \leq yd + \lambda \leq k\}. \end{aligned}$$

Since x and y are positive, we can rewrite this product as

$$\begin{aligned} N_{k,\lambda}(x, y) &= \#\left\{c \in \mathbb{Z}, c \neq 0 : \frac{-k-\lambda}{x} \leq c \leq \frac{k-\lambda}{x}, -\frac{k}{y} \leq c \leq \frac{k}{y}\right\} \\ &\quad \times \#\left\{d \in \mathbb{Z}, d \neq 0 : -\frac{k}{x} \leq d \leq \frac{k}{x}, \frac{-k-\lambda}{y} \leq d \leq \frac{k-\lambda}{y}\right\} \\ &= \#\left\{c \in \mathbb{Z}, c \neq 0 : -\frac{k}{y} \leq c \leq \min\left\{\frac{k-\lambda}{x}, \frac{k}{y}\right\}\right\} \\ &\quad \times \#\left\{d \in \mathbb{Z}, d \neq 0 : \max\left\{-\frac{k}{x}, \frac{-k-\lambda}{y}\right\} \leq d \leq \frac{k-\lambda}{y}\right\}, \quad (8) \end{aligned}$$

where we have used $\lambda \geq 0$ and $x \leq y$ to simplify the inequalities slightly. The first factor on the right-hand side of (8) is

$$\min\left\{\frac{k-\lambda}{x}, \frac{k}{y}\right\} - \left(-\frac{k}{y}\right) + O(1) = k \min\left\{\frac{1-\lambda/k}{x} + \frac{1}{y}, \frac{2}{y}\right\} + O(1)$$

if this expression is positive, and 0 otherwise; it is thus precisely $kC(\lambda/k; x, y) + O(1)$. Similarly, the second factor on the right-hand side of (8) is

$$\frac{k-\lambda}{y} - \max\left\{-\frac{k}{x}, \frac{-k-\lambda}{y}\right\} + O(1) = k \min\left\{\frac{1-\lambda/k}{y} + \frac{1}{x}, \frac{2}{y}\right\} + O(1)$$

(note that this expression is always positive under the hypotheses of the lemma), which is simply $kD(\lambda/k; x, y) + O(1)$. Multiplying these two factors yields

$$N_{k,\lambda}(x, y) = k^2 C\left(\frac{\lambda}{k}; x, y\right) D\left(\frac{\lambda}{k}; x, y\right) + k \cdot O\left(C\left(\frac{\lambda}{k}; x, y\right) + D\left(\frac{\lambda}{k}; x, y\right)\right) + O(1).$$

The lemma follows upon noting that both $C(\lambda/k; x, y)$ and $D(\lambda/k; x, y)$ are $\ll 1/y$ by definition, so the second summand becomes simply $O(k/y)$, and the $O(1)$ term may be subsumed into $O(k/y)$ since $y \leq k$. \square

We have already used the trivial estimate

$$\sum_{L \leq a < U} 1 = (U - L) + O(1),$$

provided $0 < L < U$. We will also use, without further comment, the estimates

$$\sum_{L \leq a < U} \frac{1}{a} = \log \frac{U}{L} + O\left(\frac{1}{L}\right),$$

with its particular case

$$\sum_{1 \leq a < U} \frac{1}{a} = \log U + O(1),$$

and

$$\sum_{L \leq a < U} \frac{1}{a^2} = \frac{1}{L} - \frac{1}{U} + O\left(\frac{1}{L^2}\right).$$

These estimates (also valid for $0 < L < U$) follow readily from comparison to the integrals $\int_L^U dx/x$ and $\int_L^U dx/x^2$.

Most of the technical work in proving Theorem 1 lies in establishing an estimate on a sum of the form $\sum_{1 \leq \alpha < \beta} C(\delta; \alpha, \beta) D(\delta; \alpha, \beta)$ for a fixed β . The following proposition provides an asymptotic formula for this sum; we defer the proof until the next section. Assuming this proposition, though, we can complete the proof of Theorem 1, as well as Corollary 2.

Proposition 10. *Let $\beta \geq 1$ and $0 \leq \delta \leq 2$ be real numbers, and let C and D be the functions defined in (7). Then*

$$\sum_{1 \leq \alpha < \beta} C(\delta; \alpha, \beta) D(\delta; \alpha, \beta) = \frac{V(\delta)}{\beta} + O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

where $V(\delta)$ was defined in (1).

Proof of Theorem 1 assuming Proposition 10. The functions C and D defined in (7) are homogeneous of degree -1 in the variables x and y , so that Lemma 9 implies

$$N_{k,\lambda}(d\alpha, d\beta) = \frac{k^2}{d^2} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O\left(\frac{k}{d\beta}\right).$$

Inserting this formula into the conclusion of Lemma 8 yields

$$|\mathcal{M}_2^\lambda(k)| = 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \alpha < \beta \leq k/d} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O\left(\sum_{d \leq k} \sum_{1 \leq \alpha < \beta \leq k/d} \frac{k}{d\beta}\right) + O(k^2).$$

