Documenta Mathematica

Gegründet 1996 durch die

Deutsche Mathematiker-Vereinigung

Extra Volume

Proceedings

of the

Conference on Quadratic Forms

and

Related Topics

Baton Rouge, Louisiana, USA

March 26 – 30, 2001

Documenta Mathematica veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert.

Documenta Mathematica erscheint am World Wide Web unter:
`http://www.mathematik.uni-bielefeld.de/documenta`

Artikel können als TEX-Dateien per E-Mail bei einem der Herausgeber eingereicht werden. Hinweise für die Vorbereitung der Artikel können unter der obigen WWW-Adresse gefunden werden.

Documenta Mathematica publishes research manuscripts out of all mathematical fields and is refereed in the traditional manner.

Documenta Mathematica is published on the World Wide Web under:
`http://www.mathematik.uni-bielefeld.de/documenta`

Manuscripts should be submitted as TEX -files by e-mail to one of the editors. Hints for manuscript preparation can be found under the above WWW-address.

ISSN 1431-0635 (Print), ISSN 1431-0643 (Internet)

SPARC
Leading Edge

Documenta Mathematica is a Leading Edge Partner of SPARC, the Scholarly Publishing and Academic Resource Coalition of the Association of Research Libraries (ARL), Washington DC, USA.

Address of Technical Managing Editor: Ulf Rehmann, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld, Typesetting in TEX, Printing: media print services GmbH, D-83064 Raubling, Germany.

Documenta Mathematica
Extra Volume
Quadratic Forms and Related Topics, LSU, Baton Rouge, 2001

Preface

A conference on Quadratic forms and Related Topics was held at Louisiana State University, Baton Rouge, Louisiana, USA, from March 26 to March 30, 2001. This meeting was jointly supported by the National Science Foundation, the Louisiana Education Quality Support Fund, the LSU Office of Research and Graduate Studies, the LSU College of Arts and Sciences, and the LSU Department of Mathematics. The conference was organized by J. William Hoffman, Jurgen Hurrelbrink, Jorge Morales, Robert Perlis, and Paul van Wamelen, all at LSU.

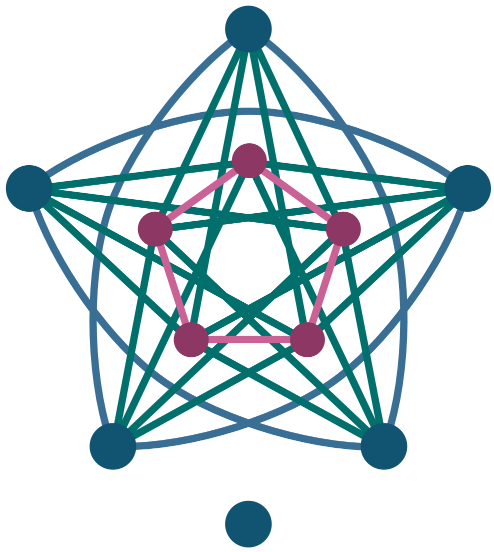
This book is the volume of the proceedings for that meeting. The majority of the articles published here record details of talks delivered at the conference. All contributions have been refereed independently according to Documenta Mathematica standards.

The papers in this volume are representative of the current state of the subject. In the recent past, the field of Quadratic Forms has enjoyed breakthrough results such as the confirmation of the Milnor Conjecture on relations between the theory of quadratic forms and algebraic K-theory. Topics of the articles in the proceedings include Witt groups, Brauer groups, Galois cohomology, generic splitting of quadratic forms, Hasse principles, and the theory of involutions.

It is a pleasure for us to give thanks to the agencies involved for their support of this conference. We would also like to take the opportunity to thank our colleagues, graduate students and staff at LSU for their untiring and alert assistance before and during the meeting, and all speakers and participants for their contributions to the success of the conference.

The Organizers
Baton Rouge, October 2001

The Logo



It is a classical problem to determine the structure of the ideal class group of a number field. Several students of the Algebraic Number Theory/Quadratic Forms group at LSU have been working on this problem.

The students associated graphs to quadratic fields so that the properties of the graphs yielded results about the ideal class groups. One of the graphs they came across is this beautiful graph. Except for the isolated vertex, the graph is the edge complement of the Petersen graph, one of the most fundamental and well-known of all graphs.

The LSU Department of Mathematics has adopted this graph as the logo for its website to symbolize the many years of achievement by its graduate students.

Editors

The proceedings are edited by the conference organizers J. W. Hoffman, J. Hurrelbrink, J. Morales, R. Perlis, P. van Wamelen in cooperation with the editors of Documenta Mathematica.

List of Talks

1. Jon Arason
   Witt Groups of Projective Line Bundles
2. Ricardo Baeza
   Behaviour of Bilinear Forms Under Function Field Extensions in Characteristic 2
3. Paul Balmer
   Survey of Quadratic Forms over Varieties
4. Pedro Benjamin Barquero
   On the Norm and Corestriction Principles for Reductive Algebraic Groups
5. Eva Bayer-Fluckiger
   Ideal Lattices
6. Karim Johannes Becher
   Non-Real Fields With Finite Square Class Number
7. Anthony Bevelacqua
   Global Isotropy of 4-Dimensional Quadratic Forms over $F(X)$
8. Eric Brussel
   The Brauer Group of a Strictly Henselian Field
9. Wai Kiu Chan
   On Almost Strong Approximation for Algebraic Groups
10. Vladimir Chernousov
    On the Rost Invariant for Quasi-Split Exceptional Groups
11. Andreas Dress
    Metric Spaces in Pure and Applied Mathematics
12. Martin Epkenhans
    On the Annihilating Ideal for Trace Forms
13. Robert W. Fitzgerald
    Isotropy and Factorization in Reduced Witt Rings
14. R. Skip Garibaldi
    Exterior Algebras and a Quadratic Form Invariant of Central Simple Algebras

15. Stefan Gille
    The Witt Group of the Relative Projective Line over a Regular Local Ring
16. Alexander J. Hahn
    Length Problems in Groups
17. Detlev Hoffmann
    Dimensions of Indefinite Quadratic Forms
18. Jens Hornbostel
    Localization in Hermitian K-Theory
19. John Hsia
    Representations of Quadratic Forms
20. Nikita Karpenko
    On the First Witt Index
21. Max A. Knus
    Quaternion Quadratic Forms, Involutions and Triality
22. Manfred Kolster
    Special Values of Zeta-Functions
23. Przemysław Koprowski
    Witt Equivalence of Function Fields of Real Algebraic Varieties
24. Ahmed Laghribi
    The Witt Kernel of a Multiquadratic Extension in Characteristic 2
25. David Leep
    Several Diverse Results in the Algebraic Theory
26. Alar Leibak
    On Venkov's Reduction of Positive Definite Unary Quadratic Forms
27. Louis Mahé
    Local-Global Principle for $\mathbf{R}(X, Y)$
28. Sean McGarraghy
    Exterior Powers for Quadratic Forms and Annihilating Polynomials
29. Alexander Merkurjev
    Unramified Cohomology of Classifying Varieties
30. Jan Minac
    Additive Properties of Multiplicative Subgroups of Fields
31. Jun Morita
    Bruhat-, Birkhoff-, Gauss-Decompositions and Their Variants

32. Ibrahim Mostafa
On Expanding $X^N + Y^N$ in Terms of Quadratic Binary Forms
33. William Pardon
The Filtered Gersten-Witt Resolution for Regular Schemes
34. R. Parimala
Quadratic Forms over 2-Dimensional Henselian Fields
35. Albrecht Pfister
Small Zeros of Quadratic Forms over Algebraic Function Fields
36. Anne Quéguiner-Mathieu
Decomposability of Degree 8 Algebras With Orthogonal Involution
37. Ulf Rehmann
Anisotropic Splitting Towers of Orthogonal Groups
38. Konstantin Rybnikov
Voronoi's Theories of Perfect Domains and L-Types for Positive Quadratic Forms
39. Claus Scheiderer
Sums of Squares in 2-Dimensional Local Rings
40. Tara L. Smith
Galois Groups over Nonrigid Fields
41. T. A. Springer
Large Schubert Varieties
42. Marek Szyjewski
Generalized Discriminant
43. Jean-Pierre Tignol
Multipliers of Similitudes
44. Tuong Ton-That
A Generalized Poincaré Theorem for Dual Group Actions
45. Thomas Unger
A Weak Hasse Principle for Algebras With Involution and Sums of Hermitian Squares
46. Jerzy Urbanowicz
Remarks on Linear Congruence Relations for Kubota-Leopoldt 2-adic L-Functions
47. Christiaan Van de Woestijne
Generalizing the Gram-Schmidt Orthogonalization Algorithm
48. Adrian Wadsworth
The Semihereditary Order of an Involution

List of Participants

1. Jon Arason (University of Iceland, Reykjavik, Iceland)
2. Ricardo Baeza (University of Talca, Chile)
3. Paul Balmer (University of Münster, Germany)
4. Pedro Barquero (Santa Monica College, USA)
5. Eva Bayer-Fluckiger (EPF Lausanne, Switzerland)
6. Karim Johannes Becher (University of Besançon, France)
7. Anthony Bevelacqua (University of North Dakota, USA)
8. Eric Brussel (Emory University, USA)
9. Juliusz Brzezinski (University of Gothenburg, Sweden)
10. Wai Kiu Chan (Wesleyan University, USA)
11. Vladimir Chernousov (University of Bielefeld, Germany)
12. Anne Cortella (University of Besançon, France)
13. Charles Delzell (Louisiana State University, USA)
14. Andreas Dress (University of Bielefeld, Germany)
15. Martin Epkenhans (University of Paderborn, Germany)
16. Laura Fainsilber (University of Gothenburg, Sweden)
17. Robert Fitzgerald (Southern Illinois University, USA)
18. Skip Garibaldi (University of California, Los Angeles, USA)
19. Stefan Gille (University of Münster, Germany)
20. Alexander Hahn (University of Notre Dame)
21. Sidney Hawkins (Alcorn State University, USA)
22. Detlev Hoffmann (University of Besançon, France)
23. Jens Hornbostel (University of Münster, Germany)
24. John Hsia (Ohio State University, USA)
25. Seva Joukhovitski (Northwestern University, USA)
26. Changheon Kang (Louisiana State University, USA)
27. Nikita Karpenko (University d'Artois, Lens, France)
28. Myung-Hwan Kim (Seoul National University, South Korea)
29. Max Knus (ETH Zürich, Switzerland)
30. Manfred Kolster (McMaster University, Canada)
31. Przemyslaw Koprowski (Silesian University, Poland)
32. Ahmed Laghribi (University d'Artois, Lens, France)
33. Douglas Larmour (Colorado College, USA)
34. David Leep (University of Kentucky, USA)
35. Alar Leibak (Tallinn Technical University, Estonia)

36. David Lewis (University College Dublin, Ireland)
37. Louis Mahé (University of Rennes, France)
38. Sean McGarraghy (University College Dublin, Ireland)
39. Alexander Merkurjev (University of California, Los Angeles, USA)
40. Jan Minac (University of Western Ontario, Canada)
41. Jun Morita (University of Tsukuba, Japan)
42. Ibrahim Mostafa (October 6 University, Giza, Egypt)
43. Brian Murray (Louisiana State University, USA)
44. Robert Osburn (Louisiana State University, USA)
45. William Pardon (Duke University, USA)
46. Raman Parimala (Tata Institute, India)
47. Albrecht Pfister (University of Mainz, Germany)
48. Alexandre Prestel (University of Konstanz, Germany)
49. Hourong Qin (Nanjing University, PRC)
50. Anne Quéguiner-Mathieu (University of Paris, France)
51. Ulf Rehmann (University of Bielefeld, Germany)
52. Konstantin Rybnikov (Cornell University, USA)
53. Claus Scheiderer (University of Duisburg, Germany)
54. Daniel Shapiro (Ohio State University, USA)
55. Tara Smith (University of Cincinnati, USA)
56. Marius Somodi (Louisiana State University, USA)
57. Tonny A. Springer (University of Utrecht, The Netherlands)
58. Marek Szyjewski (Silesian University, Poland)
59. Jean-Pierre Tignol (Catholic University of Louvain, Belgium)
60. Tuong Ton-That (University of Iowa, USA)
61. Thomas Unger (University College Dublin, Ireland)
62. Jerzy Urbanowicz (Polish Academy of Sciences, Poland)
63. Christiaan Van de Woestijne (University of Leiden, The Netherlands)
64. Jan Van Geel (University of Ghent, Belgium)
65. Sergei V. Vostokov (University of St. Petersburg, Russia)
66. Adrian Wadsworth (UC San Diego, USA)
67. Uroyoan Walker (Louisiana State University, USA)

Witt Groups of Projective Line Bundles

Jón Kr. Arason

Received: May 16, 2001

Communicated by Ulf Rehmann

Abstract. We construct an exact sequence for the Witt group of a projective line bundle.

2000 Mathematics Subject Classification: 11E81, 19G12
Keywords and Phrases: Witt groups. Projective line bundles.

Let X be a noetherian scheme over which 2 is invertible, let $\mathcal{S}$ be a vector bundle of rank 2 over X, and let $Y = \mathbf{P}(\mathcal{S})$ be the corresponding projective line bundle in the sense of Grothendieck. The structure morphism $f : Y \to X$ induces a morphism $f^* : W(X) \to W(Y)$ of Witt rings. In this paper we shall show that there is an exact sequence

$$W(X) \to W(Y) \to M_\top(X)$$

where $M_\top(X)$ is a Witt group of formations over X like the one defined by Ranicki in the affine case. (Cf. [R].) The subscript is meant to show that in the definition of $M_\top(X)$ we use a duality functor that might differ from the usual one.
In his work, Ranicki shows that in the affine case the Witt group $M(X)$ of formations is naturally isomorphic to his L-group $L^1(X)$. We therefore could have used the notation $L^1_\top(X)$. Furthermore, according to Walter, $M(X)$ is also the higher Witt group $W^{-1}(X)$ as defined by Balmer using derived categories. (Cf. [B].) And Walter, [W], has announced very interesting results on higher Witt groups of general projective space bundles over X of which our result is just a special case.

The paper has two main parts. In the first one we study the obstruction for an element in $W(Y)$ to come from $W(X)$. In the second part we define and study $M_\top(X)$. In a short third part we prove our main theorem and make some remarks.

Besides the notation already introduced, we shall use the following. We denote by $\mathcal{O}_Y(1)$ the tautological line bundle on Y. We shall, of course, use the usual

notation for twistings by $\mathcal{O}_Y(1)$. We denote by ω the relative canonical bundle $\omega_{Y/X}$. We also write $\mathcal{L} = \mathcal{S} \wedge \mathcal{S}$. Then $\omega = f^*(\mathcal{L})(-2)$. We shall write $\mathcal{S}_k$ for $f_*(\mathcal{O}_Y(k))$. In particular, $\mathcal{S}_0 = \mathcal{O}_X$ and $\mathcal{S}_1 = \mathcal{S}$.
There is a natural short exact sequence

$$0 \to \omega \to f^*(\mathcal{S})(-1) \to \mathcal{O}_Y \to 0$$

that we shall use often. As other results that we need on algebraic geometry, it can be found in [H].

Section 1.1

In this section we shall use higher direct images to check whether a symmetric bilinear space over Y comes from X.
The main fact used is the corresponding result in the linear case. It must be well known although we don't have a reference handy. We shall, however, give an elementary proof here.

Proposition 1: Let $\mathcal{E}$ be a coherent Y-module. If $R^1 f_*(\mathcal{E}(-1)) = 0$ and $f_*(\mathcal{E}(-1)) = 0$ then the canonical morphism $f^*(f_*(\mathcal{E})) \to \mathcal{E}$ is an isomorphism.
Proof: Let $\mathcal{E}$ be a coherent Y-module. We look at the tensor product

$$0 \to \omega \otimes \mathcal{E} \to f^*(\mathcal{S})(-1) \otimes \mathcal{E} \to \mathcal{E} \to 0$$

of $\mathcal{E}$ and the natural short exact sequence above. Twisting by $k+1$ and taking higher direct images we get the exact sequence

$$\begin{aligned} 0 &\to f_*(\omega \otimes \mathcal{E}(k+1)) \to \mathcal{S} \otimes f_*(\mathcal{E}(k)) \to f_*(\mathcal{E}(k+1)) \\ &\to R^1 f_*(\omega \otimes \mathcal{E}(k+1)) \to \mathcal{S} \otimes R^1 f_*(\mathcal{E}(k)) \to R^1 f_*(\mathcal{E}(k+1)) \to 0 \end{aligned}$$

From it we first get:
Fact 1: If $R^1 f_*(\mathcal{E}(k)) = 0$ then also $R^1 f_*(\mathcal{E}(k+1)) = 0$.
Noting that $\omega \otimes \mathcal{E}(k+1) = f^*(\mathcal{L}) \otimes \mathcal{E}(k-1)$ we also get:
Fact 2: If $R^1 f_*(\mathcal{E}(k-1)) = 0$ then the natural morphism $\mathcal{S} \otimes f_*(\mathcal{E}(k)) \to f_*(\mathcal{E}(k+1))$ is an epimorphism.
Using the two previous facts and induction on k we see that if $R^1 f_*(\mathcal{E}(-1)) = 0$ then $\mathcal{S}_k \otimes f_*(\mathcal{E}) \to f_*(\mathcal{E}(k))$ is an epimorphism for every $k \geq 0$. This implies:
Fact 3: If $R^1 f_*(\mathcal{E}(-1)) = 0$ then the canonical morphism $f^*(f_*(\mathcal{E})) \to \mathcal{E}$ is an epimorphism.
In the situation of Fact 3 we have a natural short exact sequence $0 \to \mathcal{N} \to f^*(f_*(\mathcal{E})) \to \mathcal{E} \to 0$. We note that $\mathcal{N}$ is coherent because $f_*(\mathcal{E})$ is coherent. Taking higher direct images, using that $f_*(f^*(f_*(\mathcal{E}))) \to f_*(\mathcal{E})$ is an isomorphism and that $R^1 f_*(f^*(f_*(\mathcal{E}))) = 0$, we get that $f_*(\mathcal{N}) = 0$ and $R^1 f_*(\mathcal{N}) = 0$.
Twisting the short exact sequence by -1 and then taking higher direct images,

using that $f_*(f^*(f_*(\mathcal{E}))(-1)) = 0$ and $R^1 f_*(f^*(f_*(\mathcal{E}))(-1)) = 0$, we get that $R^1 f_*(\mathcal{N}(-1))$ is naturally isomorphic to $f_*(\mathcal{E}(-1))$. So if $f_*(\mathcal{E}(-1)) = 0$ then $R^1 f_*(\mathcal{N}(-1)) = 0$. As $f_*(\mathcal{N}) = 0$ it then follows from Fact 3 that $\mathcal{N} = 0$. The proposition follows.
PROPOSITION 1, CNTD: Furthermore, if $\mathcal{E}$ is a vector bundle on Y then $f_*(\mathcal{E})$ is a vector bundle on X.
Proof: Clearly, Y is flat over X, so $\mathcal{E}$ is flat over X. Using the Theorem of Cohomology and Base Change, (cf. [H], Theorem III.12.11), we therefore see that if $\mathcal{E}$ is a vector bundle on Y such that $R^1 f_*(\mathcal{E}) = 0$ then $f_*(\mathcal{E})$ is a vector bundle on X.