We bound the first error term by summing over $1 \leq \alpha < \beta$ to obtain

$$\sum_{d \leq k} \sum_{1 \leq \alpha < \beta \leq k/d} \frac{k}{d\beta} \leq \sum_{d \leq k} \sum_{1 < \beta \leq k/d} \frac{k}{d} < \sum_{d \leq k} \frac{k^2}{d^2} \ll k^2,$$

so that we have the estimate

$$|\mathcal{M}_2^\lambda(k)| = 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \alpha < \beta \leq k/d} C\left(\frac{\lambda}{k}; \alpha, \beta\right) D\left(\frac{\lambda}{k}; \alpha, \beta\right) + O(k^2).$$

We now apply Proposition 10 to obtain

$$\begin{aligned} |\mathcal{M}_2^\lambda(k)| &= 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \sum_{1 \leq \beta \leq k/d} \left(\frac{V(\lambda/k)}{\beta} + O\left(\frac{1 + \log \beta}{\beta^2}\right)\right) + O(k^2) \\ &= 4k^2 \sum_{d \leq k} \frac{\mu(d)}{d^2} \left(V\left(\frac{\lambda}{k}\right) \left(\log \frac{k}{d} + O(1)\right) + O(1)\right) + O(k^2) \\ &= 4k^2 \left(V\left(\frac{\lambda}{k}\right) \log k \sum_{d \leq k} \frac{\mu(d)}{d^2} + O\left(\sum_{d \leq k} \frac{\log d}{d^2}\right)\right) + O(k^2) \\ &= 4k^2 \left(V\left(\frac{\lambda}{k}\right) \log k \left(\frac{6}{\pi^2} + O\left(\frac{1}{k}\right)\right) + O(1)\right) + O(k^2) \\ &= \frac{24}{\pi^2} V\left(\frac{\lambda}{k}\right) k^2 \log k + O(k^2), \end{aligned}$$

where we have used (5) and the convergence of $\sum 1/n^2$ and $\sum (\log n)/n^2$ (so the partial sums are uniformly bounded). \square

Proof of Corollary 2 from Theorem 1. Note that for any $M \in \mathcal{M}_2(k)$, if one eigenvalue is an integer then both of them are (since the trace of M is an integer). Thus if we add up the cardinalities of all of the $\mathcal{M}_2^\lambda(k)$, we get twice the cardinality of $\mathcal{M}_2^{\mathbb{Z}}(k)$, except that matrices with repeated eigenvalues only get counted once. However, the number of such matrices is $\ll_\varepsilon k^{2+\varepsilon}$ by Lemma 7. Therefore

$$\begin{aligned} 2|\mathcal{M}_2^{\mathbb{Z}}(k)| &= \sum_{\lambda \in \mathbb{Z}} |\mathcal{M}_2^\lambda(k)| + O_\varepsilon(k^{2+\varepsilon}) \\ &= \sum_{-2k \leq \lambda \leq 2k} \left(\frac{24}{\pi^2} V\left(\frac{\lambda}{k}\right) k^2 \log k + O(k^2) \right) + O_\varepsilon(k^{2+\varepsilon}) \\ &= \frac{24}{\pi^2} k^3 \log k \sum_{-2k \leq \lambda \leq 2k} \frac{V\left(\frac{\lambda}{k}\right)}{k} + O(k^3). \end{aligned}$$

The sum is a Riemann sum of a function of bounded variation, so this becomes

$$2|\mathcal{M}_2^{\mathbb{Z}}(k)| = \frac{24}{\pi^2} k^3 \log k \left(\int_{-2}^2 V(\delta) d\delta + O\left(\frac{1}{k}\right) \right) + O(k^3).$$

The corollary then follows from the straightforward computation of the integral

$$\int_{-2}^2 V(\delta) d\delta = \frac{4}{9}(7\sqrt{2} + 4 + 3 \log(\sqrt{2} + 1)),$$

noting that $\log(\sqrt{2} - 1) = -\log(\sqrt{2} + 1)$. □

4. Proving the key proposition

It remains to prove Proposition 10. Recalling that the functions C and D defined in (7) are formed by combinations of minima and maxima, we need to separate our arguments into several cases depending on the range of δ . The following lemma addresses a sum that occurs in two of these cases ($0 < \delta < 1$ and $1 < \delta < \sqrt{2}$). Note that because of the presence of terms like $\log(\delta - 1)$ in the formula for $V(\delta)$, we need to exercise some caution near $\delta = 1$.

Lemma 11. *Let $\beta \geq 1$ and $0 \leq \delta \leq \sqrt{2}$ be real numbers, with $\delta \neq 1$. Then*

$$\sum'_\alpha \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left(\frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left(\frac{1+\log \beta}{\beta^2}\right),$$

where \sum'_α is the sum subject to the condition $\max\{1, |1-\delta|\beta\} \leq \alpha < \beta/(1+\delta)$.

Proof. Suppose first that $|1-\delta|\beta \geq 1$. Then the sum in question is

$$\begin{aligned} & \frac{2(1-\delta)}{\beta} \sum_{\substack{\alpha \geq |1-\delta|\beta \\ \alpha < \beta/(1+\delta)}} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{\substack{\alpha \geq |1-\delta|\beta \\ \alpha < \beta/(1+\delta)}} 1 \\ &= \frac{2(1-\delta)}{\beta} \left(\log \frac{\beta}{(1+\delta)|1-\delta|\beta} + O\left(\frac{1}{|1-\delta|\beta}\right) \right) + \frac{2}{\beta^2} \left(\frac{\beta}{1+\delta} - |1-\delta|\beta + O(1) \right) \\ &= \frac{2}{\beta} \left(\frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left(\frac{1}{\beta^2}\right), \end{aligned}$$

which establishes the lemma in this case. On the other hand, if $|1-\delta|\beta < 1$ then the sum in question is

$$\begin{aligned} & \frac{2(1-\delta)}{\beta} \sum_{1 \leq \alpha < \beta/(1+\delta)} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{1 \leq \alpha < \beta/(1+\delta)} 1 \\ &= \frac{2(1-\delta)}{\beta} \left(\log \frac{\beta}{1+\delta} + O(1) \right) + \frac{2}{\beta^2} \left(\frac{\beta}{1+\delta} + O(1) \right) \\ &= \frac{2}{\beta} \left(\frac{1}{1+\delta} - (1-\delta) \log(1+\delta) \right) + O\left(\frac{1}{\beta^2} + \frac{|1-\delta| \log \beta}{\beta}\right). \end{aligned}$$

We subtract $2(|1-\delta| + (1-\delta) \log |1-\delta|)/\beta$ from the main term and compensate in the error term to obtain

$$\begin{aligned} & \frac{2(1-\delta)}{\beta} \sum_{1 \leq \alpha < \beta/(1+\delta)} \frac{1}{\alpha} + \frac{2}{\beta^2} \sum_{1 \leq \alpha < \beta/(1+\delta)} 1 \\ &= \frac{2}{\beta} \left(\frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) \\ & \quad + O\left(\frac{1}{\beta^2} + \frac{|1-\delta| \log \beta}{\beta} + \frac{|1-\delta| + |1-\delta| \log |1-\delta|^{-1}}{\beta}\right) \\ &= \frac{2}{\beta} \left(\frac{1}{1+\delta} - |1-\delta| - (1-\delta) \log |1-\delta^2| \right) + O\left(\frac{1 + \log \beta}{\beta^2} + \frac{|1-\delta| \log |1-\delta|^{-1}}{\beta}\right), \end{aligned}$$

since we are working with the assumption that $|1-\delta| < 1/\beta$. Because the function $t \log t^{-1}$ is increasing on the interval $(0, 1/e)$ and bounded on the interval $(0, 1]$, we have $|1-\delta| \log |1-\delta|^{-1} < (1/\beta) \log \beta$ if $\beta > e$ and $|1-\delta| \log |1-\delta|^{-1} \ll 1 \ll \frac{1}{\beta}$ if $1 \leq \beta \leq e$. In either case, the last error term can be simplified to

$$O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

which establishes the lemma in second case. □

Proof of Proposition 10. We consider separately the four cases corresponding to the different parts of the definition (1) of $V(\delta)$. To lighten the expressions, we use the \sum'_α notation from the statement of Lemma 11 and omit the dependence on δ , α , and β from the notation for C and D .

Case 1: $0 \leq \delta < 1$. In this case we have $0 < 1 - \delta < \frac{1}{1+\delta} \leq 1$ and

$$C = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq (1 - \delta)\beta, \\ \frac{1-\delta}{\alpha} + \frac{1}{\beta}, & \text{if } (1 - \delta)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{1+\delta}, \\ \frac{1-\delta}{\beta} + \frac{1}{\alpha}, & \text{if } \frac{\beta}{1+\delta} \leq \alpha. \end{cases}$$

Therefore,

$$\begin{aligned} & \sum_{1 \leq \alpha < \beta} CD \\ &= \sum_{1 \leq \alpha < (1-\delta)\beta} \frac{2}{\beta} \cdot \frac{2}{\beta} + \sum'_a \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} + \sum_{\beta/(1+\delta) \leq \alpha < \beta} \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \left(\frac{1-\delta}{\beta} + \frac{1}{\alpha} \right). \end{aligned} \tag{9}$$

(The first sum might be empty, but this does not invalidate the argument that follows.) The first sum is simply

$$\frac{4}{\beta^2} \sum_{1 \leq \alpha < (1-\delta)\beta} 1 = \frac{4}{\beta^2} ((1 - \delta)\beta + O(1)) = \frac{4(1-\delta)}{\beta} + O\left(\frac{1}{\beta^2}\right).$$

By Lemma 11, the second sum is

$$\sum'_a \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left(\frac{1}{1+\delta} + \delta - 1 - (1-\delta) \log(1-\delta^2) \right) + O\left(\frac{1+\log \beta}{\beta^2}\right),$$

while the third sum is

$$\begin{aligned} & \sum_{\beta/(1+\delta) \leq \alpha < \beta} \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta} \right) \left(\frac{1-\delta}{\beta} + \frac{1}{\alpha} \right) \\ &= (1 - \delta) \sum_{\beta/(1+\delta) \leq \alpha < \beta} \frac{1}{\alpha^2} + \frac{\delta^2 - 2\delta + 2}{\beta} \sum_{\beta/(1+\delta) \leq \alpha < \beta} \frac{1}{\alpha} + \frac{1-\delta}{\beta^2} \sum_{\beta/(1+\delta) \leq \alpha < \beta} 1 \\ &= (1 - \delta) \left(\frac{1+\delta}{\beta} - \frac{1}{\beta} + O\left(\frac{1}{\beta^2}\right) \right) \\ &\quad + \frac{\delta^2 - 2\delta + 2}{\beta} \left(\log \frac{\beta}{(1+\delta)^{-1}\beta} + O\left(\frac{1}{\beta}\right) \right) + \frac{1-\delta}{\beta^2} \left(\beta - \frac{\beta}{1+\delta} + O(1) \right) \\ &= \frac{1}{\beta} \left(2 - \delta^2 - \frac{2}{1+\delta} + (\delta^2 - 2\delta + 2) \log(1 + \delta) \right) + O\left(\frac{1}{\beta^2}\right). \end{aligned} \tag{10}$$