Although we really do not need it here we bring the following generalization of Proposition 1.

PROPOSITION 2: Let $\mathcal{E}$ be a coherent Y-module. If $R^1 f_*(\mathcal{E}(-1)) = 0$ then there is a natural short exact sequence

$$0 \to f^*(f_*(\omega(1) \otimes \mathcal{E}))(-1) \to f^*(f_*(\mathcal{E})) \to \mathcal{E} \to 0$$

Proof: In the proof of Proposition 1 we had, even without the hypothesis $f_*(\mathcal{E}(-1)) = 0$, that $f_*(\mathcal{N}) = 0$ and $R^1 f_*(\mathcal{N}) = 0$. By Proposition 1 the canonical morphism $f^*(f_*(\mathcal{N}(1)) \to \mathcal{N}(1)$ is therefore an isomorphism. Taking the tensor product of this isomorphism with ω and using $R^1 f_*$ on the resulting isomorphism, noting that $R^1 f_*(\omega \otimes f^*(f_*(\mathcal{N}(1))))$ is naturally isomorphic to $f_*(\mathcal{N}(1))$, we see that $f_*(\mathcal{N}(1))$ is naturally isomorphic to $R^1 f_*(\omega \otimes \mathcal{N}(1)) = \mathcal{L} \otimes R^1 f_*(\mathcal{N}(-1))$. But we saw in the proof of Proposition 1 that $R^1 f_*(\mathcal{N}(-1))$ is naturally isomorphic to $f_*(\mathcal{E}(-1))$, so this means that $f_*(\mathcal{N}(1))$ is naturally isomorphic to $\mathcal{L} \otimes f_*(\mathcal{E}(-1))$. But $\mathcal{L} \otimes f_*(\mathcal{E}(-1)) = f_*(f^*(\mathcal{L}) \otimes \mathcal{E}(-1)) = f_*(\omega(1) \otimes \mathcal{E})$. The proposition follows.
PROPOSITION 2, CNTD: Furthermore, if $\mathcal{E}$ is a vector bundle on Y then $f_*(\mathcal{E})$ and $f_*(\omega(1) \otimes \mathcal{E})$ are vector bundles on X.
Proof: Noting that $f_*(\omega(1) \otimes \mathcal{E}) = \mathcal{L} \otimes f_*(\mathcal{E}(-1))$, this follows as in the proof of Proposition 1.

We shall, however, use the following corollary of Proposition 1.

PROPOSITION 3: Let $\mathcal{E}$ be a coherent Y-module. If $R^1 f_*(\mathcal{E}) = 0$ and $f_*(\mathcal{E}(-1)) = 0$ then there is a natural short exact sequence

$$0 \to f^*(f_*(\mathcal{E})) \to \mathcal{E} \to f^*(R^1 f_*(\omega(1) \otimes \mathcal{E}))(-1) \to 0$$

Proof: We let $\mathcal{C} = f_*(\mathcal{E})$. From the canonical morphism $f^*(f_*(\mathcal{E})) \to \mathcal{E}$ we then get an exact sequence

$$0 \to \mathcal{N} \to f^*(\mathcal{C}) \to \mathcal{E} \to \mathcal{Q} \to 0$$

of coherent Y-modules. As the direct image functor is left-exact and the induced morphism $f_*(f^*(\mathcal{C})) \to f_*(\mathcal{E})$ is an isomorphism, we see that $f_*(\mathcal{N}) = 0$. We now break the exact sequence up into two short exact sequences

$$0 \to \mathcal{N} \to f^*(\mathcal{C}) \to \mathcal{M} \to 0$$

and

$$0 \to \mathcal{M} \to \mathcal{E} \to \mathcal{Q} \to 0$$

Using the hypothesis $f_*(\mathcal{E}(-1)) = 0$, we get from the second short exact sequence that $f_*(\mathcal{M}(-1)) = 0$. Using that and the fact that $R^1 f_*(f^*(\mathcal{C})(-1)) = 0$, we get from the first one that $R^1 f_*(\mathcal{N}(-1)) = 0$. As we already saw that $f_*(\mathcal{N}) = 0$, it follows from Proposition 2 that $\mathcal{N} = 0$. (Fact 3 in the proof of Proposition 1 suffices.) So we have the short exact sequence

$$0 \to f^*(\mathcal{C}) \to \mathcal{E} \to \mathcal{Q} \to 0$$

As $R^1 f_*(\mathcal{E}) = 0$ and $f_*(f^*(\mathcal{C})) \to f_*(\mathcal{E})$ is an isomorphism, we get, using that $R^1 f_*(f^*(\mathcal{C})) = 0$, that $f_*(\mathcal{Q}) = 0$ and $R^1 f_*(\mathcal{Q}) = 0$. By Proposition 1 this means that the canonical morphism $f^*(f_*(\mathcal{Q}(1))) \to \mathcal{Q}(1)$ is an isomorphism. Writing $\mathcal{B} = f_*(\mathcal{Q}(1))$, we therefore get that $\mathcal{Q} \cong f^*(\mathcal{B})(-1)$.
Taking the tensor product of the short exact sequence

$$0 \to f^*(\mathcal{C}) \to \mathcal{E} \to f^*(\mathcal{B})(-1) \to 0$$

with $\omega(1)$ and then taking higher direct images, noting that $R^1 f_*(\omega(1) \otimes f^*(\mathcal{C})) = 0$, we get that $R^1 f_*(\omega(1) \otimes \mathcal{E}) \to R^1 f_*(\omega(1) \otimes f^*(\mathcal{B})(-1))$ is an isomorphism. But $R^1 f_*(\omega(1) \otimes f^*(\mathcal{B})(-1)) = R^1 f_*(\omega \otimes f^*(\mathcal{B}))$, which is canonically isomorphic to $\mathcal{B}$. So we have a natural isomorphism $\mathcal{B} \cong R^1 f_*(\omega(1) \otimes \mathcal{E})$.
Note: If $\mathcal{E}$ is a vector bundle on Y then we see as before that $f_*(\mathcal{E})$ is a vector bundle on X. But we don't know whether $R^1 f_*(\omega(1) \otimes \mathcal{E})$ is also a vector bundle on X.

There is, in fact, a natural exact sequence

$$\begin{aligned} 0 &\to f^*(f_*(\omega(1) \otimes \mathcal{E}))(-1) \to f^*(f_*(\mathcal{E})) \to \mathcal{E} \\ &\to f^*(R^1 f_*(\omega(1) \otimes \mathcal{E}))(-1) \to f^*(R^1 f_*(\mathcal{E})) \to 0 \end{aligned}$$

for any coherent Y-module $\mathcal{E}$. But we do not need that here. What we need is the following bilinear version of Proposition 1.

Proposition 4: Let $(\mathcal{E}, \chi)$ be a symmetric bilinear space over Y. If $R^1 f_*(\mathcal{E}(-1)) = 0$ then there is a symmetric bilinear space $(\mathcal{G}, \psi)$ over X such that $(\mathcal{E}, \chi) \cong f^*(\mathcal{G}, \psi)$.
Proof: For any morphism $f : Y \to X$ of schemes and any Y-module $\mathcal{F}$ and any X-module $\mathcal{G}$ there is a canonical isomorphism $f_*(\mathcal{H}om_Y(f^*(\mathcal{G}), \mathcal{F})) \cong$

$\mathcal{H}om_X(\mathcal{G}, f_*(\mathcal{F}))$. In our case $f_*(\mathcal{O}_Y) = \mathcal{O}_X$ hence, in particular, there is a canonical isomorphism $f_*(f^*(\mathcal{G})^\vee) \cong \mathcal{G}^\vee$. It then follows that there are canonical isomorphisms $\mathrm{Hom}_Y(f^*(\mathcal{G}), f^*(\mathcal{G})^\vee) \cong \mathrm{Hom}_X(\mathcal{G}, f_*(f^*(\mathcal{G})^\vee)) \cong \mathrm{Hom}_X(\mathcal{G}, \mathcal{G}^\vee)$.
In the case at hand we first note that as $\mathcal{E}$ is self dual, Serre duality shows that $R^1 f_*(\mathcal{E}(-1)) = 0$ implies that $f_*(\mathcal{E}(-1)) = 0$. So we can use Proposition 1 to write $\mathcal{E} \cong f^*(\mathcal{G})$ with the vector bundle $\mathcal{G} = f_*(\mathcal{E})$ over X. The proposition follows.

Section 1.2

In this section we shall prove a useful condition for the Witt class of a symmetric bilinear space over Y to come from $W(X)$.

Let $(\mathcal{E}, \chi)$ be a symmetric bilinear space over Y and let $\mathcal{U}$ be a totally isotropic subbundle of $(\mathcal{E}, \chi)$. Denote by $\mathcal{V}$ the orthogonal subbundle to $\mathcal{U}$ in $(\mathcal{E}, \chi)$ and by $\mathcal{F}$ the quotient bundle of $\mathcal{V}$ by $\mathcal{U}$. Then χ induces a symmetric bilinear form φ on $\mathcal{F}$ and the symmetric bilinear space $(\mathcal{F}, \varphi)$ has the same class in $W(Y)$ as $(\mathcal{E}, \chi)$. We also have the commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{U} & \to & \mathcal{V} & \to & \mathcal{F} & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{U} & \to & \mathcal{E} & \to & \mathcal{V}^\vee & \to & 0 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & \mathcal{U}^\vee & = & \mathcal{U}^\vee & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

with exact rows and columns. It is self-dual up to the isomorphisms χ and φ. It is natural to say that $(\mathcal{F}, \varphi)$ is a quotient of $(\mathcal{E}, \chi)$ by the totally isotropic subbundle $\mathcal{U}$. But then one can also say that $(\mathcal{E}, \chi)$ is an extension of $(\mathcal{F}, \varphi)$ by $\mathcal{U}$. Extensions of symmetric bilinear spaces in this sense are studied in [A]. One of the main results there is that the set of equivalence classes of extensions of $(\mathcal{F}, \varphi)$ by $\mathcal{U}$ is functorial in $\mathcal{U}$.

Proposition 1: Let $\mathcal{M}$ be a metabolic space over Y. Then there is a metabolic space $\mathcal{N}$ over X such that $\mathcal{M}$ is a quotient of $f^*(\mathcal{N})$.
Proof: $\mathcal{M}$ is clearly a quotient of $\mathcal{M} \oplus -\mathcal{M}$. As 2 is invertible over Y, this latter space is hyperbolic. Hence it suffices to prove the assertion for hyperbolic spaces $H(\mathcal{U})$ over Y.
By Serre's Theorem (cf. [H], Theorem III.8.8) and the Theorem of Cohomology and Base Change ([H], Theorem III.12.11), we have for every sufficiently large N that $f_*(\mathcal{U}(N))$ is locally free and that the canonical morphism $f^*(f_*(\mathcal{U}(N))) \to \mathcal{U}(N)$ is an epimorphism. This means that there

is a vector bundle $\mathcal{A} = f_*(\mathcal{U}(N))$ over X such that $\mathcal{U}$ is a quotient of $f^*(\mathcal{A})(-N)$. But then, clearly, $H(\mathcal{U})$ is, as a symmetric bilinear space, a quotient of $H(f^*(\mathcal{A})(-N))$. Now, if N is sufficiently large, $f_*(\mathcal{O}_Y(N))$ is locally free and $f^*(f_*(\mathcal{O}_Y(N))) \to \mathcal{O}_Y(N)$ is an epimorphism, hence $f^*(\mathcal{A}^\vee)(N) = \mathcal{O}_Y(N) \otimes_{\mathcal{O}_Y} f^*(\mathcal{A}^\vee)$ is a quotient of $f^*(\mathcal{B})$ for the vector bundle $\mathcal{B} = f_*(\mathcal{O}_Y(N)) \otimes_{\mathcal{O}_X} \mathcal{A}^\vee$ over X. It follows that $H(f^*(\mathcal{A})(-N)) = H((f^*(\mathcal{A})(-N))^\vee) = H(f^*(\mathcal{A}^\vee)(N))$ is, as a symmetric bilinear space, a quotient of $H(f^*(\mathcal{B})) = f^*(H(\mathcal{B}))$. But then also $H(\mathcal{U})$ is a quotient of $f^*(H(\mathcal{B}))$.

Corollary: Let $\mathcal{F}$ be a symmetric bilinear space over Y such that the class of $\mathcal{F}$ in $W(Y)$ lies in the image of $f^* : W(X) \to W(Y)$. Then there is a symmetric bilinear space $\mathcal{G}$ over X such that $\mathcal{F}$ is a quotient of $f^*(\mathcal{G})$.
Proof: Write $\mathcal{F} \oplus \mathcal{M}_1 \cong f^*(\mathcal{G}_0) \oplus \mathcal{M}_2$ with a symmetric bilinear space $\mathcal{G}_0$ over X and metabolic spaces $\mathcal{M}_1$ and $\mathcal{M}_2$ over Y. Using the proposition on $\mathcal{M}_2$, we get that $\mathcal{F} \oplus \mathcal{M}_1$ is a quotient of $f^*(\mathcal{G})$, where $\mathcal{G} = \mathcal{G}_0 \oplus \mathcal{N}_2$ for some metabolic space $\mathcal{N}_2$ over X. Then also $\mathcal{F}$ is a quotient of $f^*(\mathcal{G})$.

In fact, the same proofs show that Proposition 1 and its Corollary hold for every projective scheme Y over X which is flat over X. But in the case at hand we can make the Corollary more specific:

Theorem 2: Let $\mathcal{F}$ be a symmetric bilinear space over Y such that the class of $\mathcal{F}$ in $W(Y)$ lies in the image of $f^* : W(X) \to W(Y)$. Then there is a symmetric bilinear space $\mathcal{G}$ over X and a vector bundle $\mathcal{Z}$ over X such that $\mathcal{F}$ is a quotient of $f^*(\mathcal{G})$ by $f^*(\mathcal{Z})(-1)$.
Proof: By the Corollary to Proposition 1, there is a symmetric bilinear space $\mathcal{G}$ over X such that $\mathcal{F}$ is a quotient of $f^*(\mathcal{G})$. Let the diagram at the beginning of this section be a presentation of $\mathcal{E} := f^*(\mathcal{G})$ as an extension of $\mathcal{F}$. As $\mathcal{E}$ comes from X, we have $R^1f_*(\mathcal{E}(-1)) = 0$. It follows that also $R^1f_*(\mathcal{V}^\vee(-1)) = 0$ and $R^1f_*(\mathcal{U}^\vee(-1)) = 0$. From the latter fact it follows that $f_*(\mathcal{U}^\vee(-1))$ is a vector bundle over X. We let $\mathcal{Z}$ be the dual bundle, so that $\mathcal{Z}^\vee = f_*(\mathcal{U}^\vee(-1))$. From the canonical morphism $f^*(f_*(\mathcal{U}^\vee(-1))) \to \mathcal{U}^\vee(-1)$ we get a morphism $f^*(\mathcal{Z}^\vee)(1) \to \mathcal{U}^\vee$. We let $\alpha : \mathcal{U} \to \mathcal{U}_1 := f^*(\mathcal{Z})(-1)$ be the dual morphism. By [A], there is an extension $\mathcal{E}_1$ of $\mathcal{F}$ by $\mathcal{U}_1$ with a corresponding presentation

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{U}_1 & \to & \mathcal{V}_1 & \to & \mathcal{F} & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{U}_1 & \to & \mathcal{E}_1 & \to & \mathcal{V}_1^\vee & \to & 0 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & \mathcal{U}_1^\vee & = & \mathcal{U}_1^\vee & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & \mathcal{U} & \to & \mathcal{V} & \to & \mathcal{F} & \to & 0 \\ & & \downarrow \alpha & & \downarrow & & \| & & \\ 0 & \to & \mathcal{U}_1 & \to & \mathcal{V}_1 & \to & \mathcal{F} & \to & 0 \end{array}$$

a vector bundle $\mathcal{W}$ over Y and a commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & \mathcal{U} & \to & \mathcal{E} & \to & \mathcal{V}^\vee & \to & 0 \\ & & \| & & \uparrow & & \uparrow & & \\ 0 & \to & \mathcal{U} & \to & \mathcal{W} & \to & \mathcal{V}_1^\vee & \to & 0 \\ & & \downarrow \alpha & & \downarrow & & \| & & \\ 0 & \to & \mathcal{U}_1 & \to & \mathcal{E}_1 & \to & \mathcal{V}_1^\vee & \to & 0 \end{array}$$

where the middle row is also exact. From the dual of the former diagram we get, after taking the tensor product with $\mathcal{O}_Y(-1)$ and taking higher direct images, the commutative diagram

$$\begin{array}{ccccccc} 0 & \to & f_*(\mathcal{F}(-1)) & \to & f_*(\mathcal{V}^\vee(-1)) & \to & f_*(\mathcal{U}^\vee(-1)) \\ & & \| & & \uparrow & & \uparrow \\ 0 & \to & f_*(\mathcal{F}(-1)) & \to & f_*(\mathcal{V}_1^\vee(-1)) & \to & f_*(\mathcal{U}_1^\vee(-1)) \end{array}$$

$$\begin{array}{ccccccccc} \to & R^1f_*(\mathcal{F}(-1)) & \to & R^1f_*(\mathcal{V}^\vee(-1)) & \to & R^1f_*(\mathcal{U}^\vee(-1)) & \to & 0 \\ & \| & & \uparrow & & \uparrow & & \\ \to & R^1f_*(\mathcal{F}(-1)) & \to & R^1f_*(\mathcal{V}_1^\vee(-1)) & \to & R^1f_*(\mathcal{U}_1^\vee(-1)) & \to & 0 \end{array}$$

with exact rows. As already mentioned, $R^1f_*(\mathcal{V}^\vee(-1)) = 0$ and $R^1f_*(\mathcal{U}^\vee(-1)) = 0$. Also, $R^1f_*(\mathcal{U}_1^\vee(-1)) = R^1f_*(f^*(\mathcal{Z}^\vee)) = 0$. Furthermore, the morphism $f_*(\mathcal{U}_1^\vee(-1)) \to f_*(\mathcal{U}^\vee(-1))$ is, by construction, an isomorphism. We conclude that $R^1f_*(\mathcal{V}_1^\vee(-1)) = 0$ and that the morphism $f_*(\mathcal{V}_1^\vee(-1)) \to f_*(\mathcal{V}^\vee(-1))$ is an isomorphism.
Doing the same with the latter diagram we get the commutative diagram

$$\begin{array}{ccccccc} 0 & \to & f_*(\mathcal{U}(-1)) & \to & f_*(\mathcal{E}(-1)) & \to & f_*(\mathcal{V}^\vee(-1)) \\ & & \| & & \uparrow & & \uparrow \\ 0 & \to & f_*(\mathcal{U}(-1)) & \to & f_*(\mathcal{W}(-1)) & \to & f_*(\mathcal{V}_1^\vee(-1)) \\ & & \downarrow & & \downarrow & & \| \\ 0 & \to & f_*(\mathcal{U}_1(-1)) & \to & f_*(\mathcal{E}_1(-1)) & \to & f_*(\mathcal{V}_1^\vee(-1)) \end{array}$$