This case of the proposition then follows from (9) by noting that

$$\begin{aligned} & 4 - 4\delta + \frac{2}{1+\delta} + 2\delta - 2 - 2(1 - \delta) \log(1 - \delta^2) + 2 - \delta^2 - \frac{2}{1+\delta} + (\delta^2 - 2\delta + 2) \log(1 + \delta) \\ &= 4 - 2\delta - \delta^2 + \delta^2 \log(1 + \delta) - 2(1 - \delta) \log(1 - \delta). \end{aligned}$$

Case 2: $\delta = 1$. In this case we have

$$C = \frac{1}{\beta}, \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{2}, \\ \frac{1}{\alpha} & \text{if } \frac{\beta}{2} \leq \alpha. \end{cases}$$

Therefore,

$$\begin{aligned} \sum_{1 \leq \alpha < \beta} CD &= \sum_{1 \leq \alpha < \beta/2} \frac{1}{\beta} \cdot \frac{2}{\beta} + \sum_{\beta/2 \leq \alpha < \beta} \frac{1}{\beta} \cdot \frac{1}{\alpha} \\ &= \frac{2}{\beta^2} \left(\frac{\beta}{2} + O(1) \right) + \frac{1}{\beta} \left(\log \frac{\beta}{\beta/2} + O\left(\frac{1}{\beta}\right) \right) = \frac{1 + \log 2}{\beta} + O\left(\frac{1}{\beta^2}\right), \end{aligned}$$

as desired.

Case 3: $1 < \delta \leq \sqrt{2}$. In this case we have

$$C = \begin{cases} 0 & \text{if } \alpha \leq (\delta - 1)\beta, \\ \frac{1 - \delta}{\alpha} + \frac{1}{\beta} & \text{if } (\delta - 1)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{\delta + 1}, \\ \frac{1 - \delta}{\beta} + \frac{1}{\alpha} & \text{if } \frac{\beta}{\delta + 1} \leq \alpha. \end{cases}$$

Therefore,

$$\sum_{1 \leq \alpha < \beta} CD = \sum_{\alpha}' \left(\frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} + \sum_{\beta/(\delta + 1) \leq \alpha < \beta} \left(\frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \left(\frac{1 - \delta}{\beta} + \frac{1}{\alpha} \right). \quad (11)$$

(We note that $(\delta - 1)\beta \leq \beta/(\delta + 1)$ for δ between 1 and $\sqrt{2}$. For very small β we might have $1 > \beta/(\delta + 1)$, in which case the first sum is empty, but that does not invalidate the argument that follows.) By Lemma 11, the first sum is

$$\sum_{\alpha}' \left(\frac{1 - \delta}{\alpha} + \frac{1}{\beta} \right) \frac{2}{\beta} = \frac{2}{\beta} \left(\frac{1}{1 + \delta} + 1 - \delta - (1 - \delta) \log(\delta^2 - 1) \right) + O\left(\frac{1 + \log \beta}{\beta^2}\right),$$

while the second sum has already been evaluated in (10) above. This case of the proposition then follows from (11) by noting that

$$\begin{aligned} \frac{2}{1 + \delta} + 2 - 2\delta - 2(1 - \delta) \log(\delta^2 - 1) + 2 - \delta^2 - \frac{2}{\delta + 1} + (\delta^2 - 2\delta + 2) \log(\delta + 1) \\ = 4 - 2\delta - \delta^2 + \delta^2 \log(\delta + 1) + 2(\delta - 1) \log(\delta - 1). \end{aligned}$$

Case 4: $\sqrt{2} < \delta \leq 2$. Just as in Case 3, we have

$$C = \begin{cases} 0 & \text{if } \alpha \leq (\delta - 1)\beta, \\ \frac{1 - \delta}{\alpha} + \frac{1}{\beta} & \text{if } (\delta - 1)\beta \leq \alpha, \end{cases} \quad D = \begin{cases} \frac{2}{\beta} & \text{if } \alpha \leq \frac{\beta}{\delta + 1}, \\ \frac{1 - \delta}{\beta} + \frac{1}{\alpha} & \text{if } \frac{\beta}{\delta + 1} \leq \alpha. \end{cases}$$

However, the inequality $(\delta - 1)\beta \leq \alpha$ automatically implies that $\frac{\beta}{\delta+1} \leq \alpha$ when $\delta \geq \sqrt{2}$. Therefore,

$$\sum_{1 \leq \alpha < \beta} CD = \sum_{(\delta-1)\beta \leq \alpha < \beta} \left(\frac{1-\delta}{\alpha} + \frac{1}{\beta}\right) \left(\frac{1-\delta}{\beta} + \frac{1}{\alpha}\right).$$