$$\begin{array}{cccccccc} \to & R^1f_*(\mathcal{U}(-1)) & \to & R^1f_*(\mathcal{E}(-1)) & \to & R^1f_*(\mathcal{V}^\vee(-1)) & \to & 0 \\ & \| & & \uparrow & & \uparrow & & \\ \to & R^1f_*(\mathcal{U}(-1)) & \to & R^1f_*(\mathcal{W}(-1)) & \to & R^1f_*(\mathcal{V}_1^\vee(-1)) & \to & 0 \\ & \downarrow & & \downarrow & & \| & & \\ \to & R^1f_*(\mathcal{U}_1(-1)) & \to & R^1f_*(\mathcal{E}_1(-1)) & \to & R^1f_*(\mathcal{V}_1^\vee(-1)) & \to & 0 \end{array}$$

with exact rows. Using that $R^1f_*(\mathcal{E}(-1)) = 0$, $R^1f_*(\mathcal{V}_1^\vee(-1)) = 0$, and that the morphism $f_*(\mathcal{V}_1^\vee(-1)) \to f_*(\mathcal{V}^\vee(-1))$ is an isomorphism, we see from the upper half of this diagram that $R^1f_*(\mathcal{W}(-1)) = 0$.
The fact that $R^1f_*(\mathcal{U}^\vee(-1)) = 0$, in particular locally free, implies that the Serre duality morphism $R^1f_*(\omega \otimes \mathcal{U}(1)) \to f_*(\mathcal{U}^\vee(-1))^\vee$ is an isomorphism. The same applies to $\mathcal{U}_1$ instead of $\mathcal{U}$. As the morphism $f_*(\mathcal{U}_1^\vee(-1)) \to f_*(\mathcal{U}^\vee(-1))$ is, by construction, an isomorphism, it follows that the induced morphism $R^1f_*(\omega \otimes \mathcal{U}(1)) \to R^1f_*(\omega \otimes \mathcal{U}_1(1))$ is an isomorphism. But $\omega \otimes \mathcal{U}(1) = f^*(\mathcal{L}) \otimes \mathcal{U}(-1)$ and correspondingly for $\mathcal{U}_1$, and modulo the tensor product with $\mathrm{id}_{f^*(\mathcal{L})}$ the morphism $R^1f_*(\omega \otimes \mathcal{U}(1)) \to R^1f_*(\omega \otimes \mathcal{U}_1(1))$ is the morphism $R^1f_*(\mathcal{U}(-1)) \to R^1f_*(\mathcal{U}_1(-1))$ in the diagram. Hence this is an isomorphism. Using that, and the fact that $R^1f_*(\mathcal{W}(-1)) = 0$ and $R^1f_*(\mathcal{V}_1^\vee(-1)) = 0$, we see from the lower half of the diagram that $R^1f_*(\mathcal{E}_1(-1)) = 0$. By Proposition 4 in Section 1.1, it follows that $\mathcal{E}_1 = f^*(\mathcal{G}_1)$ for some symmetric bilinear space $\mathcal{G}_1$ over X.

Section 1.3

In this section we shall show that every element in $W(Y)$ is represented by a symmetric bilinear space over Y that has relatively simple higher direct images.

The natural short exact sequence $0 \to \omega \to f^*(\mathcal{S})(-1) \to \mathcal{O}_Y \to 0$, representing an extension of the trivial vector bundle $\mathcal{O}_Y$ by ω, played a major role in the proof of Proposition 1 in Section 1.1. We next construct something similar for symmetric bilinear spaces.
Using $\frac{1}{2}$ times the natural morphism $(\mathcal{S} \otimes \mathcal{S}^\vee) \times (\mathcal{S} \otimes \mathcal{S}^\vee) \to \mathcal{L} \otimes \mathcal{L}^\vee \cong \mathcal{O}_X$ induced by the exterior product, we get a regular symmetric bilinear form δ on $\mathcal{S} \otimes \mathcal{S}^\vee$. (The corresponding quadratic form on $\mathcal{S} \otimes \mathcal{S}^\vee \cong \mathcal{E}nd(\mathcal{S})$ then is the determinant.) In what follows $\mathcal{S} \otimes \mathcal{S}^\vee$ carries this form.
We have natural morphisms $\varepsilon : \mathcal{O}_X \to \mathcal{S} \otimes \mathcal{S}^\vee$, mapping 1 to the element e corresponding to the identity on $\mathcal{S}$, and $\sigma : \mathcal{S} \otimes \mathcal{S}^\vee \to \mathcal{O}_X$, the contraction (corresponding to the trace). Furthermore, the composition $\sigma \circ \varepsilon$ is 2 times the identity on $\mathcal{O}_X$. It follows that $\mathcal{S} \otimes \mathcal{S}^\vee$ is, as a vector bundle, the direct sum of $\mathcal{O}_X e$ and $\mathcal{T}$, where $\mathcal{T}$ is the kernel of σ. Computations show that this is even a decomposition of $\mathcal{S} \otimes \mathcal{S}^\vee$ as a symmetric bilinear space (and that the induced form on $\mathcal{O}_X$ is the multiplication). We let $-\psi_0$ be the induced form on $\mathcal{T}$. In what follows $\mathcal{T}$ carries the form ψ_0.
From the dual morphism $\pi^\vee : \mathcal{O}_Y \to f^*(\mathcal{S}^\vee)(1)$ to the morphism $\pi : f^*(\mathcal{S})(-1) \to \mathcal{O}_Y$ of the natural short exact sequence we get a morphism

$$f^*(\mathcal{S})(-1) = f^*(\mathcal{S})(-1) \otimes \mathcal{O}_Y \to$$

$$f^*(\mathcal{S})(-1) \otimes f^*(\mathcal{S}^\vee)(1) = f^*(\mathcal{S}) \otimes f^*(\mathcal{S}^\vee) = f^*(\mathcal{S} \otimes \mathcal{S}^\vee)$$

(Easy computations show that this makes $f^*(\mathcal{S})(-1)$ to a Lagrangian of $f^*(\mathcal{S} \otimes \mathcal{S}^\vee)$.) This morphism, composed with the projection $f^*(\mathcal{S} \otimes \mathcal{S}^\vee) \cong$

$f^*(\mathcal{O}_X e) \oplus f^*(\mathcal{T}) \to f^*(\mathcal{T})$, gives us a morphism $\kappa : f^*(\mathcal{S})(-1) \to f^*(\mathcal{T})$. Computations show that $\kappa \circ \iota : \omega \to f^*(\mathcal{T})$ makes ω a totally isotropic subbundle of $f^*(\mathcal{T})$ and that $\kappa : f^*(\mathcal{S})(-1) \to f^*(\mathcal{T})$ makes $f^*(\mathcal{S})(-1)$ the corresponding orthogonal subbundle. (By the way, the image of ω under the morphism $f^*(\mathcal{S})(-1) \to f^*(\mathcal{S} \otimes \mathcal{S}^\vee)$ is, in fact, contained in $f^*(\mathcal{T})$.) Computations now show that the induced bilinearform on $\mathrm{Coker}(\iota) \cong \mathcal{O}_Y$ is precisely the multiplication. This mean that

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \omega & \xrightarrow{\iota} & f^*(\mathcal{S})(-1) & \xrightarrow{\pi} & \mathcal{O}_Y & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \kappa} & & \downarrow{\scriptstyle \pi^\vee} & & \\
0 & \to & \omega & \xrightarrow{\kappa\circ\iota} & f^*(\mathcal{T}) & \xrightarrow{\kappa^\vee\circ\psi} & f^*(\mathcal{S}^\vee)(1) & \to & 0 \\
 & & & & \downarrow{\scriptstyle \iota^\vee\circ\kappa^\vee\circ\psi} & & \downarrow{\scriptstyle \iota^\vee} & & \\
 & & & & \omega^\vee & = & \omega^\vee & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

is a presentation of the bilinear space $f^*(\mathcal{T})$ as an extension of the unit bilinear space $\mathcal{O}_Y$ by ω. Here we have written ψ for the morphism $f^*(\psi_0)$.

For every symmetric bilinear space $\mathcal{F}$ over Y we get through the tensor product a presentation

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \omega\otimes\mathcal{F} & \to & f^*(\mathcal{S})\otimes\mathcal{F}(-1) & \to & \mathcal{F} & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & \omega\otimes\mathcal{F} & \to & f^*(\mathcal{T})\otimes\mathcal{F} & \to & f^*(\mathcal{S}^\vee)\otimes\mathcal{F}^\vee(1) & \to & 0 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & \omega^\vee\otimes\mathcal{F}^\vee & = & \omega^\vee\otimes\mathcal{F}^\vee & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

of $f^*(\mathcal{T}) \otimes \mathcal{F}$ as an extension of $\mathcal{F}$ by $\omega \otimes \mathcal{F}$.

We now assume that $k \geq -1$ and $R^1 f_*(\mathcal{F}(j)) = 0$ for every $j > k$. We have $\omega = f^*(\mathcal{L})(-2)$, hence

$$R^1 f_*(\omega^\vee \otimes \mathcal{F}^\vee(k)) = R^1 f_*(f^*(\mathcal{L}^\vee) \otimes \mathcal{F}^\vee(k+2)) = \mathcal{L}^\vee \otimes R^1 f_*(\mathcal{F}^\vee(k+2)) = 0$$

as $\mathcal{F}^\vee \cong \mathcal{F}$. With the Theorem on Cohomology and Base Change it follows that the coherent $\mathcal{O}_X$-module $\mathcal{W} := f_*(\omega^\vee \otimes \mathcal{F}^\vee(k))$ is locally free. The canonical morphism $f^*(\mathcal{W}) = f^*(f_*(\omega^\vee \otimes \mathcal{F}^\vee(k))) \to \omega^\vee \otimes \mathcal{F}^\vee(k)$ induces a morphism $f^*(\mathcal{W})(-k) \to \omega^\vee \otimes \mathcal{F}^\vee$. Using the dual morphism $\omega \otimes \mathcal{F} \to f^*(\mathcal{W}^\vee)(k)$ on the extension of $\mathcal{F}$ by $\omega \otimes \mathcal{F}$ described above, we get an extension

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & f^*(\mathcal{W}^\vee)(k) & \to & \mathcal{V}_1 & \to & \mathcal{F} & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & f^*(\mathcal{W}^\vee)(k) & \to & \mathcal{E}_1 & \to & \mathcal{V}_1^\vee & \to & 0 \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & f^*(\mathcal{W})(-k) & = & f^*(\mathcal{W})(-k) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

of $\mathcal{F}$ by $f^*(\mathcal{W}^\vee)(k)$ and a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & \mathcal{F} & \to & \mathcal{V}_1^\vee & \to & f^*(\mathcal{W})(-k) & \to & 0 \\
 & & \| & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{F} & \to & f^*(\mathcal{S}^\vee) \otimes \mathcal{F}^\vee(1) & \to & \omega^\vee \otimes \mathcal{F}^\vee & \to & 0
\end{array}$$

Now, $R^1 f_*(f^*(\mathcal{W})(j-k)) = \mathcal{W} \otimes R^1 f_*(\mathcal{O}_Y(j-k)) = 0$ for $j-k \geq -1$. From the exactness of the upper row of this diagram it therefore follows at once that also $R^1 f_*(\mathcal{V}_1^\vee(j)) = 0$ for $j > k$. For $j = k$ we get, as $f_*(f^*(\mathcal{W})) = \mathcal{W}$, the commutative diagram

$$\begin{array}{ccccccc}
\mathcal{W} & \to & R^1 f_*(\mathcal{F}(k)) & \to & R^1 f_*(\mathcal{V}_1^\vee(k)) & \to & 0 \\
\downarrow & & \| & & \downarrow & & \\
f_*(\omega^\vee \otimes \mathcal{F}^\vee(k)) & \to & R^1 f_*(\mathcal{F}(k)) & \to & R^1 f_*(f^*(\mathcal{S}^\vee) \otimes \mathcal{F}(k+1)) & \to &
\end{array}$$

with exact rows. As $R^1 f_*(f^*(\mathcal{S}^\vee) \otimes \mathcal{F}^\vee(k+1)) = \mathcal{S}^\vee \otimes R^1 f_*(\mathcal{F}^\vee(k+1)) = 0$, the connecting morphism $f_*(\omega^\vee \otimes \mathcal{F}^\vee(k)) \to R^1 f_*(\mathcal{F}(k))$ is an epimorphism. But, by construction, the morphism $\mathcal{W} \to f_*(\omega^\vee \otimes \mathcal{F}^\vee(k))$ is the identity, so the connecting morphism $\mathcal{W} \to R^1 f_*(\mathcal{F}(k))$ must also be an epimorphism. It follows that even $R^1 f_*(\mathcal{V}_1^\vee(k)) = 0$.
As $R^1 f_*(f^*(\mathcal{W}^\vee)(k+j)) = \mathcal{W}^\vee \otimes R^1 f_*(\mathcal{O}_Y(k+j)) = 0$ for $k+j \geq -1$, it now follows from the exactness of the sequence $0 \to f^*(\mathcal{W}^\vee)(k) \to \mathcal{E}_1 \to \mathcal{V}_1^\vee \to 0$ that $R^1 f_*(\mathcal{E}_1(j)) = 0$ for $j > k$ and that also $R^1 f_*(\mathcal{E}_1(k)) = 0$ if $k \geq 0$.
By induction on k downwards to $k = 0$ we get:

Theorem 1: Any symmetrical bilinear space $\mathcal{F}$ over Y is equivalent to a symmetric bilinear space $\mathcal{E}$ over Y with $R^1 f_*(\mathcal{E}(j)) = 0$ for every $j \geq 0$
Remark: We know that it follows that $f_*(\mathcal{E}(j))$ is locally free for every $j \geq 0$. Using the duality, it follows that $f_*(\mathcal{E}(j)) = 0$ and $R^1 f_*(\mathcal{E}(j))$ is locally free for every $j \leq -2$

In the case $k = -1$ also we had $R^1 f_*(\mathcal{V}_1^\vee(-1)) = 0$ (but not necessarily $R^1 f_*(\mathcal{E}_1(-1)) = 0$). But by the remark above we have in that case that $f_*(\omega\otimes\mathcal{F})=0$ and $R^1f_*(\omega\otimes\mathcal{F})$ is locally free. As $\omega(1)\otimes f^*(\mathcal{W})=f^*(\mathcal{L}\otimes\mathcal{W})(-1)$, we have $f_*(\omega(1)\otimes f^*(\mathcal{W}))=0$ and $R^1f_*(\omega(1)\otimes f^*(\mathcal{W}))=0$. From the tensor product of the short exact sequence $0\to\mathcal{F}\to\mathcal{V}_1^\vee\to f^*(\mathcal{W})(1)\to 0$ and ω it therefore follows that also $f_*(\omega\otimes\mathcal{V}_1^\vee)=0$ and that $R^1f_*(\omega\otimes\mathcal{V}_1^\vee)$ is isomorphic to $R^1f_*(\omega\otimes\mathcal{F})$, hence locally free. We therefore have:

Theorem 1, cntd.: Furthermore, $\mathcal{E}$ can be chosen to have a totally isotropic subbundle $\mathcal{U}$, isomorphic to $f^*(\mathcal{A})(-1)$ for some vector bundle $\mathcal{A}$ over X, such that $R^1f_*((\mathcal{E}/\mathcal{U})(-1))=0$ and $f_*(\omega\otimes(\mathcal{E}/\mathcal{U}))=0$ and such that $R^1f_*(\omega\otimes(\mathcal{E}/\mathcal{U}))$ is locally free.
Remark: By Proposition 3 in Section 1.1 there is then a short exact sequence $0\to f^*(\mathcal{C})(1)\to\mathcal{E}/\mathcal{U}\to f^*(\mathcal{B})\to 0$ with vector bundles $\mathcal{B}$ and $\mathcal{C}$ over X. If X is affine then it even follows that $\mathcal{E}/\mathcal{U}\cong f^*(\mathcal{B})\oplus f^*(\mathcal{C})(1)$.

Section 1.4

In this section we study higher direct images of the special representatives of elements in $W(Y)$ gotten in the last section. We also study what happens for these under extensions like those considered in Theorem 2 in Section 1.2.

An NN-pair is a pair $((\mathcal{E},\chi),(\mathcal{A},\mu))$, where $(\mathcal{E},\chi)$ is a symmetric bilinear space over Y, $\mathcal{A}$ is a vector bundle over X, and $\mu: f^*(\mathcal{A})(-1)\to\mathcal{E}$ is an embedding of $f^*(\mathcal{A})(-1)$ in $\mathcal{E}$ as a totally isotropic subbundle of $(\mathcal{E},\chi)$ such that for the cokernel $\rho:\mathcal{E}\to\overline{\mathcal{E}}$ we have $R^1f_*(\overline{\mathcal{E}}(-1))=0$, $f_*(\omega\otimes\overline{\mathcal{E}})=0$ and $R^1f_*(\omega\otimes\overline{\mathcal{E}})$ is locally free. Note that it follows that $R^1f_*(\mathcal{E})=0$, hence $R^1f_*(\mathcal{E}(j))=0$ for every $j\geq 0$.
There is an obvious notion of isomorphisms of NN-pairs. Furthermore, we can define the direct sum of two NN-pairs in an obvious way. It follows that we have the Grothendieck group of isomorphism classes of NN-pairs. We denote it here simply by $K(NN)$.
Forgetting the second object in an NN-pair we get a morphism $K(NN)\to W(X)$ of groups. By Theorem 1 in Section 1.3 this is an epimorphism.

Let $((\mathcal{E},\chi),(\mathcal{A},\mu))$ be an NN-pair and let $\rho:\mathcal{E}\to\overline{\mathcal{E}}$ be a cokernel of μ. We write $\mathcal{C}=f_*(\overline{\mathcal{E}}(-1))$ and $\mathcal{B}=R^1f_*(\omega\otimes\overline{\mathcal{E}})$. Then $\mathcal{C}$ and $\mathcal{B}$ are vector bundles over X and there is a natural short exact sequence $0\to f^*(\mathcal{C})(1)\to\overline{\mathcal{E}}\to f^*(\mathcal{B})\to 0$.
As $f^*(\mathcal{A})(-1)$ is a totally isotropic subbundle of $(\mathcal{E},\chi)$, there is a unique morphism $\tau:\overline{\mathcal{E}}\to f^*(\mathcal{A}^\vee)(1)$ making the diagram

$$\begin{array}{ccc}\mathcal{E} & \xrightarrow{\rho} & \overline{\mathcal{E}}\\ \downarrow\chi & & \downarrow\tau\\ \mathcal{E}^\vee & \xrightarrow{\mu^\vee} & f^*(\mathcal{A}^\vee)(1)\end{array}$$

commutative. Using also the dual diagram, we get the commutative diagram

$$\begin{array}{ccccccccc} 0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\mu} & \mathcal{E} & \xrightarrow{\rho} & \overline{\mathcal{E}} & \to & 0 \\ & & \Big\downarrow \tau^\vee & & \Big\downarrow \chi & & \Big\downarrow \tau & & \\ 0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\rho^\vee} & \mathcal{E}^\vee & \xrightarrow{\mu^\vee} & f^*(\mathcal{A}^\vee)(1) & \to & 0 \end{array}$$

with exact rows, the second row being the dual of the first one.
We have $f_*(\overline{\mathcal{E}}(-1)) = \mathcal{C}$ and $R^1 f_*(\overline{\mathcal{E}}(-1)) = 0$. Using the dual of the short exact sequence $0 \to f^*(\mathcal{C})(1) \to \overline{\mathcal{E}} \to f^*(\mathcal{B}) \to 0$, we see that $f_*(\overline{\mathcal{E}}^\vee(-1)) = 0$ and that we can write $R^1 f_*(\overline{\mathcal{E}}^\vee(-1)) = \mathcal{L}^\vee \otimes \mathcal{C}^\vee$. Twisting the last diagram above and taking higher direct images, we therefore get the commutative diagram

$$\begin{array}{ccccccccccc} 0 & \to & f_*(\mathcal{E}(-1)) & \to & \mathcal{C} & \to & \mathcal{L}^\vee \otimes \mathcal{A} & \to & R^1 f_*(\mathcal{E}(-1)) & \to & 0 \\ & & \Big\downarrow \cong & & \downarrow & & \downarrow & & \Big\downarrow \cong & & \\ 0 & \to & f_*(\mathcal{E}^\vee(-1)) & \to & \mathcal{A}^\vee & \to & \mathcal{L}^\vee \otimes \mathcal{C}^\vee & \to & R^1 f_*(\mathcal{E}^\vee(-1)) & \to & 0 \end{array}$$

with exact rows.
We denote by $\alpha : \mathcal{C} \to \mathcal{L}^\vee \otimes \mathcal{A}$ the connecting morphism in the upper row and by $\varepsilon : \mathcal{C} \to \mathcal{A}^\vee$ the second vertical morphism. Then, by Serre duality, the connecting morphism in the lower row is $-1_{\mathcal{L}^\vee} \otimes \alpha^\vee : \mathcal{A}^\vee \to \mathcal{L}^\vee \otimes \mathcal{C}^\vee$ and the third vertical morphism is $1_{\mathcal{L}^\vee} \otimes \varepsilon^\vee : \mathcal{L}^\vee \otimes \mathcal{A} \to \mathcal{L}^\vee \otimes \mathcal{C}^\vee$. The exactness of the diagram is therefore seen to mean that $\begin{bmatrix} \alpha \\ \varepsilon \end{bmatrix} : \mathcal{C} \to (\mathcal{L}^\vee \otimes \mathcal{A}) \oplus \mathcal{A}^\vee$ is an embedding of $\mathcal{C}$ in $(\mathcal{L}^\vee \otimes \mathcal{A}) \oplus \mathcal{A}^\vee$ as a Lagrangian of the hyperbolic $\mathcal{L}^\vee$-valued symmetric bilinear space $\left((\mathcal{L}^\vee \otimes \mathcal{A}) \oplus \mathcal{A}^\vee, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\right)$.

Let $((\mathcal{E}, \chi), (\mathcal{A}, \mu))$ be an NN-pair and let $\mathcal{Z}$ be a vector bundle over X. Let $(\mathcal{E}_1, \chi_1)$ be an extension of $(\mathcal{E}, \chi)$ by $f^*(\mathcal{Z})(-1)$ with presentation

$$\begin{array}{ccccccccc} & & & & 0 & & 0 & & \\ & & & & \downarrow & & \downarrow & & \\ 0 & \to & f^*(\mathcal{Z})(-1) & \xrightarrow{\iota} & \mathcal{V} & \xrightarrow{\pi} & \mathcal{E} & \to & 0 \\ & & \| & & \Big\downarrow \kappa & & \Big\downarrow \pi^\vee \circ \chi & & \\ 0 & \to & f^*(\mathcal{Z})(-1) & \longrightarrow & \mathcal{E}_1 & \xrightarrow{\kappa^\vee \circ \chi_1} & \mathcal{V}^\vee & \to & 0 \\ & & & & \Big\downarrow & & \Big\downarrow \iota^\vee & & \\ & & & & f^*(\mathcal{Z}^\vee)(1) & = & f^*(\mathcal{Z}^\vee)(1) & & \\ & & & & \downarrow & & \downarrow & & \\ & & & & 0 & & 0 & & \end{array}$$

From Proposition 1 in Section 1.1 it follows at once that any extension of $f^*(\mathcal{A})(-1)$ by $f^*(\mathcal{Z})(-1)$ can be written as $f^*(\mathcal{A}_1)(-1)$ for some vector bundle $\mathcal{A}_1$ over X. It then comes from a unique extension

$$0 \to \mathcal{Z} \xrightarrow{\iota_{\mathcal{A}}} \mathcal{A}_1 \xrightarrow{\pi_{\mathcal{A}}} \mathcal{A} \to 0$$

of vector bundles over X. Taking the pull-back of

$$\begin{tikzcd}
 & f^*(\mathcal{A})(-1) \arrow[d, "\mu"] \\
\mathcal{V} \arrow[r, "\pi"] & \mathcal{E}
\end{tikzcd}$$

we therefore get an exact commutative diagram

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & f^*(\mathcal{Z})(-1) \arrow[r, "f^*(\iota_{\mathcal{A}})(-1)"] \arrow[d, equal] & f^*(\mathcal{A}_1)(-1) \arrow[r, "f^*(\pi_{\mathcal{A}})(-1)"] \arrow[d, "\mu_{\mathcal{V}}"] & f^*(\mathcal{A})(-1) \arrow[r] \arrow[d, "\mu"] & 0 \\
0 \arrow[r] & f^*(\mathcal{Z})(-1) \arrow[r, "\iota"] & \mathcal{V} \arrow[r, "\pi"] \arrow[d] & \mathcal{E} \arrow[r] \arrow[d, "\rho"] & 0 \\
 & & \overline{\mathcal{E}} \arrow[r, equal] \arrow[d] & \overline{\mathcal{E}} \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

uniquely determined up to an isomorphism of $\mathcal{A}_1$. As $f^*(\mathcal{Z})(-1)$ is a totally isotropic subbundle of $(\mathcal{E}_1, \chi_1)$ and the quotient $f^*(\mathcal{A})(-1)$ is a totally isotropic subbundle of $(\mathcal{E}, \chi)$, it is clear that the composition $\mu_1 = \kappa \circ \mu_{\mathcal{V}} : f^*(\mathcal{A}_1)(-1) \to \mathcal{E}_1$ is an embedding of $f^*(\mathcal{A}_1)(-1)$ in $\mathcal{E}_1$ as a totally isotropic subbundle of $(\mathcal{E}_1, \chi_1)$.

Taking the push-out of

$$\begin{tikzcd}
\mathcal{V} \arrow[r, "\rho\circ\pi"] \arrow[d, "\kappa"] & \overline{\mathcal{E}} \\
\mathcal{E}_1 &
\end{tikzcd}$$

we now get an exact commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_{\mathcal{V}}} & \mathcal{V} & \xrightarrow{\rho\circ\pi} & \overline{\mathcal{E}} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle\kappa} & & \downarrow{\scriptstyle\lambda} & & \\
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_1} & \mathcal{E}_1 & \xrightarrow{\rho_1} & \overline{\mathcal{E}}_1 & \to & 0 \\
 & & & & \downarrow{\scriptstyle\iota^\vee\circ\kappa^\vee\circ\chi_1} & & \downarrow{\scriptstyle\sigma} & & \\
 & & & & f^*(\mathcal{Z}^\vee)(1) & = & f^*(\mathcal{Z}^\vee)(1) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

From the right hand column in this diagram we get, using our hypotheses on $\overline{\mathcal{E}}$, that $R^1f_*(\overline{\mathcal{E}}_1(-1)) = 0$, $f_*(\omega\otimes\overline{\mathcal{E}}_1) = 0$ and $R^1f_*(\omega\otimes\overline{\mathcal{E}}_1)$ is locally free. (In fact, $R^1f_*(\omega\otimes\overline{\mathcal{E}}_1)$ is isomorphic to $R^1f_*(\omega\otimes\overline{\mathcal{E}})$.) So $((\mathcal{E}_1,\chi_1),(\mathcal{A}_1,\mu_1))$ is an NN-pair.
In this situation we say that the NN-pair $((\mathcal{E}_1,\chi_1),(\mathcal{A}_1,\mu_1))$ is an extension of the NN-pair $((\mathcal{E},\chi),(\mathcal{A},\mu))$ by $\mathcal{Z}$.

Let the NN-pair $((\mathcal{E}_1,\chi_1),(\mathcal{A}_1,\mu_1))$ be an extension of the NN-pair $((\mathcal{E},\chi),(\mathcal{A},\mu))$. We keep the notations from above and extend them in the obvious way. In particular, we have the short exact sequence

$$0 \to \mathcal{Z} \xrightarrow{\iota_{\mathcal{A}}} \mathcal{A}_1 \xrightarrow{\pi_{\mathcal{A}}} \mathcal{A} \to 0$$

of vector bundles over X. Furthermore, by twisting and taking higher direct images, the right hand column of the third big diagram induces a short exact sequence

$$0 \to \mathcal{C} \xrightarrow{\iota_{\mathcal{C}}} \mathcal{C}_1 \xrightarrow{\pi_{\mathcal{C}}} \mathcal{Z}^\vee \to 0$$

of vector bundles over X.
We have $\tau_1\circ\lambda\circ\rho\circ\pi = \tau_1\circ\rho_1\circ\kappa = \mu_1^\vee\circ\chi_1\circ\kappa = \mu_{\mathcal{V}}^\vee\circ\kappa^\vee\circ\chi_1\circ\kappa = \mu_{\mathcal{V}}^\vee\circ\pi^\vee\circ\chi\circ\pi = f^*(\pi_{\mathcal{A}}^\vee)(1)\circ\mu^\vee\circ\chi\circ\pi = f^*(\pi_{\mathcal{A}}^\vee)(1)\circ\tau\circ\rho\circ\pi$. As $\rho\circ\pi$ is an epimorphism, it follows that $\tau_1\circ\lambda = f^*(\pi_{\mathcal{A}}^\vee)(1)\circ\tau$. We also have $f^*(\iota_{\mathcal{A}}^\vee)(1)\circ\tau_1\circ\rho_1 = f^*(\iota_{\mathcal{A}}^\vee)(1)\circ\mu_1^\vee\circ\chi_1 = f^*(\iota_{\mathcal{A}}^\vee)(1)\circ\mu_{\mathcal{V}}^\vee\circ\kappa^\vee\circ\chi_1 = \iota^\vee\circ\kappa^\vee\circ\chi_1 = \sigma\circ\rho_1$. As ρ_1 is an epimorphism, it follows that $f^*(\iota_{\mathcal{A}}^\vee)(1)\circ\tau_1 = \sigma$. This shows that the diagram

$$\begin{array}{ccccccccc}
0 & \to & \overline{\mathcal{E}} & \xrightarrow{\lambda} & \overline{\mathcal{E}}_1 & \xrightarrow{\sigma} & f^*(\mathcal{Z}^\vee)(1) & \to & 0 \\
 & & \downarrow{\scriptstyle\tau} & & \downarrow{\scriptstyle\tau_1} & & \| & & \\
0 & \to & f^*(\mathcal{A}^\vee)(1) & \xrightarrow{f^*(\pi_{\mathcal{A}}^\vee)(1)} & f^*(\mathcal{A}_1^\vee)(1) & \xrightarrow{f^*(\iota_{\mathcal{A}}^\vee)(1)} & f^*(\mathcal{Z}^\vee)(1) & \to & 0
\end{array}$$

is commutative. Twisting by -1 and taking higher direct images, we therefore get the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C} \arrow[r, "\iota_{\mathcal{C}}"] \arrow[d, "\varepsilon"] & \mathcal{C}_1 \arrow[r, "\pi_{\mathcal{C}}"] \arrow[d, "\varepsilon_1"] & \mathcal{Z}^\vee \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{A}^\vee \arrow[r, "\pi_{\mathcal{A}}^\vee"] & \mathcal{A}_1^\vee \arrow[r, "\iota_{\mathcal{A}}^\vee"] & \mathcal{Z}^\vee \arrow[r] & 0
\end{tikzcd}$$

with exact rows.
We have the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & f^*(\mathcal{A})(-1) \arrow[r, "\mu"] & \mathcal{E} \arrow[r, "\rho"] & \overline{\mathcal{E}} \arrow[r] & 0 \\
0 \arrow[r] & f^*(\mathcal{A}_1)(-1) \arrow[r, "\mu_{\mathcal{V}}"] \arrow[u, "f^*(\pi_{\mathcal{A}})(-1)"'] \arrow[d, equal] & \mathcal{V} \arrow[r, "\rho\circ\pi"] \arrow[u, "\pi"'] \arrow[d, "\kappa"] & \overline{\mathcal{E}} \arrow[r] \arrow[u, equal] \arrow[d, "\lambda"] & 0 \\
0 \arrow[r] & f^*(\mathcal{A}_1)(-1) \arrow[r, "\mu_1"] & \mathcal{E}_1 \arrow[r, "\rho_1"] & \overline{\mathcal{E}}_1 \arrow[r] & 0
\end{tikzcd}$$

with exact rows. Twisting it by -1 and looking at the connecting morphisms for the higher direct images, we get the commutative diagram

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "\alpha"] & \mathcal{L}^\vee \otimes \mathcal{A} \\
\mathcal{C} \arrow[r] \arrow[u, equal] \arrow[d, "\iota_{\mathcal{C}}"] & \mathcal{L}^\vee \otimes \mathcal{A}_1 \arrow[u, "1_{\mathcal{L}^\vee}\otimes\pi_{\mathcal{A}}"'] \arrow[d, equal] \\
\mathcal{C}_1 \arrow[r, "\alpha_1"] & \mathcal{L}^\vee \otimes \mathcal{A}_1
\end{tikzcd}$$

It follows that the diagram

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "\alpha"] \arrow[d, "\iota_{\mathcal{C}}"] & \mathcal{L}^\vee \otimes \mathcal{A} \\
\mathcal{C}_1 \arrow[r, "\alpha_1"] & \mathcal{L}^\vee \otimes \mathcal{A}_1 \arrow[u, "1_{\mathcal{L}^\vee}\otimes\pi_{\mathcal{A}}"']
\end{tikzcd}$$

is commutative.

We close this section with an example that we shall need later.

Let $\mathcal{C}$ be a vector bundle over X. We shall use the natural short exact sequence

$$0 \rightarrow f^*(\mathcal{L}\otimes\mathcal{C})(-1) \xrightarrow{\iota} f^*(\mathcal{S}\otimes\mathcal{C}) \xrightarrow{\pi} f^*(\mathcal{C})(1) \rightarrow 0$$

of vector bundles over Y.

Let $\eta : f^*(\mathcal{C})(1) \to f^*(\mathcal{L}^\vee \otimes \mathcal{C}^\vee)(1)$ be a morphism. As morphisms from $f^*(\mathcal{S} \otimes \mathcal{C})$ are uniquely given by their direct images and as $f_*(\iota^\vee)$ is an isomorphism, there is a unique morphism $\xi : f^*(\mathcal{S} \otimes \mathcal{C}) \to f^*(\mathcal{S}^\vee \otimes \mathcal{C}^\vee)$ such that $\iota^\vee \circ \xi = \eta \circ \pi$. Clearly, we then have $\iota^\vee \circ \xi \circ \iota = 0$.
Conversely, if $\xi : f^*(\mathcal{S} \otimes \mathcal{C}) \to f^*(\mathcal{S}^\vee \otimes \mathcal{C}^\vee)$ is a morphism such that $\iota^\vee \circ \xi \circ \iota = 0$ then, by the exactness of the natural sequence, there is a unique morphism $\eta : f^*(\mathcal{C})(1) \to f^*(\mathcal{L}^\vee \otimes \mathcal{C}^\vee)(1)$ such that $\iota^\vee \circ \xi = \eta \circ \pi$.

Let $\xi : f^*(\mathcal{S} \otimes \mathcal{C}) \to f^*(\mathcal{S}^\vee \otimes \mathcal{C}^\vee)$ be given. Let η be the corresponding morphism, so $\iota^\vee \circ \xi = \eta \circ \pi$. Also let η' be the morphism corresponding to $\xi^\vee$, so $\iota^\vee \circ \xi^\vee = \eta' \circ \pi$.
Write $\mathcal{E} = f^*(\mathcal{S} \otimes \mathcal{C}) \oplus f^*(\mathcal{S}^\vee \otimes \mathcal{C}^\vee)$, $\mathcal{A} = \mathcal{L} \otimes \mathcal{C}$, $\overline{\mathcal{E}} = f^*(\mathcal{C})(1) \oplus f^*(\mathcal{S}^\vee \otimes \mathcal{C}^\vee)$, $\mu = \begin{bmatrix} \iota \\ 0 \end{bmatrix} : f^*(\mathcal{A})(-1) \to \mathcal{E}$, and $\rho = \begin{bmatrix} \pi & 0 \\ 0 & 1 \end{bmatrix} : \mathcal{E} \to \overline{\mathcal{E}}$. Then the sequence

$$0 \to f^*(\mathcal{A})(-1) \xrightarrow{\mu} \mathcal{E} \xrightarrow{\rho} \overline{\mathcal{E}} \to 0$$

is exact. Furthermore, $f_*(\overline{\mathcal{E}}(-1)) = \mathcal{C}$, $R^1 f_*(f^*(\mathcal{A})(-2)) = \mathcal{L}^\vee \otimes \mathcal{A} = \mathcal{C}$ and the connecting morphism for this sequence twisted by -1 is $1_{\mathcal{C}}$.
Let $\chi = \begin{bmatrix} \xi & 1 \\ 1 & 0 \end{bmatrix} : \mathcal{E} \to \mathcal{E}^\vee$. Then χ is an isomorphism. Let $\tau = \begin{bmatrix} \eta & \iota^\vee \end{bmatrix} : \overline{\mathcal{E}} \to f^*(\mathcal{A}^\vee)(1)$ and $\tau' = \begin{bmatrix} (\eta')^\vee \\ \iota \end{bmatrix} : f^*(\mathcal{A})(-1) \to \overline{\mathcal{E}}^\vee$. Then we have the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & f^*(\mathcal{A})(-1) \arrow[r, "\mu"] \arrow[d, "\tau'"] & \mathcal{E} \arrow[r, "\rho"] \arrow[d, "\chi"] & \overline{\mathcal{E}} \arrow[r] \arrow[d, "\tau"] & 0 \\
0 \arrow[r] & \overline{\mathcal{E}}^\vee \arrow[r, "\rho^\vee"] & \mathcal{E}^\vee \arrow[r, "\mu^\vee"] & f^*(\mathcal{A}^\vee)(1) \arrow[r] & 0
\end{tikzcd}$$

where the bottom row is the dual of the top one. Twisting by -1 and taking higher direct images, we get the commutative diagram

$$\begin{tikzcd}
\mathcal{C} \arrow[r, "1"] \arrow[d, "f_*(\eta(-1))"] & \mathcal{C} \arrow[d, "R^1 f_*((\eta')^\vee(-1))"] \\
\mathcal{L}^\vee \otimes \mathcal{C}^\vee \arrow[r, "-1"] & \mathcal{L}^\vee \otimes \mathcal{C}^\vee
\end{tikzcd}$$

for the connecting morphisms. This means that $R^1 f_*((\eta')^\vee(-1)) = -f_*(\eta(-1))$.
There are unique morphisms $\varepsilon, \varepsilon' : \mathcal{C} \to \mathcal{L}^\vee \otimes \mathcal{C}^\vee$ such that $\eta = f^*(\varepsilon)(1)$ and $\eta' = f^*(\varepsilon')(1)$. Then $f_*(\eta(-1)) = \varepsilon$ and $R^1 f_*((\eta')^\vee(-1)) = R^1 f_*(f^*((\varepsilon')^\vee)(-2)) = 1_{\mathcal{L}^\vee} \otimes (\varepsilon')^\vee$. So we have $1_{\mathcal{L}^\vee} \otimes (\varepsilon')^\vee = -\varepsilon$, which is equivalent to $\varepsilon' = -1_{\mathcal{L}^\vee} \otimes \varepsilon^\vee$.
We have $\xi^\vee = \xi$ if and only if $\eta' = \eta$. But $\eta' = \eta$ means exactly that $\varepsilon' = \varepsilon$. We conclude that $\xi^\vee = \xi$ if and only if $1_{\mathcal{L}^\vee} \otimes \varepsilon^\vee = -\varepsilon$. In that case the

computations above show that $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{C},\mu))$ is an NN-pair such that the corresponding α is the identity on $\mathcal{C}$ and the corresponding ε is the given one.