(In this case we will not need to use $\max\{1, (\delta - 1)\beta\} \leq \alpha$, which is the more precise lower bound.) This yields

$$\begin{aligned} & \sum_{1 \leq \alpha < \beta} CD \\ &= (1 - \delta) \sum_{(\delta-1)\beta \leq \alpha < \beta} \frac{1}{\alpha^2} + \frac{\delta^2 - 2\delta + 2}{\beta} \sum_{(\delta-1)\beta \leq \alpha < \beta} \frac{1}{\alpha} + \frac{1 - \delta}{\beta^2} \sum_{(\delta-1)\beta \leq \alpha < \beta} 1 \\ &= (1 - \delta) \left(\frac{1}{(\delta-1)\beta} - \frac{1}{\beta} + O\left(\frac{1}{(\delta-1)^2\beta^2}\right) \right) \\ &\quad + \frac{\delta^2 - 2\delta + 2}{\beta} \left(\log \frac{\beta}{(\delta-1)\beta} + O\left(\frac{1}{(\delta-1)\beta}\right) \right) + \frac{1 - \delta}{\beta^2} (\beta - (\delta - 1)\beta + O(1)) \\ &= \frac{1}{\beta} \left(\delta^2 - 2\delta - (\delta^2 - 2\delta + 2) \log(\delta - 1) \right) + O\left(\frac{1}{\beta^2}\right), \end{aligned}$$

the error terms having been simplified since $\delta - 1$ is bounded away from 0. □

5. Distribution of real eigenvalues

In proving Theorem 3, it will be convenient to define the odd function

$$G(z) = \int_0^z -\log |t| dt = z(1 - \log |z|), \tag{12}$$

whose relevance is demonstrated by the following lemma.

Lemma 12. *If B and C are independent random variables uniformly distributed on $[-1, 1]$, the product BC is a random variable whose distribution function is*

$$F_{BC}(z) = \Pr(BC < z) = \frac{1}{2}(1 + G(z))$$

for $z \in [-1, 1]$.

Of course, for $z < -1$ we have $F_{BC}(z) = 0$, and likewise $F_{BC}(z) = 1$ for $z > 1$.

Proof. Note that $|B|$ and $|C|$ are uniformly distributed on $[0, 1]$. For $0 \leq z \leq 1$, we easily check that

$$\Pr(|BC| < z) = \int_0^1 \min\left\{1, \frac{z}{s}\right\} ds = G(z).$$

Thus $|BC|$ is distributed on $[0, 1]$ with density $f_{|BC|}(z) = -\log z$, and by symmetry BC has density $f_{BC}(z) = -\frac{1}{2} \log |z|$ on $[-1, 1]$. The lemma follows upon computing

$$F_{BC}(z) = \int_{-1}^z f_{BC}(s) ds. \quad \square$$

It will also be helpful to define the following functions, which are symmetric in x and y :

$$v_1(x, y) = \frac{1}{2} + \frac{1}{2}G(xy) + G\left(\frac{(x-y)^2}{4}\right), \quad (13)$$

$$v_2(x, y) = \frac{1}{2} - \frac{1}{2}G(xy), \quad (14)$$

$$v(x, y) = \begin{cases} v_1(x, y), & \text{if } xy < 1 \text{ and } x + y < 0, \\ v_2(x, y), & \text{if } xy < 1 \text{ and } x + y > 0, \\ 1 + G\left(\frac{(x-y)^2}{4}\right), & \text{if } xy > 1 \text{ and } x + y < 0, \\ 0, & \text{otherwise.} \end{cases} \quad (15)$$

To prove Theorem 3, we first consider the distribution function

$$F_W(\delta) = \int_{t < \delta} W(t) dt$$

associated to the density $W(\delta)$. For a random matrix M in $\mathcal{M}_2([-1, 1])$ and a real number δ , we will derive an expression for the expected number of real eigenvalues of M falling below δ , then differentiate it to obtain $W(\delta)$.

It is clear that $W(-\delta) = W(\delta)$ since the set $\mathcal{M}_2([-1, 1])$ is closed under negation, so it suffices to compute $W(\delta)$ for $\delta \in [0, 2]$. It turns out that our calculations for F_W will be somewhat simplified by considering $F_W(-\delta)$ rather than $F_W(\delta)$.

Proposition 13. *We have*

$$F_W(-\delta) = \frac{1}{4} \int_{-1+\delta}^{1+\delta} \int_{-1+\delta}^{1+\delta} v(x, y) dx dy$$

for all $0 \leq \delta \leq 2$, where v is defined in (15).