Section 1.5

In this section we show how to construct extensions with given behaviour, as in Section 1.4, for the higher direct images. (This turned out to be the hardest part of all.)

Let $((\mathcal{E},\chi),(\mathcal{A},\mu))$ be an NN-pair. We keep the notations from Section 1.4.

Let $\mathcal{A}_1$ and $\mathcal{C}_1$ be vector bundles over X and let $\begin{bmatrix}\alpha_1\\ \varepsilon_1\end{bmatrix} : \mathcal{C}_1 \to (\mathcal{L}^\vee\otimes\mathcal{A}_1)\oplus\mathcal{A}_1^\vee$ be an embedding of $\mathcal{C}_1$ in $(\mathcal{L}^\vee\otimes\mathcal{A}_1)\oplus\mathcal{A}_1^\vee$ as a Lagrangian of the hyperbolic $\mathcal{L}^\vee$-valued symmetric bilinear space $\left((\mathcal{L}^\vee\otimes\mathcal{A}_1)\oplus\mathcal{A}_1^\vee, \begin{bmatrix}0&1\\1&0\end{bmatrix}\right)$. Assume also that there is a vector bundle $\mathcal{Z}$ over X and short exact sequences

$$0 \to \mathcal{Z} \xrightarrow{\iota_\mathcal{A}} \mathcal{A}_1 \xrightarrow{\pi_\mathcal{A}} \mathcal{A} \to 0$$

and

$$0 \to \mathcal{C} \xrightarrow{\iota_\mathcal{C}} \mathcal{C}_1 \xrightarrow{\pi_\mathcal{C}} \mathcal{Z}^\vee \to 0$$

such that the diagrams

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{C} \arrow[r, "\iota_\mathcal{C}"] \arrow[d, "\varepsilon"] & \mathcal{C}_1 \arrow[r, "\pi_\mathcal{C}"] \arrow[d, "\varepsilon_1"] & \mathcal{Z}^\vee \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & \mathcal{A}^\vee \arrow[r, "\pi_\mathcal{A}^\vee"] & \mathcal{A}_1^\vee \arrow[r, "\iota_\mathcal{A}^\vee"] & \mathcal{Z}^\vee \arrow[r] & 0 \end{tikzcd}$$

and

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\alpha"] \arrow[d, "\iota_\mathcal{C}"] & \mathcal{L}^\vee\otimes\mathcal{A} \\ \mathcal{C}_1 \arrow[r, "\alpha_1"] & \mathcal{L}^\vee\otimes\mathcal{A}_1 \arrow[u, "1_{\mathcal{L}^\vee}\otimes\pi_\mathcal{A}"'] \end{tikzcd}$$

are commutative.

We want to show that there is an extension $((\mathcal{E}_1,\chi_1),(\mathcal{A}_1,\mu_1))$ of $((\mathcal{E},\chi),(\mathcal{A},\mu))$ giving rise to this data as in Section 1.4.

As $R^1f_*(\overline{\mathcal{E}}(-1)) = 0$, we have $\mathrm{Ext}_Y(f^*(\mathcal{Z}^\vee)(1),\overline{\mathcal{E}}) = \mathrm{Ext}_Y(f^*(\mathcal{Z}^\vee),\overline{\mathcal{E}}(-1)) \cong \mathrm{Ext}_X(\mathcal{Z}^\vee, f_*(\overline{\mathcal{E}}(-1))) = \mathrm{Ext}_X(\mathcal{Z}^\vee,\mathcal{C})$. We therefore have a unique extension

$$0 \to \overline{\mathcal{E}} \xrightarrow{\lambda} \overline{\mathcal{E}}_1 \xrightarrow{\sigma} f^*(\mathcal{Z}^\vee)(1) \to 0$$

such that $f_*(\overline{\mathcal{E}}_1(-1)) = \mathcal{C}_1$ and such that the given sequence $0\to\mathcal{C}\to\mathcal{C}_1\to\mathcal{Z}^\vee\to 0$ is precisely the sequence of direct images for the twisted sequence

$0 \to \overline{\mathcal{E}}(-1) \to \overline{\mathcal{E}}_1(-1) \to f^*(\mathcal{Z}^\vee) \to 0$. It is clear that $R^1 f_*(\overline{\mathcal{E}}(-1)) = 0$ implies that also $R^1 f_*(\overline{\mathcal{E}}_1(-1)) = 0$. Looking at the tensor product $0 \to \omega \otimes \overline{\mathcal{E}} \to \omega \otimes \overline{\mathcal{E}}_1 \to \omega(1) \otimes f^*(\mathcal{Z}^\vee) \to 0$, we also get at once that $f_*(\omega \otimes \overline{\mathcal{E}}_1) \cong f_*(\omega \otimes \overline{\mathcal{E}})$ and $R^1 f_*(\omega \otimes \overline{\mathcal{E}}_1) \cong R^1 f_*(\omega \otimes \overline{\mathcal{E}})$. In particular, $f_*(\omega \otimes \overline{\mathcal{E}}_1) = 0$ and $R^1 f_*(\omega \otimes \overline{\mathcal{E}}_1)$ is locally free.

Instead of $\overline{\mathcal{E}}$ and the given sequence $0 \to \mathcal{C} \to \mathcal{C}_1 \to \mathcal{Z}^\vee \to 0$ we could have used $f^*(\mathcal{A}^\vee)(1)$ and the dual of the given sequence $0 \to \mathcal{Z} \to \mathcal{A}_1 \to \mathcal{A} \to 0$ in the construction above. But we know that the resulting extension then is represented by $0 \to f^*(\mathcal{A}^\vee)(1) \to f^*(\mathcal{A}_1^\vee)(1) \to f^*(\mathcal{Z}^\vee)(1) \to 0$. By hypothesis, $\mathrm{Ext}_X(\mathcal{Z}^\vee, \varepsilon)$ maps the class of $0 \to \mathcal{C} \to \mathcal{C}_1 \to \mathcal{Z}^\vee \to 0$ to the class of $0 \to \mathcal{A}^\vee \to \mathcal{A}_1^\vee \to \mathcal{Z}^\vee \to 0$. As $f_*(\tau(-1)) = \varepsilon$, it follows that $\mathrm{Ext}_Y(f^*(\mathcal{Z}^\vee(1), \tau)$ maps the class of $0 \to \overline{\mathcal{E}} \to \overline{\mathcal{E}}_1 \to f^*(\mathcal{Z}^\vee)(1) \to 0$ to the class of $0 \to f^*(\mathcal{A}^\vee)(1) \to f^*(\mathcal{A}_1^\vee)(1) \to f^*(\mathcal{Z}^\vee)(1) \to 0$. So we have a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & \overline{\mathcal{E}} & \xrightarrow{\lambda} & \overline{\mathcal{E}}_1 & \xrightarrow{\sigma} & f^*(\mathcal{Z}^\vee)(1) & \to & 0 \\
 & & \downarrow{\scriptstyle\tau} & & \downarrow{\scriptstyle\tau_1} & & \| & & \\
0 & \to & f^*(\mathcal{A}^\vee)(1) & \xrightarrow{f^*(\pi_{\mathcal{A}}^\vee)(1)} & f^*(\mathcal{A}_1^\vee)(1) & \xrightarrow{f^*(\iota_{\mathcal{A}}^\vee)(1)} & f^*(\mathcal{Z}^\vee)(1) & \to & 0
\end{array}$$

Here τ_1 is not uniquely determined. But using that the diagram

$$\begin{array}{ccccccccc}
0 & \to & \mathcal{C} & \xrightarrow{\iota_{\mathcal{C}}} & \mathcal{C}_1 & \xrightarrow{\pi_{\mathcal{C}}} & \mathcal{Z}^\vee & \to & 0 \\
 & & \downarrow{\scriptstyle\varepsilon} & & \downarrow{\scriptstyle\varepsilon_1} & & \| & & \\
0 & \to & \mathcal{A}^\vee & \xrightarrow{\pi_{\mathcal{A}}^\vee} & \mathcal{A}_1^\vee & \xrightarrow{\iota_{\mathcal{A}}^\vee} & \mathcal{Z}^\vee & \to & 0
\end{array}$$

is commutative and that $\mathrm{Hom}_Y(f^*(\mathcal{Z}^\vee)(1), f^*(\mathcal{A}^\vee)(1)) \cong \mathrm{Hom}_X(\mathcal{Z}^\vee, \mathcal{A}^\vee)$, we see that we can choose τ_1 uniquely in such a way that $f_*(\tau_1(-1)) = \varepsilon_1$.

We now use the results on "Special extensions" in the appendix to this section. By hypothesis,

$$0 \to f^*(\mathcal{A})(-1) \xrightarrow{\mu} \mathcal{E} \xrightarrow{\rho} \overline{\mathcal{E}} \to 0$$

corresponds to $\alpha : \mathcal{C} \to \mathcal{L}^\vee \otimes \mathcal{A}$. Of course, we let

$$0 \to f^*(\mathcal{A}_1)(-1) \xrightarrow{\mu_1} \mathcal{E}_1 \xrightarrow{\rho_1} \overline{\mathcal{E}}_1 \to 0$$

correspond to $\alpha_1 : \mathcal{C}_1 \to \mathcal{L}^\vee \otimes \mathcal{A}_1$. We also let

$$0 \to f^*(\mathcal{A}_1)(-1) \xrightarrow{\mu_{\mathcal{V}}} \mathcal{V} \xrightarrow{\rho_{\mathcal{V}}} \overline{\mathcal{E}} \to 0$$

correspond to $\alpha_1 \circ \iota_{\mathcal{C}} : \mathcal{C} \to \mathcal{L}^\vee \otimes \mathcal{A}_1$ and

$$0 \to f^*(\mathcal{A})(-1) \xrightarrow{\widetilde{\mu_{\mathcal{V}}}} \widetilde{\mathcal{V}} \xrightarrow{\widetilde{\rho_{\mathcal{V}}}} \overline{\mathcal{E}}_1 \to 0$$

correspond to $(1_{\mathcal{L}^\vee} \otimes \pi_{\mathcal{A}}) \circ \alpha_1 : \mathcal{C}_1 \to \mathcal{L}^\vee \otimes \mathcal{A}$. Using λ and $f^*(\pi_{\mathcal{A}})(-1)$ we then get the commutative diagrams

$$\begin{array}{ccccccccc}
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_{\mathcal{V}}} & \mathcal{V} & \xrightarrow{\rho_{\mathcal{V}}} & \overline{\mathcal{E}} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \kappa} & & \downarrow{\scriptstyle \lambda} & & \\
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_1} & \mathcal{E}_1 & \xrightarrow{\rho_1} & \overline{\mathcal{E}}_1 & \to & 0 \\
 & & \downarrow{\scriptstyle f^*(\pi_A)(-1)} & & \downarrow{\scriptstyle \widetilde{\kappa}} & & \| & & \\
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\widetilde{\mu_{\mathcal{V}}}} & \widetilde{\mathcal{V}} & \xrightarrow{\widetilde{\rho_{\mathcal{V}}}} & \overline{\mathcal{E}}_1 & \to & 0
\end{array}$$

and

$$\begin{array}{ccccccccc}
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_{\mathcal{V}}} & \mathcal{V} & \xrightarrow{\rho_{\mathcal{V}}} & \overline{\mathcal{E}} & \to & 0 \\
 & & \downarrow{\scriptstyle f^*(\pi_A)(-1)} & & \downarrow{\scriptstyle \pi} & & \| & & \\
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\mu} & \mathcal{E} & \xrightarrow{\rho} & \overline{\mathcal{E}} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \widetilde{\pi}} & & \downarrow{\scriptstyle \lambda} & & \\
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\widetilde{\mu_{\mathcal{V}}}} & \widetilde{\mathcal{V}} & \xrightarrow{\widetilde{\rho_{\mathcal{V}}}} & \overline{\mathcal{E}}_1 & \to & 0
\end{array}$$

We also get that the compositions $\widetilde{\kappa} \circ \kappa$ and $\widetilde{\pi} \circ \pi$ are equal.

We know the kernel and cokernel of λ. Using that we can extend the bottom half of the last diagram to the exact commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\mu} & \mathcal{E} & \xrightarrow{\rho} & \overline{\mathcal{E}} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \widetilde{\pi}} & & \downarrow{\scriptstyle \lambda} & & \\
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\widetilde{\mu_{\mathcal{V}}}} & \widetilde{\mathcal{V}} & \xrightarrow{\widetilde{\rho_{\mathcal{V}}}} & \overline{\mathcal{E}}_1 & \to & 0 \\
 & & & & \downarrow{\scriptstyle \sigma \circ \widetilde{\rho_{\mathcal{V}}}} & & \downarrow{\scriptstyle \sigma} & & \\
 & & & & f^*(\mathcal{Z}^\vee)(1) & = & f^*(\mathcal{Z}^\vee)(1) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

Taking the composition of the right hand half of this diagram with the diagram defining τ_1 and using the exactness of the sequence

$$0 \to \overline{\mathcal{E}}^\vee \xrightarrow{\chi^{-1} \circ \rho^\vee} \mathcal{E} \xrightarrow{\mu^\vee \circ \chi} f^*(\mathcal{A}^\vee)(1) \to 0$$

noting that $\tau \circ \rho = \mu^\vee \circ \chi$, we get the exact commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\chi^{-1}\circ\rho^\vee} & \mathcal{E} & \xrightarrow{\tau\circ\rho} & f^*(\mathcal{A}^\vee)(1) & \tau o & 0 \\
 & & \| & & \big\downarrow{\scriptstyle \widetilde{\pi}} & & \big\downarrow{\scriptstyle f^*(\pi_{\mathcal{A}}^\vee)(1)} & & \\
0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\widetilde{\pi}\circ\chi^{-1}\circ\rho^\vee} & \widetilde{\mathcal{V}} & \xrightarrow{\tau_1\circ\widetilde{\rho_{\mathcal{V}}}} & f^*(\mathcal{A}_1^\vee)(1) & \to & 0 \\
 & & & & \big\downarrow{\scriptstyle \sigma\circ\widetilde{\rho_{\mathcal{V}}}} & & \big\downarrow{\scriptstyle f^*(\iota_{\mathcal{A}}^\vee)(1)} & & \\
 & & & & f^*(\mathcal{Z}^\vee)(1) & = & f^*(\mathcal{Z}^\vee)(1) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

In particular, the middle row is exact. Using that $\chi^{-1} \circ \rho^\vee \circ \tau^\vee = \mu$, we now have the commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & f^*(\mathcal{A})(-1) & \xrightarrow{\widetilde{\mu_{\mathcal{V}}}} & \widetilde{\mathcal{V}} & \xrightarrow{\widetilde{\rho_{\mathcal{V}}}} & \overline{\mathcal{E}}_1 & \to & 0 \\
 & & \big\downarrow{\scriptstyle \tau^\vee} & & \| & & \big\downarrow{\scriptstyle \tau_1} & & \\
0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\widetilde{\pi}\circ\chi^{-1}\circ\rho^\vee} & \widetilde{\mathcal{V}} & \xrightarrow{\tau_1\circ\widetilde{\rho_{\mathcal{V}}}} & f^*(\mathcal{A}_1^\vee)(1) & \to & 0
\end{array}$$

with exact rows. Twisting by -1 and taking higher direct images we get the commutative diagram

$$\begin{array}{ccc}
\mathcal{C}_1 & \xrightarrow{(1_{\mathcal{L}^\vee}\otimes\pi_{\mathcal{A}})\circ\alpha_1} & \mathcal{L}^\vee \otimes \mathcal{A} \\
\big\downarrow{\scriptstyle \varepsilon_1} & & \big\downarrow{\scriptstyle 1_{\mathcal{L}^\vee}\otimes\varepsilon^\vee} \\
\mathcal{A}_1^\vee & \xrightarrow{\alpha'} & \mathcal{L}^\vee \otimes \mathcal{C}^\vee
\end{array}$$

where α' is a connecting morphism. Now $\varepsilon^\vee \circ \alpha_1 = \iota_{\mathcal{C}}^\vee \circ \varepsilon_1^\vee$ and $(1_{\mathcal{L}^\vee} \otimes \varepsilon_1^\vee) \circ \alpha_1 = -(1_{\mathcal{L}^\vee} \otimes \alpha_1^\vee) \circ \varepsilon_1$, so this commutativity means that $\alpha' \circ \varepsilon_1 = -(1_{\mathcal{L}^\vee} \otimes (\alpha_1 \circ \iota_C)^\vee) \circ \varepsilon_1$. Twisting the commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\chi^{-1}\circ\rho^\vee} & \mathcal{E} & \xrightarrow{\tau\circ\rho} & f^*(\mathcal{A}^\vee)(1) & \tau o & 0 \\
 & & \| & & \big\downarrow{\scriptstyle \widetilde{\pi}} & & \big\downarrow{\scriptstyle f^*(\pi_{\mathcal{A}}^\vee)(1)} & & \\
0 & \to & \overline{\mathcal{E}}^\vee & \xrightarrow{\widetilde{\pi}\circ\chi^{-1}\circ\rho^\vee} & \widetilde{\mathcal{V}} & \xrightarrow{\tau_1\circ\widetilde{\rho_{\mathcal{V}}}} & f^*(\mathcal{A}_1^\vee)(1) & \to & 0
\end{array}$$

by -1 and taking higher direct images, we get the commutative diagram

$$\begin{array}{ccc}
\mathcal{A}^\vee & \xrightarrow{-1_{\mathcal{L}^\vee}\otimes\alpha^\vee} & \mathcal{L}^\vee \otimes \mathcal{C}^\vee \\
\big\downarrow{\scriptstyle \pi_{\mathcal{A}}^\vee} & & \| \\
\mathcal{A}_1^\vee & \xrightarrow{\alpha'} & \mathcal{L}^\vee \otimes \mathcal{C}^\vee
\end{array}$$

for the connecting morphisms. As $\alpha = (1_{\mathcal{L}^\vee} \otimes \pi_{\mathcal{A}}) \circ \alpha_1 \circ \iota_C$, this commutativity means that $\alpha' \circ \pi_{\mathcal{A}}^\vee = -(1_{\mathcal{L}^\vee} \otimes (\alpha_1 \circ \iota_C)^\vee) \circ \pi_{\mathcal{A}}^\vee$. It follows from the second diagram in this section that $[\varepsilon_1 \ \pi_{\mathcal{A}}^\vee]$ is an epimorphism. From both these commutativity relations for α' we therefore get that $\alpha' = -1_{\mathcal{L}^\vee} \otimes (\alpha_1 \circ \iota_C)^\vee$.