Proof. Denote the entries of M by A, B, C, D , which by assumption are independent random variables uniformly distributed in $[-1, 1]$. Let δ be fixed in the range $[0, 2]$, and consider the shifted matrix $M' = M + \delta I$, which we write as

$$M' = \begin{pmatrix} X & B \\ C & Y \end{pmatrix},$$

where X, Y range independently and uniformly in $[-1 + \delta, 1 + \delta]$ and B, C are as before. Clearly the eigenvalues of M less than $-\delta$ correspond to the negative (real) eigenvalues of M' . By Lemma 7, we are free to exclude the null set where

M' is singular or has repeated eigenvalues. Outside of this null set, M' has exactly one negative eigenvalue if and only if

$$\det M' = XY - BC < 0,$$

by Lemma 6 (c). Likewise by Lemma 6 (d), M' has exactly two negative eigenvalues if and only if

$$XY - BC > 0, \quad X + Y < 0 \quad \text{and} \quad \text{disc } M' = (X - Y)^2 + 4BC > 0.$$

We thus have:

$$F_W(-\delta) = \Pr(BC > XY) + 2\Pr\left(X + Y < 0 \text{ and } -\frac{(X - Y)^2}{4} < BC < XY\right).$$

We may express this probability as the average value

$$F_W(-\delta) = \frac{1}{4} \int_{-1+\delta}^{1+\delta} \int_{-1+\delta}^{1+\delta} \rho(x, y) \, dx \, dy,$$

where for fixed x and y ,

$$\begin{aligned} \rho(x, y) &= \Pr(BC > xy) + 2\Pr\left(x + y < 0 \text{ and } -\frac{(x - y)^2}{4} < BC < xy\right) \\ &= \Pr(BC > xy) + 2\Pr\left(-\frac{(x - y)^2}{4} < BC < xy\right) \mathbf{1}\{x + y < 0\} \end{aligned} \tag{16}$$

(here $\mathbf{1}\{\cdot\}$ denotes the indicator function of the indicated relation). To complete the proof it suffices to show that ρ equals the function ν defined in (15).

The probabilities appearing in (16) are effectively given by Lemma 12. However, there is some case-checking involved in applying this lemma, since the value of, say,

$$\Pr(BC > xy) = 1 - F_{BC}(xy)$$

depends on whether $xy < -1$, $-1 \leq xy \leq 1$, or $xy > 1$. We make some observations to reduce the number of cases we need to examine.

Note that $(x - y)^2/4$ is bounded between 0 and 1 for any $x, y \in [-1 + \delta, 1 + \delta]$, so that $-\frac{(x - y)^2}{4}$ always lies in the interval $[-1, 1]$ prescribed by Lemma 12. From the identity

$$(x + y)^2 - (x - y)^2 = 4xy$$

we see also that $xy \geq -(x - y)^2/4$. Thus xy is never lower than -1 , and we need only consider whether $xy > 1$ (in which case $F_{BC}(xy) = 1$). We therefore have

$$\begin{aligned} \Pr(BC > xy) &= 1 - F_{BC}(xy) = \frac{1}{2}(1 - G(xy)) \mathbf{1}\{xy < 1\}, \\ 2\Pr\left(-\frac{(x - y)^2}{4} < BC < xy\right) &= 2F_{BC}(xy) - 2F_{BC}\left(-\frac{(x - y)^2}{4}\right) \\ &= \mathbf{1}\{xy > 1\} + G(xy) \mathbf{1}\{xy < 1\} + G\left(\frac{(x - y)^2}{4}\right). \end{aligned}$$

Inserting these two evaluations into the formula (16), we obtain

$$\rho(x, y) = \frac{1}{2}(1 - G(xy))\mathbf{1}\{xy < 1\} + \left(\mathbf{1}\{xy > 1\} + G(xy)\mathbf{1}\{xy < 1\} + G\left(\frac{(x-y)^2}{4}\right)\right)\mathbf{1}\{x + y < 0\}.$$

It can be verified that this last expression is indeed equal to the right-hand side of the definition (15) of ν . \square

Since

$$W(\delta) = W(-\delta) = -\frac{d}{d\delta}F_W(-\delta),$$

to finish the proof of Theorem 3 it therefore suffices to prove that $-\frac{d}{d\delta}F_W(-\delta)$ equals the formula given in Theorem 3.

6. The derivative of the distribution

Proposition 13 expresses $F_W(\delta)$ as an integral, of a function ν that is independent of δ , over the square

$$S_\delta = [-1 + \delta, 1 + \delta]^2.$$

Since the region S_δ varies continuously with δ , we can compute the derivative

$$-\frac{d}{d\delta}F_W(-\delta)$$

by an appropriate line integral around the boundary of S_δ . Indeed, by the fundamental theorem of calculus, we have

$$\begin{aligned} -\frac{d}{d\delta}F_W(-\delta) &= -\frac{1}{4}\frac{d}{d\delta}\left(\int_{-1+\delta}^{1+\delta}\int_{-1+\delta}^{1+\delta}\nu(x, y) dx dy\right) \\ &= -\frac{1}{4}\left(\int_{-1+\delta}^{1+\delta}\nu(1 + \delta, y) dy - \int_{-1+\delta}^{1+\delta}\nu(-1 + \delta, y) dy \right. \\ &\quad \left. + \int_{-1+\delta}^{1+\delta}\nu(x, 1 + \delta) dx - \int_{-1+\delta}^{1+\delta}\nu(x, -1 + \delta) dx\right) \\ &= \frac{1}{2}\int_{-1+\delta}^{1+\delta}\nu(x, -1 + \delta) dx - \frac{1}{2}\int_{-1+\delta}^{1+\delta}\nu(x, 1 + \delta) dx, \end{aligned} \tag{17}$$

where we have used the symmetry $\nu(x, y) = \nu(y, x)$ to reduce the integral to just the top and bottom edges of S_δ (where $y = 1 + \delta$ and $y = -1 + \delta$, respectively).

The evaluation of (17) divides into three cases depending on the behavior of the indicator functions $\mathbf{1}\{x + y < 0\}$ and $\mathbf{1}\{xy < 1\}$ on the boundary of S_δ (Figure 2).

Case 1: $0 \leq \delta \leq 1$. For this range of δ , the line $x + y = 0$ intersects the bottom edge of S_δ at $x = 1 - \delta$, while the hyperbola $xy = 1$ intersects the top edge at

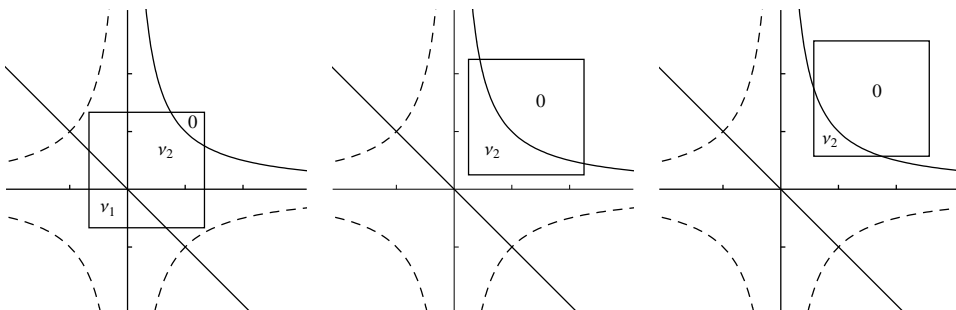


Figure 2. The three cases: $0 \leq \delta \leq 1$, $1 \leq \delta \leq \sqrt{2}$ and $\sqrt{2} \leq \delta \leq 2$.

$x = (1 + \delta)^{-1}$. Thus by the definition of v , (17) becomes

$$-\frac{d}{d\delta}F_W(-\delta) = \frac{1}{2} \left(\int_{-1+\delta}^{1-\delta} v_1(x, -1 + \delta) dx + \int_{1-\delta}^{1+\delta} v_2(x, -1 + \delta) dx - \int_{-1+\delta}^{(1+\delta)^{-1}} v_2(x, 1 + \delta) dx \right).$$

The following elementary antiderivatives, which are readily obtained by substitution and integration by parts, follow for any fixed nonzero real number y from the definitions (12), (13), and (14) of G , v_1 , and v_2 :

$$\begin{aligned} \int v_1(x, y) dx &= \frac{1}{2}x + \frac{1}{8}x^2y(3 - 2 \log |xy|) + \frac{1}{36}(x-y)^3 \left(5 - 6 \log \left| \frac{x-y}{2} \right| \right), \\ \int v_2(x, y) dx &= \frac{1}{2}x - \frac{1}{8}x^2y(3 - 2 \log |xy|). \end{aligned} \tag{18}$$

Therefore in this case,

$$\begin{aligned} -\frac{d}{d\delta}F_W(-\delta) &= \frac{1}{2} \left(\left(\frac{1}{2}x + \frac{1}{8}x^2(-1 + \delta)(3 - 2 \log |x(-1 + \delta)|) \right. \right. \\ &\quad \left. \left. + \frac{1}{36}(x + 1 - \delta)^3 \left(5 - 6 \log \left| \frac{x+1-\delta}{2} \right| \right) \right) \Big|_{x=-1+\delta}^{1-\delta} \right. \\ &\quad \left. + \left(\frac{1}{2}x - \frac{1}{8}x^2(-1 + \delta)(3 - 2 \log |x(-1 + \delta)|) \right) \Big|_{x=1-\delta}^{1+\delta} \right. \\ &\quad \left. - \left(\frac{1}{2}x - \frac{1}{8}x^2(1 + \delta)(3 - 2 \log |x(1 + \delta)|) \right) \Big|_{x=-1+\delta}^{(1+\delta)^{-1}} \right) \\ &= \frac{80 + 20\delta + 90\delta^2 + 52\delta^3 - 107\delta^4}{144(1 + \delta)} \\ &\quad - \frac{(5 - 7\delta + 8\delta^2)(1 - \delta)}{12} \log(1 - \delta) - \frac{\delta(1 - \delta^2)}{4} \log(1 + \delta) \end{aligned}$$

(after some algebraic simplification), which verifies the first case of Theorem 3. (Note that the integrands really are continuous, despite terms that look like $\log 0$, because the function G is continuous at 0; hence evaluating the integrals by antiderivatives is valid.)

Case 2: $1 \leq \delta \leq \sqrt{2}$. Now, the line $x + y = 0$ does not intersect S_δ , while the hyperbola $xy = 1$ intersects the top edge at $x = (1 + \delta)^{-1}$. Thus by the definition of ν and the antiderivative (18) of ν_2 , (17) becomes

$$\begin{aligned} -\frac{d}{d\delta}F_W(-\delta) &= \frac{1}{2} \left(\int_{-1+\delta}^{1+\delta} \nu_2(x, -1 + \delta) dx - \int_{-1+\delta}^{(1+\delta)^{-1}} \nu_2(x, 1 + \delta) dx \right) \\ &= \frac{1}{2} \left(\left(\frac{1}{2}x - \frac{1}{8}x^2(-1 + \delta)(3 - 2 \log |x(-1 + \delta)|) \right) \Big|_{x=-1+\delta}^{1+\delta} \right. \\ &\quad \left. - \left(\frac{1}{2}x - \frac{1}{8}x^2(1 + \delta)(3 - 2 \log |x(1 + \delta)|) \right) \Big|_{x=-1+\delta}^{(1+\delta)^{-1}} \right) \\ &= \frac{\delta(20 + 10\delta - 12\delta^2 - 3\delta^3)}{16(1 + \delta)} + \frac{(3\delta - 1)(\delta - 1)}{4} \log(\delta - 1) \\ &\quad + \frac{\delta(\delta^2 - 1)}{4} \log(\delta + 1), \end{aligned}$$

which verifies the second case of Theorem 3.