We now look at the dual sequence

$$0 \to f^*(\mathcal{A}_1)(-1) \xrightarrow{\widetilde{\rho}_{\mathcal{V}}^\vee \circ \tau_1^\vee} \widetilde{\mathcal{V}}^\vee \xrightarrow{\rho \circ \chi^{-1} \circ \widetilde{\pi}^\vee} \overline{\mathcal{E}} \to 0$$

By Serre duality, the result just proved means that the connecting morphism for this short exact sequence twisted by -1 equals $\alpha_1 \circ \iota_C$. By our results on "Special extensions" it follows that there is a unique isomorphism $\widetilde{\mathcal{V}}^\vee \cong \mathcal{V}$ such that $\widetilde{\rho}_{\mathcal{V}}^\vee \circ \tau_1^\vee$ corresponds to $\mu_{\mathcal{V}}$ and $\rho \circ \chi^{-1} \circ \widetilde{\pi}^\vee$ corresponds to $\rho_{\mathcal{V}}$. Using the commutativity of a diagram above, we also see that then $\chi^{-1} \circ \widetilde{\pi}^\vee$ corresponds to π.

We use this to identify $\widetilde{\mathcal{V}}$ with $\mathcal{V}^\vee$. Then $\tau_1 \circ \widetilde{\rho}_{\mathcal{V}} = \mu_{\mathcal{V}}^\vee$, $\widetilde{\pi} \circ \chi^{-1} \circ \rho^\vee = \rho_{\mathcal{V}}^\vee$ and $\widetilde{\pi} \circ \chi^{-1} = \pi^\vee$. The last equation means that $\widetilde{\pi} = \pi^\vee \circ \chi$. Using that, the second one reduces to $\pi^\vee \circ \rho^\vee = \rho_{\mathcal{V}}^\vee$, which we already knew.

We next do something similar for $\mathcal{E}_1$. Using the diagram defining κ, instead of the one defining $\widetilde{\pi}$, we first get the exact commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_{\mathcal{V}}} & \mathcal{V} & \xrightarrow{\rho_{\mathcal{V}}} & \overline{\mathcal{E}} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \kappa} & & \downarrow{\scriptstyle \lambda} & & \\
0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_1} & \mathcal{E}_1 & \xrightarrow{\rho_1} & \overline{\mathcal{E}}_1 & \to & 0 \\
 & & & & \downarrow{\scriptstyle \sigma \circ \rho_1} & & \downarrow{\scriptstyle \sigma} & & \\
 & & & & f^*(\mathcal{Z}^\vee)(1) & = & f^*(\mathcal{Z}^\vee)(1) & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

Using the exactness of the sequence

$$0 \to \overline{\mathcal{E}}_1^\vee \xrightarrow{\widetilde{\rho}_{\mathcal{V}}^\vee} \mathcal{V} \xrightarrow{\widetilde{\mu}_{\mathcal{V}}^\vee} f^*(\mathcal{A}^\vee)(1) \to 0 \ cr$$

noting that $\tau \circ \rho_{\mathcal{V}} = \tau \circ \rho \circ \chi^{-1} \circ \widetilde{\pi}^\vee = \mu^\vee \circ \widetilde{\pi}^\vee = \widetilde{\mu}_{\mathcal{V}}^\vee$, we get the exact commutative diagram

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \overline{\mathcal{E}}_1^\vee \arrow[r, "\widetilde{\rho}_{\mathcal{V}}^\vee"] \arrow[d, equal] & \mathcal{V} \arrow[r, "\tau\circ\rho_{\mathcal{V}}"] \arrow[d, "\kappa"] & f^*(\mathcal{A}^\vee)(1) \arrow[r] \arrow[d, "f^*(\pi_{\mathcal{A}}^\vee)(1)"] & 0 \\
0 \arrow[r] & \overline{\mathcal{E}}_1^\vee \arrow[r, "\kappa\circ\widetilde{\rho}_{\mathcal{V}}^\vee"] & \mathcal{E}_1 \arrow[r, "\tau_1\circ\rho_1"] \arrow[d, "\sigma\circ\rho_1"] & f^*(\mathcal{A}_1^\vee)(1) \arrow[r] \arrow[d, "f^*(\iota_{\mathcal{A}}^\vee)(1)"] & 0 \\
 & & f^*(\mathcal{Z}^\vee)(1) \arrow[r, equal] \arrow[d] & f^*(\mathcal{Z}^\vee)(1) \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

Using that $\kappa\circ\widetilde{\rho}_{\mathcal{V}}^\vee\circ\tau_1^\vee=\kappa\circ\mu_{\mathcal{V}}=\mu_1$, we now have the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & f^*(\mathcal{A}_1)(-1) \arrow[r, "\mu_1"] \arrow[d, "\tau_1^\vee"] & \mathcal{E}_1 \arrow[r, "\rho_1"] \arrow[d, equal] & \overline{\mathcal{E}}_1 \arrow[r] \arrow[d, "\tau_1"] & 0 \\
0 \arrow[r] & \overline{\mathcal{E}}_1^\vee \arrow[r, "\kappa\circ\widetilde{\rho}_{\mathcal{V}}^\vee"] & \mathcal{E}_1 \arrow[r, "\tau_1\circ\rho_1"] & f^*(\mathcal{A}_1^\vee)(1) \arrow[r] & 0
\end{tikzcd}$$

with exact rows. Twisting and taking higher direct images we now get the commutative diagram

$$\begin{tikzcd}
\mathcal{C}_1 \arrow[r, "\alpha_1"] \arrow[d, "\varepsilon_1"] & \mathcal{L}^\vee\otimes\mathcal{A}_1 \arrow[d, "1_{\mathcal{L}^\vee}\otimes\varepsilon_1^\vee"] \\
\mathcal{A}_1^\vee \arrow[r, "\alpha_1'"] & \mathcal{L}^\vee\otimes\mathcal{C}_1^\vee
\end{tikzcd}$$

where α_1' is a connecting morphism. As $(1_{\mathcal{L}^\vee}\otimes\varepsilon_1^\vee)\circ\alpha_1=-(1_{\mathcal{L}^\vee}\otimes\alpha_1^\vee)\circ\varepsilon_1$, this commutativity means that $\alpha_1'\circ\varepsilon_1=-(1_{\mathcal{L}^\vee}\otimes\alpha_1^\vee)\circ\varepsilon_1$. Using the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & \overline{\mathcal{E}}_1^\vee \arrow[r, "\widetilde{\rho}_{\mathcal{V}}^\vee"] \arrow[d, equal] & \mathcal{V} \arrow[r, "\tau\circ\rho_{\mathcal{V}}"] \arrow[d, "\kappa"] & f^*(\mathcal{A}^\vee)(1) \arrow[r] \arrow[d, "f^*(\pi_{\mathcal{A}}^\vee)(1)"] & 0 \\
0 \arrow[r] & \overline{\mathcal{E}}_1^\vee \arrow[r, "\kappa\circ\widetilde{\rho}_{\mathcal{V}}^\vee"] & \widetilde{\mathcal{V}} \arrow[r, "\tau_1\circ\rho_1"] & f^*(\mathcal{A}_1^\top l)(1) \arrow[r] & 0
\end{tikzcd}$$

we get the commutative diagram

$$\begin{tikzcd}
\mathcal{A}^\vee \arrow[r, "-(1_{\mathcal{L}^\vee}\otimes\alpha_1^\vee)\circ\pi_{\mathcal{A}}^\vee"] \arrow[d, "\pi_{\mathcal{A}}^\vee"] & \mathcal{L}^\vee\otimes\mathcal{C}_1\ \mathit{dual} \arrow[d, equal] \\
\mathcal{A}_1^\vee \arrow[r, "\alpha_1'"] & \mathcal{L}^\vee\otimes\mathcal{C}_1^\vee
\end{tikzcd}$$

for connecting morphisms. So $\alpha_1'\circ\pi_{\mathcal{A}}^\vee=-(1_{\mathcal{L}^\vee}\otimes\alpha_1^\vee)\circ\pi_{\mathcal{A}}^\vee$. As before we get from these two commutativity relations for α_1' that $\alpha_1'=-1_{\mathcal{L}^\vee}\otimes\alpha_1^\vee$.

We now look at the dual sequence

$$0 \to f^*(\mathcal{A}_1)(-1) \xrightarrow{\rho_1^\vee \circ \tau_1^\vee} \mathcal{E}_1^\vee \xrightarrow{\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee} \overline{\mathcal{E}}_1 \to 0$$

By Serre duality, the result just proved means that the connecting morphism for this short exact sequence twisted by -1 equals α_1. It follows that there is a unique isomorphism $\chi_1 : \mathcal{E}_1 \to \mathcal{E}_1^\vee$ making the diagram

$$\begin{array}{ccccccccc} 0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_1} & \mathcal{E}_1 & \xrightarrow{\rho_1} & \overline{\mathcal{E}}_1 & \to & 0 \\ & & \| & & \downarrow{\scriptstyle \chi_1} & & \| & & \\ 0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\rho_1^\vee \circ \tau_1^\vee} & \mathcal{E}_1^\vee & \xrightarrow{\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee} & \overline{\mathcal{E}}_1 & \to & 0 \end{array}$$

commutative.

We have $\widetilde{\kappa}^\vee \circ \mu_\mathcal{V} = \widetilde{\kappa}^\vee \circ \widetilde{\rho}_\mathcal{V}^\vee \circ \tau_1^\vee = \rho_1^\vee \circ \tau_1^\vee$. Furthermore, $\widetilde{\kappa} \circ \kappa = \widetilde{\pi} \circ \pi = \pi^\vee \circ \chi \circ \pi$ is symmetric, hence $\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee \circ \widetilde{\kappa}^\vee = \widetilde{\rho}_\mathcal{V} \circ \widetilde{\kappa} \circ \kappa = \rho_1 \circ \kappa = \lambda \circ \rho_\mathcal{V}$. So the diagram

$$\begin{array}{ccccccccc} 0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\mu_\mathcal{V}} & \mathcal{V} & \xrightarrow{\rho_\mathcal{V}} & \overline{\mathcal{E}} & \to & 0 \\ & & \| & & \downarrow{\scriptstyle \widetilde{\kappa}^\vee} & & \downarrow{\scriptstyle \lambda} & & \\ 0 & \to & f^*(\mathcal{A}_1)(-1) & \xrightarrow{\rho_1^\vee \circ \tau_1^\vee} & \mathcal{E}_1^\vee & \xrightarrow{\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee} & \overline{\mathcal{E}}_1 & \to & 0 \end{array}$$

is commutative. Because of the uniqueness of κ it follows that $\chi_1^{-1} \circ \widetilde{\kappa}^\vee = \kappa$, i.e., $\widetilde{\kappa} = \kappa^\vee \circ \chi_1^\vee$. We then get $\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee \circ \chi_1^\vee = \widetilde{\rho}_\mathcal{V} \circ \widetilde{\kappa} = \rho_1$. We also have $\chi_1^\vee \circ \mu_1 = \chi_1^\vee \circ \kappa \circ \mu_\mathcal{V} = \chi_1^\vee \circ \kappa \circ \widetilde{\rho}_\mathcal{V}^\vee \circ \tau_1^\vee = (\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee \circ \chi_1)^\vee \circ \tau_1^\vee$. As $\widetilde{\rho}_\mathcal{V} \circ \kappa^\vee \circ \chi_1 = \rho_1$ by the above, we get $\chi_1^\vee \circ \mu_1 = \rho_1^\vee \circ \tau_1^\vee$. This shows that the diagram defining χ_1 remains commutative if we replace χ_1 by $\chi_1^\vee$. As χ_1 is uniquely determined, this means that $\chi_1^\vee = \chi_1$. So $(\mathcal{E}_1, \chi_1)$ is a symmetric bilinear space. Also note that we can now write the identity $\widetilde{\kappa} = \kappa^\vee \circ \chi_1^\vee$ as $\widetilde{\kappa} = \kappa^\vee \circ \chi_1$.

It is now easy to check that $((\mathcal{E}_1, \chi_1), (\mathcal{A}_1, \mu_1))$ is an NN-pair extending $((\mathcal{E}, \chi), (\mathcal{A}, \mu))$ in the way we wanted.

Appendix on Special Extensions

Let $\mathcal{X}$ be a vector bundle over Y such that $R^1 f_*(\mathcal{X}(-1)) = 0$. Let $\mathcal{M}$ be a vector bundle over X. Then $\mathrm{Hom}_Y(\mathcal{X}, f^*(\mathcal{M})(-1)) = 0$ and there is a natural isomorphism $\mathrm{Ext}_Y(\mathcal{X}, f^*(\mathcal{M})(-1)) \cong \mathrm{Hom}_X(f_*(\mathcal{X}(-1)), \mathcal{L}^\vee \otimes \mathcal{M})$. This isomorphism maps the class of a short exact sequence

$$0 \to f^*(\mathcal{M})(-1) \to \mathcal{Y} \to \mathcal{X} \to 0$$

to the connecting morphism

$$f_*(\mathcal{X}(-1)) \to R^1 f_*(f^*(\mathcal{M})(-2)) = \mathcal{L}^\vee \otimes \mathcal{M}$$

for the short exact sequence twisted by -1.
There are various ways to prove this. One way is to use the natural short exact sequence

$$0 \to \omega(1) \otimes f^*(f_*(\mathcal{X}(-1))) \to f^*(f_*(\mathcal{X})) \to \mathcal{X} \to 0$$

given by Proposition 2 in Section 1.1 and the corresponding long exact sequence of higher Ext groups.
Note that $\mathrm{Hom}_Y(\mathcal{X}, f^*(\mathcal{M})(-1)) = 0$ implies that a short exact sequence as above, representing a given element in $\mathrm{Ext}_Y(\mathcal{X}, f^*(\mathcal{M})(-1))$, is determined up to a unique isomorphism of $\mathcal{Y}$.

Let $\mathcal{X}_1$ and $\mathcal{M}_1$ be another pair satisfying the hypotheses above. Let the short exact sequence

$$0 \to f^*(\mathcal{M})(-1) \to \mathcal{Y} \to \mathcal{X} \to 0$$

correspond to $\alpha : f_*(\mathcal{X}(-1)) \to \mathcal{L}^\vee \otimes \mathcal{M}$ and let

$$0 \to f^*(\mathcal{M}_1)(-1) \to \mathcal{Y}_1 \to \mathcal{X}_1 \to 0$$

correspond to $\alpha_1 : f_*(\mathcal{X}_1(-1)) \to \mathcal{L}^\vee \otimes \mathcal{M}_1$. Let $\xi : \mathcal{X} \to \mathcal{X}_1$ and $\mu : \mathcal{M} \to \mathcal{M}_1$ be morphisms making the diagram

$$\begin{array}{ccc}
f_*(\mathcal{X}(-1)) & \xrightarrow{\alpha} & \mathcal{L}^\vee \otimes \mathcal{M} \\
\Big\downarrow{\scriptstyle f_*(\xi(-1))} & & \Big\downarrow{\scriptstyle 1\otimes\mu} \\
f_*(\mathcal{X}_1(-1)) & \xrightarrow{\alpha_1} & \mathcal{L}^\vee \otimes \mathcal{M}_1
\end{array}$$

commutative. Then there is a unique morphism $\eta : \mathcal{Y} \to \mathcal{Y}_1$ making the diagram

$$\begin{array}{ccccccccc}
0 & \to & f^*(\mathcal{M})(-1) & \to & \mathcal{Y} & \to & \mathcal{X} & \to & 0 \\
 & & \Big\downarrow{\scriptstyle f^*(\mu)(-1)} & & \Big\downarrow{\scriptstyle \eta} & & \Big\downarrow{\scriptstyle \xi} & & \\
0 & \to & f^*(\mathcal{M}_1)(-1) & \to & \mathcal{Y}_1 & \to & \mathcal{X}_1 & \to & 0
\end{array}$$

commutative.
Indeed, the uniqueness follows from the fact that $\mathrm{Hom}_Y(\mathcal{X}, f^*(\mathcal{M}_1)(-1)) = 0$. The existence follows from the following commutative diagram.

$$\begin{array}{ccccccccc}
0 & \to & f^*(\mathcal{M})(-1) & \to & \mathcal{Y} & \to & \mathcal{X} & \to & 0 \\
 & & \Big\downarrow{\scriptstyle f^*(\mu)(-1)} & & \downarrow & & \| & & \\
0 & \to & f^*(\mathcal{M}_1)(-1) & \to & \mathcal{Z} & \to & \mathcal{X} & \to & 0 \\
 & & \| & & \Big\downarrow{\scriptstyle \cong} & & \| & & \\
0 & \to & f^*(\mathcal{M}_1)(-1) & \to & \mathcal{Z}_1 & \to & \mathcal{X} & \to & 0 \\
 & & \| & & \downarrow & & \Big\downarrow{\scriptstyle \xi} & & \\
0 & \to & f^*(\mathcal{M}_1)(-1) & \to & \mathcal{Y}_1 & \to & \mathcal{X}_1 & \to & 0
\end{array}$$

Here the top part is gotten by a push-out and the bottom one by pull-back. The middle part comes from the fact that both short exact sequences have the same image in $\mathrm{Hom}_X(f_*(\mathcal{X}(-1)), \mathcal{L}^\vee \otimes \mathcal{M}_1)$.

SECTION 2.1

In the affine case, Ranicki has defined the Witt group of formations and proved that it is isomorphic to his group L^1. (Cf [R].) Here we shall extend his definition to our case.

We shall use the duality functor $\top$ on vector bundles on X given by $\mathcal{E}^\top = \mathcal{L}^\vee \otimes \mathcal{E}^\vee$. But, in fact, what we do makes sense in any exact category with duality.

A (non-singular) formation is a triple $((\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{C},\gamma))$, where $(\mathcal{F},\varphi)$ is a symmetric bilinear space and $\alpha : \mathcal{A} \to \mathcal{F}$ and $\gamma : \mathcal{C} \to \mathcal{F}$ are lagrangians of $(\mathcal{F},\varphi)$. Sometimes we simply say that $(\mathcal{F},\alpha,\gamma)$ is a formation.
There is an obvious notion of isomorphisms of formations. Furthermore, we can define the direct sum of two formations in an obvious way. It follows that we have the Grothendieck group of isomorphism classes of formations.

For any vector bundle $\mathcal{Z}$ we have the formation $(\mathrm{H}_\top(\mathcal{Z}),(\mathcal{Z},\begin{bmatrix}1\\0\end{bmatrix}),(\mathcal{Z}^\top,\begin{bmatrix}0\\1\end{bmatrix}))$. Ranicki uses direct sums of formations with these special formations to define when the formations are stably isomorphic. As short exact sequences are not necessarily split in our case, we have to use something more general than direct sums.

Let $((\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{C},\gamma))$ be a formation and let $\mathcal{Z}$ be a vector bundle. We shall define what it means that a formation $((\mathcal{F}_1,\varphi_1),(\mathcal{A}_1,\alpha_1),(\mathcal{C}_1,\gamma_1))$ is an extension of $((\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{C},\gamma))$ by $\mathcal{Z}$.
The first condition is that $(\mathcal{F}_1,\varphi_1)$ is an extension of $(\mathcal{F},\varphi)$ by $\mathcal{Z}$. This means that there is a commutative diagram

$$\begin{array}{ccccccccc}
 & & & & 0 & & 0 & & \\
 & & & & \downarrow & & \downarrow & & \\
0 & \to & \mathcal{Z} & \xrightarrow{\iota} & \mathcal{V} & \xrightarrow{\pi} & \mathcal{F} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle \kappa} & & \downarrow{\scriptstyle \pi^\top\circ\varphi} & & \\
0 & \to & \mathcal{Z} & \xrightarrow{\kappa\circ\iota} & \mathcal{F}_1 & \xrightarrow{\kappa^\top\circ\varphi_1} & \mathcal{V}^\top & \to & 0 \\
 & & & & \downarrow{\scriptstyle \iota^\top\circ\kappa^\top\circ\varphi_1} & & \downarrow{\scriptstyle \iota^\top} & & \\
 & & & & \mathcal{Z}^\top & = & \mathcal{Z}^\top & & \\
 & & & & \downarrow & & \downarrow & & \\
 & & & & 0 & & 0 & &
\end{array}$$

with exact rows and columns. In a relaxed language this says that $\mathcal{Z}$ is a totally isotropic subbundle of $\mathcal{F}_1$ with the orthogonal subbundle $\mathcal{V}$ and that $\mathcal{F}$ is the quotient of $\mathcal{V}$ by $\mathcal{Z}$ with the induced form.

The second condition is that $\mathcal{A}_1$ is an extension of $\mathcal{A}$ by $\mathcal{Z}$ and $\mathcal{C}_1$ is an extension of $\mathcal{Z}^\top$ by $\mathcal{C}$. So we have short exact sequences

$$0 \to \mathcal{Z} \xrightarrow{\iota_{\mathcal{A}}} \mathcal{A}_1 \xrightarrow{\pi_{\mathcal{A}}} \mathcal{A} \to 0$$

and

$$0 \to \mathcal{C} \xrightarrow{\iota_{\mathcal{C}}} \mathcal{C}_1 \xrightarrow{\pi_{\mathcal{C}}} \mathcal{Z}^\top \to 0$$

Finally, these extensions are to be compatible in the following sense. The embedding $\mathcal{A}_1 \to \mathcal{F}_1$ factors as $\alpha_1 = \kappa \circ \underline{\alpha}$ with a morphism $\underline{\alpha} : \mathcal{A}_1 \to \mathcal{V}$ such that the diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{Z} \arrow[r, "\iota_{\mathcal{A}}"] \arrow[d, equal] & \mathcal{A}_1 \arrow[r, "\pi_{\mathcal{A}}"] \arrow[d, "\underline{\alpha}"] & \mathcal{A} \arrow[r] \arrow[d, "\alpha"] & 0 \\
0 \arrow[r] & \mathcal{Z} \arrow[r, "\iota"] & \mathcal{V} \arrow[r, "\pi"] & \mathcal{F} \arrow[r] & 0
\end{tikzcd}$$

is commutative. Also, the embedding $\mathcal{C} \to \mathcal{F}$ factors as $\gamma = \pi \circ \underline{\gamma}$ with a morphism $\underline{\gamma} : \mathcal{C} \to \mathcal{V}$ such that the diagram

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C} \arrow[r, "\iota_{\mathcal{C}}"] \arrow[d, "\underline{\gamma}"] & \mathcal{C}_1 \arrow[r, "\pi_{\mathcal{C}}"] \arrow[d, "\gamma_1"] & \mathcal{Z}^\top \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{V} \arrow[r, "\kappa"] & \mathcal{F}_1 \arrow[r, "\iota^\top \circ \kappa^\top \circ \varphi_1"] & \mathcal{Z}^\top \arrow[r] & 0
\end{tikzcd}$$

is commutative.

As we know the cokernels of α and γ_1, we can, if the conditions above hold, extend the last two diagrams to the commutative diagrams

$$\begin{tikzcd}
 & & 0 \arrow[d] & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{Z} \arrow[r, "\iota_{\mathcal{A}}"] \arrow[d, equal] & \mathcal{A}_1 \arrow[r, "\pi_{\mathcal{A}}"] \arrow[d, "\underline{\alpha}"] & \mathcal{A} \arrow[r] \arrow[d, "\alpha"] & 0 \\
0 \arrow[r] & \mathcal{Z} \arrow[r, "\iota"] & \mathcal{V} \arrow[r, "\pi"] \arrow[d, "\alpha^\top \circ \varphi \circ \pi"] & \mathcal{F} \arrow[r] \arrow[d, "\alpha^\top \circ \varphi"] & 0 \\
 & & \mathcal{A}^\top \arrow[r, equal] \arrow[d] & \mathcal{A}^\top \arrow[d] & \\
 & & 0 & 0 &
\end{tikzcd}$$

and

$$\begin{tikzcd}
 & 0 \arrow[d] & 0 \arrow[d] & & \\
0 \arrow[r] & \mathcal{C} \arrow[r, "\iota_{\mathcal{C}}"] \arrow[d, "\gamma"] & \mathcal{C}_1 \arrow[r, "\pi_{\mathcal{C}}"] \arrow[d, "\gamma_1"] & \mathcal{Z}^\top \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{V} \arrow[r, "\kappa"] \arrow[d, "\gamma_1^\top\circ\varphi_1\circ\kappa"] & \mathcal{F}_1 \arrow[r, "\iota^\top\circ\kappa^\top\circ\varphi_1"] \arrow[d, "\gamma_1^\top\circ\varphi_1"] & \mathcal{Z}^\top \arrow[r] & 0 \\
 & \mathcal{C}_1^\top \arrow[r, equal] \arrow[d] & \mathcal{C}_1^\top \arrow[d] & & \\
 & 0 & 0 & &
\end{tikzcd}$$

with exact rows and columns. It then follows that we also have the commutative diagrams

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{A}_1 \arrow[r, "\alpha_1"] & \mathcal{F}_1 \arrow[r, "\alpha_1^\top\circ\varphi_1"] & \mathcal{A}_1^\top \arrow[r] & 0 \\
0 \arrow[r] & \mathcal{A}_1 \arrow[r, "\alpha"] \arrow[u, equal] \arrow[d, "\pi_{\mathcal{A}}"] & \mathcal{V} \arrow[r, "\alpha^\top\circ\varphi\circ\pi"] \arrow[u, "\kappa"] \arrow[d, "\pi"] & \mathcal{A}^\top \arrow[r] \arrow[u, "\pi_{\mathcal{A}}^\top"] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{A} \arrow[r, "\alpha"] & \mathcal{F} \arrow[r, "\alpha^\top\circ\varphi"] & \mathcal{A}^\top \arrow[r] & 0
\end{tikzcd}$$

and

$$\begin{tikzcd}
0 \arrow[r] & \mathcal{C} \arrow[r, "\gamma"] & \mathcal{F} \arrow[r, "\gamma^\top\circ\varphi"] & \mathcal{C}^\top \arrow[r] & 0 \\
0 \arrow[r] & \mathcal{C} \arrow[r, "\gamma"] \arrow[u, equal] \arrow[d, "\iota_{\mathcal{C}}"] & \mathcal{V} \arrow[r, "\gamma_1^\top\circ\varphi_1\circ\kappa"] \arrow[u, "\pi"] \arrow[d, "\kappa"] & \mathcal{C}_1^\top \arrow[r] \arrow[u, "\iota_{\mathcal{C}}^\top"] \arrow[d, equal] & 0 \\
0 \arrow[r] & \mathcal{C}_1 \arrow[r, "\gamma_1"] & \mathcal{F}_1 \arrow[r, "\gamma_1^\top\circ\varphi_1"] & \mathcal{C}_1^\top \arrow[r] & 0
\end{tikzcd}$$

with exact rows.

Let $(\mathcal{F},\varphi)$ be a symmetric bilinear space and let $\gamma:\mathcal{C}\to\mathcal{F}$ be a lagrangian of $(\mathcal{F},\varphi)$. We then might say that $((\mathcal{F},\varphi),(\mathcal{C},\gamma))$ is a metabolic pair. Now let $\mathcal{C}_1$ be a vector bundle and let $\iota_C:\mathcal{C}\to\mathcal{C}_1$ be a morphism. Then there is, by [A], a metabolic pair $((\mathcal{F}_1,\varphi_1),(\mathcal{C}_1,\gamma_1))$, uniquely determined up to an isomorphism by the conditions that there is a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & \mathcal{C} & \xrightarrow{\gamma} & \mathcal{F} & \xrightarrow{\gamma^\top\circ\varphi} & \mathcal{C}^\top & \to & 0 \\
 & & \| & & \uparrow{\scriptstyle\pi} & & \uparrow{\scriptstyle\iota_C^\top} & & \\
0 & \to & \mathcal{C} & \xrightarrow{\gamma} & \mathcal{V} & \xrightarrow{\gamma_1^\top\circ\varphi_1\circ\kappa} & \mathcal{C}_1^\top & \to & 0 \\
 & & \downarrow{\scriptstyle\iota_C} & & \downarrow{\scriptstyle\kappa} & & \| & & \\
0 & \to & \mathcal{C}_1 & \xrightarrow{\gamma_1} & \mathcal{F}_1 & \xrightarrow{\gamma_1^\top\circ\varphi_1} & \mathcal{C}_1^\top & \to & 0
\end{array}$$

with exact rows and that $\kappa^\top \circ \varphi_1 \circ \kappa = \pi^\top \circ \varphi \circ \pi$.
Now assume that we have a short exact sequence

$$0 \to \mathcal{C} \xrightarrow{\iota_C} \mathcal{C}_1 \xrightarrow{\pi_C} \mathcal{Z}^\top \to 0$$

Then we can compute the kernels and cokernels of the vertical morphisms in the double diagram above. It easily follows that $(\mathcal{F}_1, \varphi_1)$ is an extension of $(\mathcal{F}, \varphi)$ by $\mathcal{Z}$ as in the definition above.
Now assume that $\alpha : \mathcal{A} \to \mathcal{F}$ is also a lagrangian of $(\mathcal{F}, \varphi)$. Taking the "inverse image" in $\mathcal{V}$ of the subbundle $\mathcal{A}$ of $\mathcal{F}$ we get a commutative diagram

$$\begin{array}{ccccccccc}
0 & \to & \mathcal{Z} & \xrightarrow{\iota_{\mathcal{A}}} & \mathcal{A}_1 & \xrightarrow{\pi_{\mathcal{A}}} & \mathcal{A} & \to & 0 \\
 & & \| & & \downarrow{\scriptstyle\underline{\alpha}} & & \downarrow{\scriptstyle\alpha} & & \\
0 & \to & \mathcal{Z} & \xrightarrow{\iota} & \mathcal{V} & \xrightarrow{\pi} & \mathcal{F} & \to & 0
\end{array}$$

with exact rows. Letting $\alpha_1 = \kappa \circ \underline{\alpha}$, one then checks that $\alpha_1 : \mathcal{A}_1 \to \mathcal{F}_1$ is a lagrangian of $(\mathcal{F}_1, \varphi_1)$. It then easily follows that by this we have constructed an extension $((\mathcal{F}_1, \varphi_1), (\mathcal{A}_1, \alpha_1), (\mathcal{C}_1, \gamma_1))$ of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}$.
We conclude that there is a natural bijective correspondence between isomorphism classes of extensions of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}$ and isomorphism classes of extensions of $\mathcal{Z}^\top$ by $\mathcal{C}$. Of course, the latter correspond to isomorphism classes of extensions of $\mathcal{C}^\top$ by $\mathcal{Z}$.
This makes it rather easy to work with extensions of formations. For example, if $\mathcal{C}_1$ is the trivial extension of $\mathcal{Z}^\top$ by $\mathcal{C}$, then $((\mathcal{F}_1, \varphi_1), (\mathcal{A}_1, \alpha_1), (\mathcal{C}_1, \gamma_1))$ is the trivial extension of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}$, i.e., the direct sum of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ and $(\mathrm{H}_\top(\mathcal{Z}), (\mathcal{Z}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{Z}^\top, \begin{bmatrix} 0 \\ 1 \end{bmatrix}))$. In particular, we get nothing new in the affine case.
Using the concept of the direct sum of two extensions of $\mathcal{C}^\top$, we get the following lemma as another application.

Lemma 1: Let $((\mathcal{F}_1, \varphi_1), (\mathcal{A}_1, \alpha_1), (\mathcal{C}_1, \gamma_1))$ be an extension of of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}_1$ and let $((\mathcal{F}_2, \varphi_2), (\mathcal{A}_2, \alpha_2), (\mathcal{C}_2, \gamma_2))$ be an extension of of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}_2$. Then there is an extension

$((\mathcal{F}_3, \varphi_3), (\mathcal{A}_3, \alpha_3), (\mathcal{C}_3, \gamma_3))$ of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}_1 \oplus \mathcal{Z}_2$ such that the original extension of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ are intermediate extensions in the natural way.

We shall say that two formations are stably isomorphic if they have a common extension. From the lemma it follows that this induces an equivalence relation on the set of isomorphism classes of formations. This equivalence relation is clearly compatible with direct sums.
By a remark above this coincides with Ranicki's definition in the affine case.

We now say, as Ranicki, that two formations $(\mathcal{F}_1, \alpha_1, \gamma_1)$ and $(\mathcal{F}_2, \alpha_2, \gamma_2)$ are equivalent if there is a space $\mathcal{M}_1$ with lagrangians u_1, v_1 and w_1 and a space $\mathcal{M}_2$ with lagrangians u_2, v_2 and w_2 such that the direct sum

$$(\mathcal{F}_1, \alpha_1, \gamma_1) \oplus (\mathcal{M}_1, u_1, v_1) \oplus (\mathcal{M}_1, v_1, w_1) \oplus (\mathcal{M}_2, u_2, w_2)$$

is stably isomorphic to the direct sum

$$(\mathcal{F}_2, \alpha_2, \gamma_2) \oplus (\mathcal{M}_2, u_2, v_2) \oplus (\mathcal{M}_2, v_2, w_2) \oplus (\mathcal{M}_1, u_1, w_1)$$

It is easy to check that this is an equivalence relation on formations. The direct sum induces a group structure on the set of equivalence classes. (We shall see, in a moment, how additive inverses are found.) The resulting group is called the Witt group of formations and is denoted $M(X)$ or, if we want to stress the duality functor used, $M_\top(X)$.
In the affine case Ranicki shows that $M(X)$ is isomorphic to $L^1(X)$, so we might as well have used the notation $L^1(X)$ in our case.

An equivalent way to define $M(X)$ is to consider first the Grothendieck group of isomorphism classes of formations and then to consider $M(X)$ as the qoutiont group gotten by demanding two formations to have the same class if one is an extension of the other and that the direct sum $(\mathcal{F}, \alpha, \beta) \oplus (\mathcal{F}, \beta, \gamma)$ has the same class as $(\mathcal{F}, \alpha, \gamma)$.
In this formulation it is clear that the class of $(\mathcal{F}, \alpha, \alpha)$ is trivial and then that the class of $(\mathcal{F}, \gamma, \alpha)$ is the inverse of the class of $(\mathcal{F}, \alpha, \gamma)$.

Section 2.2

A formation is said to be split if it is isomorphic to a formation of the type

$$\left((\mathcal{A} \oplus \mathcal{A}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{C}, \begin{bmatrix} \alpha \\ \varepsilon \end{bmatrix})\right)$$

In this section we study split formation and define a Witt group of these.

A split-formation (over X) is a quadruple $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$, where $\mathcal{A}$ and $\mathcal{C}$ are vector bundles over X and $\alpha:\mathcal{C}\to\mathcal{A}$ and $\varepsilon:\mathcal{C}\to\mathcal{A}^\top$ are morphisms such that $\begin{bmatrix}\alpha\\ \varepsilon\end{bmatrix}$ is an embedding of $\mathcal{C}$ in $\mathcal{A}\oplus\mathcal{A}^\top$ as a Lagrangian of the hyperbolic $\top$-symmetric bilinear space $\mathrm{H}_\top(\mathcal{A})$. This means that

$$0 \;\to\; \mathcal{C} \;\xrightarrow{\begin{bmatrix}\alpha\\ \varepsilon\end{bmatrix}}\; \mathcal{A}\oplus\mathcal{A}^\top \;\xrightarrow{\begin{bmatrix}\varepsilon^\top & \alpha^\top\end{bmatrix}}\; \mathcal{C}^\top \;\to\; 0$$

is a short exact sequence.
There is an obvious notion of isomorphisms of split-formations. Furthermore, we can define the direct sum of two formations in an obvious way. It follows that we have the Grothendieck group of isomorphism classes of split-formations.

Let $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ and $(\mathcal{A}_1,\mathcal{C}_1,\alpha_1,\varepsilon_1)$ be split-formations. We say that $(\mathcal{A}_1,\mathcal{C}_1,\alpha_1,\varepsilon_1)$ is an extension of $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ (by $\mathcal{Z}$) and that $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is a quotient of $(\mathcal{A}_1,\mathcal{C}_1,\alpha_1,\varepsilon_1)$ if there is a vector bundle $\mathcal{Z}$ over X and short exact sequences

$$0 \;\to\; \mathcal{Z} \;\xrightarrow{\iota_\mathcal{A}}\; \mathcal{A}_1 \;\xrightarrow{\pi_\mathcal{A}}\; \mathcal{A} \;\to\; 0$$

and

$$0 \;\to\; \mathcal{C} \;\xrightarrow{\iota_\mathcal{C}}\; \mathcal{C}_1 \;\xrightarrow{\pi_\mathcal{C}}\; \mathcal{Z}^\top \;\to\; 0$$

such that the diagrams

$$\begin{tikzcd} 0 \arrow[r] & \mathcal{C} \arrow[r, "\iota_\mathcal{C}"] \arrow[d, "\varepsilon"] & \mathcal{C}_1 \arrow[r, "\pi_\mathcal{C}"] \arrow[d, "\varepsilon_1"] & \mathcal{Z}^\top \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & \mathcal{A}^\top \arrow[r, "\pi_\mathcal{A}^\top"] & \mathcal{A}_1^\top \arrow[r, "\iota_\mathcal{A}^\top"] & \mathcal{Z}^\top \arrow[r] & 0 \end{tikzcd}$$

and

$$\begin{tikzcd} \mathcal{C} \arrow[r, "\alpha"] \arrow[d, "\iota_\mathcal{C}"] & \mathcal{A} \\ \mathcal{C}_1 \arrow[r, "\alpha_1"] & \mathcal{A}_1 \arrow[u, "\pi_\mathcal{A}"'] \end{tikzcd}$$

are commutative.