Case 3: $\sqrt{2} < \delta \leq 2$. As before, the line $x + y = 0$ does not intersect S_δ , while the hyperbola $xy = 1$ now intersects the bottom edge at $x = (\delta - 1)^{-1}$. Thus by the definition of ν and the antiderivative (18) of ν_2 , (17) becomes

$$\begin{aligned} -\frac{d}{d\delta}F_W(-\delta) &= \frac{1}{2} \int_{-1+\delta}^{(-1+\delta)^{-1}} \nu_2(x, -1 + \delta) dx \\ &= \frac{1}{2} \left(\frac{1}{2}x - \frac{1}{8}x^2(1 + \delta)(3 - 2 \log |x(1 + \delta)|) \right) \Big|_{x=-1+\delta}^{(-1+\delta)^{-1}} \\ &= \frac{\delta(\delta - 2)(2 - 6\delta + 3\delta^2)}{16(\delta - 1)} - \frac{(\delta - 1)^3}{4} \log(\delta - 1), \end{aligned}$$

which verifies the third case of Theorem 3.

Since the last case of Theorem 3 is a consequence of Lemma 4, the proof of the theorem is complete.

Remark. One could also use the same method to extract the individual distributions of the greater and lesser eigenvalues of M : for instance, eliminating the factor of 2 from (16) would yield an expression for the distribution of just the lesser eigenvalue of M .

References

- [Edelman 1997] A. Edelman, “The probability that a random real Gaussian matrix has k real eigenvalues, related distributions, and the circular law”, *J. Multivariate Anal.* **60**:2 (1997), 203–232. MR 98b:15025 Zbl 0886.15024
- [Gershgorin 1931] S. Gershgorin, “Über die Abgrenzung der Eigenwerte einer Matrix”, *Izv. Akad. Nauk. USSR Otd. Fiz.-Mat. Nauk* **7** (1931), 749–754.
- [Girko 1984] V. L. Girko, “The circular law”, *Teor. Veroyatnost. i Primenen.* **29**:4 (1984), 669–679. MR 87c:15042 Zbl 0565.60034
- [Hetzel et al. 2007] A. J. Hetzel, J. S. Liew, and K. E. Morrison, “The probability that a matrix of integers is diagonalizable”, *Amer. Math. Monthly* **114**:6 (2007), 491–499. MR 2008d:15054 Zbl 1140.15018
- [Katznelson 1993] Y. R. Katznelson, “Singular matrices and a uniform bound for congruence groups of $SL_n(\mathbf{Z})$ ”, *Duke Math. J.* **69**:1 (1993), 121–136. MR 94g:11083 Zbl 0785.11050
- [Martin and Wong 2008] G. Martin and E. B. Wong, “Almost all integer matrices have no integer eigenvalues”, preprint, 2008. To appear in *Amer. Math. Monthly*. arXiv 0712.3060
- [Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory. I. Classical theory*, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, Cambridge, 2007. MR 2378655 Zbl 1142.11001
- [Tao and Vu 2008] T. Tao and V. Vu, “Random matrices: Universality of ESDs and the circular law”, preprint, 2008. arXiv 0807.4898

Communicated by Andrew Granville

Received 2008-05-28

Revised 2008-08-13

Accepted 2008-10-16

gerg@math.ubc.ca

*Department of Mathematics, University of British Columbia,
Room 121, 1984 Mathematics Road,
Vancouver, BC V6T 1Z2, Canada
www.math.ubc.ca/~gerg*

erick@math.ubc.ca

*Department of Mathematics, University of British Columbia,
Room 121, 1984 Mathematics Road,
Vancouver, BC V6T 1Z2, Canada
www.math.ubc.ca/~erick*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in ANT are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use \LaTeX but submissions in other varieties of \TeX , and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of Bib \TeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@mathscipub.org with details about how your graphics were generated.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 2 No. 8 2008

Integral points on hyperelliptic curves	859
YANN BUGEAUD, MAURICE MIGNOTTE, SAMIR SIKSEK, MICHAEL STOLL and SZABOLCS TENGELY	
Smooth curves having a large automorphism p -group in characteristic $p > 0$	887
MICHEL MATIGNON and MAGALI ROCHER	
Inner derivations of alternative algebras over commutative rings	927
OTTMAR LOOS, HOLGER P. PETERSSON and MICHEL L. RACINE	
On Oliver's p -group conjecture	969
DAVID J. GREEN, LÁSZLÓ HÉTHELYI and MARKUS LILIENTHAL	
The number of 2×2 integer matrices having a prescribed integer eigenvalue	979
GREG MARTIN and ERICK WONG	



1937-0652(2008)2:8;1-8