We say that a split-formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is elementary if α is an isomorphism. Indeed, we then may (up to an isomorphism of split-formations) assume that $\mathcal{C}=\mathcal{A}$ and $\alpha=1_\mathcal{A}$. The fact that $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is a split-formation then simply means that $\varepsilon^\top=-\varepsilon$, i.e., ε is $\top$-skew-symmetric. It follows that the "elementary" automorphism $\begin{bmatrix}1 & 0\\ \varepsilon & 1\end{bmatrix}$ of $\mathrm{H}_\top(\mathcal{A})$ takes the canonical lagrangian $\begin{bmatrix}1\\ 0\end{bmatrix}:\mathcal{A}\to\mathcal{A}\oplus\mathcal{A}^\top$ to $\begin{bmatrix}\alpha\\ \varepsilon\end{bmatrix}$.
We say that a split-formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is metabolic if it has an elementary extension. It is clear that a direct sum of metabolic split-formations is metabolic.

We define the Witt group $M^{\text{spl}}(X)$ of split-formations as the Grothendieck group of split-formations modulo the subgroup generated by metabolic split-formations.

Proposition 1: If the split-formation $(\mathcal{A}_1, \mathcal{C}_1, \alpha_1, \varepsilon_1)$ is an extension of the split-formation $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$ then the direct sum $(\mathcal{A}_1, \mathcal{C}_1, \alpha_1, \varepsilon_1) \oplus (\mathcal{A}, \mathcal{C}, \alpha, -\varepsilon)$ is metabolic.

Proof: We shall use the notations used in the definition to describe $(\mathcal{A}_1, \mathcal{C}_1, \alpha_1, \varepsilon_1)$ as an extension of $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$. We let $\widetilde{\mathcal{Z}} = \mathcal{A}^\top \oplus \mathcal{C}_1$ and let $\widetilde{\mathcal{A}} = \mathcal{A}_1 \oplus \mathcal{A} \oplus \mathcal{A}^\top \oplus \mathcal{C}_1$ be the direct sum of $\mathcal{A}_1 \oplus \mathcal{A}$ and $\widetilde{\mathcal{Z}}$. So $\widetilde{\iota}_{\mathcal{A}} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\widetilde{\pi}_{\mathcal{A}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$. We also let $\widetilde{\mathcal{C}} = \widetilde{\mathcal{A}}$ and $\widetilde{\alpha} = 1$. We now let

$$\widetilde{\varepsilon} = \begin{bmatrix} 0 & 0 & -\pi_{\mathcal{A}}^\top & \varepsilon_1 \\ 0 & 0 & -1 & 0 \\ \pi_{\mathcal{A}} & 1 & 0 & -\pi_{\mathcal{A}} \circ \alpha_1 \\ -\varepsilon_1^\top & 0 & \alpha_1^\top \circ \pi_{\mathcal{A}}^\top & \varepsilon_1^\top \circ \alpha_1 \end{bmatrix}$$

Then $\widetilde{\varepsilon}$ is clearly $\top$-skew-symmetric. (Recall that $\varepsilon_1^\top \circ \alpha_1 + \alpha_1^\top \circ \varepsilon_1 = 0$.) Also

$$\widetilde{\iota}_{\mathcal{C}} = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha \\ 0 & \varepsilon \\ 1 & \iota_{\mathcal{C}} \end{bmatrix}$$

and

$$\widetilde{\pi}_{\mathcal{C}} = \widetilde{\iota}_{\mathcal{A}}^\top \circ \widetilde{\varepsilon} = \begin{bmatrix} \pi_{\mathcal{A}} & 1 & 0 & -\pi_{\mathcal{A}} \circ \alpha_1 \\ -\varepsilon_1^\top & 0 & \alpha_1^\top \circ \pi_{\mathcal{A}}^\top & \varepsilon_1^\top \circ \alpha_1 \end{bmatrix}$$

Easy computations then show that

$$\widetilde{\varepsilon} \circ \widetilde{\iota}_{\mathcal{C}} = \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & -\varepsilon \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \widetilde{\pi}_{\mathcal{A}}^\top \circ \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & -\varepsilon \end{bmatrix}$$

and, clearly, $\widetilde{\pi}_{\mathcal{A}} \circ \widetilde{\alpha} \circ \widetilde{\iota}_\gamma = \widetilde{\pi}_{\mathcal{A}} \circ \widetilde{\iota}_\gamma = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha \end{bmatrix}$. To show that $(\widetilde{\mathcal{A}}, \widetilde{\mathcal{C}}, \widetilde{\alpha}, \widetilde{\varepsilon})$ is an extension of $(\mathcal{A}_1, \mathcal{C}_1, \alpha_1, \varepsilon_1) \oplus (\mathcal{A}, \mathcal{C}, \alpha, -\varepsilon)$ there remains only to show that the sequence

$$0 \to \mathcal{C}_1 \oplus \mathcal{C} \xrightarrow{\widetilde{\iota}_{\mathcal{C}}} \widetilde{\mathcal{C}} \xrightarrow{\widetilde{\pi}_{\mathcal{C}}} \widetilde{\mathcal{Z}}^\top \to 0$$

is exact. From the definition of $\widetilde{\pi}_{\mathcal{C}}$ and the fact that $\widetilde{\varepsilon} \circ \widetilde{\iota}_{\mathcal{C}}$ equals $\widetilde{\pi}_{\mathcal{A}}^\top \circ \begin{bmatrix} \varepsilon_1 & 0 \\ 0 & -\varepsilon \end{bmatrix}$

it follows that it is a zero sequence. We now use the commutative diagram

$$\begin{tikzcd}
 & 0 \arrow[d] & & 0 \arrow[d] & & 0 \arrow[d] & \\
0 \arrow[r] & \mathcal{C}_1 \arrow[rr, "\begin{bmatrix}\alpha_1\\ 1\end{bmatrix}"] \arrow[d, "\text{nat.incl.}"] & & \mathcal{A}_1 \oplus \mathcal{C}_1 \arrow[rr, "{[1\ \ -\alpha_1]}"] \arrow[d, "\text{nat.incl.}"] & & \mathcal{A}_1 \arrow[r, phantom, "\tau o"] \arrow[d, "\begin{bmatrix}\pi_{\mathcal{A}}\\ -\varepsilon_1^{\top}\end{bmatrix}"] & 0 \\
0 \arrow[r] & \mathcal{C}_1 \oplus \mathcal{C} \arrow[rr, "\widetilde{\iota_{\mathcal{C}}}"] \arrow[d, "\text{nat.proj.}"] & & \widetilde{\mathcal{C}} \arrow[rr, "\widetilde{\pi_{\mathcal{C}}}"] \arrow[d, "\text{nat.proj.}"] & & \widetilde{\mathcal{Z}}^{\top} \arrow[r] \arrow[d, "{[\varepsilon^{\top}\ \ -\iota_{\mathcal{C}}^{\top}]}"] & 0 \\
0 \arrow[r] & \mathcal{C} \arrow[rr, "\begin{bmatrix}\alpha\\ \varepsilon\end{bmatrix}"] \arrow[d] & & \mathcal{A} \oplus \mathcal{A}^{\top} \arrow[rr, "{[\varepsilon^{\top}\ \ \alpha^{\top}]}"] \arrow[d] & & \mathcal{C}\delta ua \arrow[r] \arrow[d] & 0 \\
 & 0 & & 0 & & 0 &
\end{tikzcd}$$

Obviously, the left hand column and the middle column are exact. Using the dual of the commutative diagram connecting ε and ε_1, one sees that the right hand column is exact. The top row is clearly exact and the bottom one is exact by the definition of a split-formation. As the middle row is a zero sequence, it follows that it is exact too.

As any split-formation is trivially an extension of itself, we have the following corollary.

Corollary 2: For any split-formation $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$ the direct sum $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon) \oplus (\mathcal{A}, \mathcal{C}, \alpha, -\varepsilon)$ is metabolic.

It follows that any element in $M^{\mathrm{spl}}(X)$ is represented by a split-formation. It also follows that a split-formation $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$ has trivial class in $M^{\mathrm{spl}}(X)$ if and only if there is a metabolic split-formation $(\mathcal{A}_0, \mathcal{C}_0, \alpha_0, \varepsilon_0)$ such that $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon) \oplus (\mathcal{A}_0, \mathcal{C}_0, \alpha_0, \varepsilon_0)$ is metabolic. (In fact, it can be shown that $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$ is metabolic itself.)

Section 2.3

In this section we prove that there is a natural isomorphism from the Witt group of split-formations to the Witt group of formations. For the proof we need that 2 is invertible.

A split-formation $(\mathcal{A}, \mathcal{C}, \gamma_+, \gamma_-)$ gives rise to the formation

$$\Big((\mathcal{A} \oplus \mathcal{A}^{\top}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{C}, \begin{bmatrix} \gamma_+ \\ \gamma_- \end{bmatrix}) \Big)$$

Going from split-formations to formations in this way clearly induces a morphism of Grothendieck groups of isomorphism classes. It is also trivial to check that extensions of split-formations go to extension of formations. If a split-formation is elementary then we may assume that it is of the type $(\mathcal{A}, \mathcal{A}, 1, \gamma_-)$ with a skew-symmetric $\gamma_- : \mathcal{A} \to \mathcal{A}^\top$. The class of the corresponding formation

$$\Big((\mathcal{A} \oplus \mathcal{A}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ \gamma_- \end{bmatrix})\Big)$$

is then the difference of the classes of the formations

$$\Big((\mathcal{A} \oplus \mathcal{A}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{A}^\top, \begin{bmatrix} 0 \\ 1 \end{bmatrix})\Big)$$

and

$$\Big((\mathcal{A} \oplus \mathcal{A}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{A}, \begin{bmatrix} 1 \\ \gamma_- \end{bmatrix}), (\mathcal{A}^\top, \begin{bmatrix} 0 \\ 1 \end{bmatrix})\Big)$$

But the automorphism $\begin{bmatrix} 1 & 0 \\ \gamma_- & 1 \end{bmatrix}$ of $(\mathcal{A} \oplus \mathcal{A}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix})$ induces an isomorphism from the former formation to the latter, so the difference of the classes is 0. It follows that we get a natural morphism $M^{\mathrm{spl}}(X) \to M(X)$ of Witt groups.

Let $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ be a formation. Then the formation

$$\Big((\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} \varphi & 0 \\ 0 & -\varphi \end{bmatrix}), (\mathcal{C} \oplus \mathcal{A}, \begin{bmatrix} \gamma & 0 \\ 0 & \alpha \end{bmatrix}), (\mathcal{F}, \begin{bmatrix} 1 \\ 1 \end{bmatrix})\Big)$$

is an extension of $((\mathcal{F}, \varphi), (\mathcal{C}, \gamma), (\mathcal{A}, \alpha))$ by $\mathcal{A}$. Indeed, the quotient of $(\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} \varphi & 0 \\ 0 & -\varphi \end{bmatrix})$ by the sublagrangian $\begin{bmatrix} 0 \\ \alpha \end{bmatrix} : \mathcal{A} \to \mathcal{F} \oplus \mathcal{F}$ is isomorphic to $(\mathcal{F}, \varphi)$ in an obvious way. We also have the extensions

$$0 \to \mathcal{A} \xrightarrow{\begin{bmatrix} 0 \\ 1 \end{bmatrix}} \mathcal{C} \oplus \mathcal{A} \xrightarrow{[1\ 0]} \mathcal{A} \to 0$$

and

$$0 \to \mathcal{A} \xrightarrow{\alpha} \mathcal{F} \xrightarrow{-\alpha^\top \circ \varphi} \mathcal{A}^\top \to 0$$

of vector bundles and it is trivial check that all this fits together. It follows that the formation

$$\Big((\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} \varphi & 0 \\ 0 & -\varphi \end{bmatrix}), (\mathcal{F}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}), (\mathcal{C} \oplus \mathcal{A}, \begin{bmatrix} \gamma & 0 \\ 0 & \alpha \end{bmatrix})\Big)$$

has the same class as $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$. As we are assuming that 2 is invertible, we have the isomorphism $\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \varphi & -\varphi \end{bmatrix}$ from $(\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} \varphi & 0 \\ 0 & -\varphi \end{bmatrix})$ to $(\mathcal{F} \oplus \mathcal{F}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix})$. It follows that the former formation is isomorphic to

$$\Big((\mathcal{F} \oplus \mathcal{F}^\top, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}), (\mathcal{F}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}), (\mathcal{C} \oplus \mathcal{A}, \begin{bmatrix} \frac{1}{2}\gamma & \frac{1}{2}\alpha \\ \varphi\circ\gamma & -\varphi\circ\alpha \end{bmatrix})\Big)$$

This is the formation arising from the split-formation

$$(\mathcal{F}, \mathcal{C} \oplus \mathcal{A}, \left[\tfrac{1}{2}\gamma \ \ \tfrac{1}{2}\alpha\right], \left[\varphi\circ\gamma \ \ -\varphi\circ\alpha\right])$$

It follows that our morphism $M^{\mathrm{spl}}(X) \to M(X)$ of Witt groups is an epimorphism.

We want to show that $M^{\mathrm{spl}}(X) \to M(X)$ is an isomorphism. By mapping the formation $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ to the split-formation

$$(\mathcal{F}, \mathcal{C} \oplus \mathcal{A}, \left[\tfrac{1}{2}\gamma \ \ \tfrac{1}{2}\alpha\right], \left[\varphi\circ\gamma \ \ -\varphi\circ\alpha\right])$$

we clearly get a morphism of Grothendieck groups of isomorphism classes. We have to check that the defining relations for $M(X)$ map to valid relations in $M^{\mathrm{spl}}(X)$.

We first look at extensions. So let $((\mathcal{F}_1, \varphi_1), (\mathcal{A}_1, \alpha_1), (\mathcal{C}_1, \gamma_1))$ be an extension of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ by $\mathcal{Z}$. We use the notations from the definition of such an extension. Using the short exact sequences

$$0 \to \mathcal{Z} \xrightarrow{\kappa\circ\iota} \mathcal{F}_1 \xrightarrow{\kappa^\top\circ\varphi_1} \mathcal{V}^\top \to 0$$

and

$$0 \to \mathcal{C} \oplus \mathcal{A}_1 \xrightarrow{\left[\begin{smallmatrix} \iota_{\mathcal{C}} & 0 \\ 0 & 1 \end{smallmatrix}\right]} \mathcal{C} \oplus \mathcal{A} \xrightarrow{[\pi_{\mathcal{C}} \ 0]} \mathcal{Z}^\top \to 0$$

one can see that

$$(\mathcal{V}^\top, \mathcal{C} \oplus \mathcal{A}_1, \left[\tfrac{1}{2}\kappa^\top\circ\varphi_1\circ\gamma_1\iota_{\mathcal{C}} \ \ \tfrac{1}{2}\kappa^\top\circ\varphi_1\circ\alpha_1\right], \left[\underline{\gamma} \ \ -\underline{\alpha}\right])$$

is a quotient of

$$(\mathcal{F}_1, \mathcal{C}_1 \oplus \mathcal{A}_1, \left[\tfrac{1}{2}\gamma_1 \ \ \tfrac{1}{2}\alpha_1\right], \left[\varphi_1\circ\gamma_1 \ \ -\varphi_1\circ\alpha_1\right])$$

Using the short exact sequence

$$0 \to \mathcal{C} \xrightarrow{\left[\begin{smallmatrix} \gamma \\ \iota_{\mathcal{C}} \end{smallmatrix}\right]} \mathcal{F} \oplus \mathcal{C}_1 \xrightarrow{\left[\pi^\top\circ\varphi \ \ -\kappa^\top\circ\varphi_1\circ\gamma_1\right]} \mathcal{V}^\top \to 0$$

and the short exact sequence that we get by adding $\mathcal{C}$ to the left hand part of the short exact sequence

$$0 \to \mathcal{A}_1 \xrightarrow{\left[\begin{smallmatrix} \pi_{\mathcal{A}} \\ \gamma_1^\top\circ\varphi_1\circ\alpha_1 \end{smallmatrix}\right]} \mathcal{A} \oplus \mathcal{C}_1^\top \xrightarrow{\left[-\gamma^\top\circ\varphi\circ\alpha \ \ \iota_{\mathcal{C}}^\top\right]} \mathcal{C}^\top \to 0$$

one can see that the direct sum

$$(\mathcal{F}, \mathcal{C} \oplus \mathcal{A}, \left[\tfrac{1}{2}\gamma \ \ \tfrac{1}{2}\alpha\right], \left[\varphi\circ\gamma \ \ -\varphi\circ\alpha\right]) \oplus \left(\mathcal{C}_1, \mathcal{C}_1^\top, 0, 1\right)$$

is an extension of

$$(\mathcal{V}^\top, \mathcal{C} \oplus \mathcal{A}_1, \left[\tfrac{1}{2}\kappa^\top\circ\varphi_1\circ\gamma_1\circ\iota_{\mathcal{C}} \ \ \tfrac{1}{2}\kappa^\top\circ\varphi_1\circ\alpha_1\right], \left[\underline{\gamma} \ \ -\underline{\alpha}\right])$$

As $(\mathcal{C}_1, \mathcal{C}_1^\top, 0, 1)$ clearly is an extension of the zero split-formation, we conclude that
$$(\mathcal{F}_1, \mathcal{C}_1 \oplus \mathcal{A}_1, \left[\tfrac{1}{2}\gamma_1 \;\; \tfrac{1}{2}\alpha_1\right], \left[\varphi_1\circ\gamma_1 \;\; -\varphi_1\circ\alpha_1\right])$$
and
$$(\mathcal{F}, \mathcal{C} \oplus \mathcal{A}, \left[\tfrac{1}{2}\gamma \;\; \tfrac{1}{2}\alpha\right], \left[\varphi\circ\gamma \;\; -\varphi\circ\alpha\right])$$
have the same class in $M^{\mathrm{spl}}(X)$.
We now consider the additivity relations which we write as
$$[(\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{B},\beta)] + [(\mathcal{F},\varphi),(\mathcal{B},\beta),(\mathcal{C},\gamma)] = [(\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{C},\gamma)]$$
It is easy to see that these are equivalent to the relations
$$[(\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{B},\beta)] + [(\mathcal{F},\varphi),(\mathcal{B},\beta),(\mathcal{C},\gamma)] + [(\mathcal{F},\varphi),(\mathcal{C},\gamma),(\mathcal{A},\alpha)] = 0$$
and
$$[(\mathcal{F},\varphi),(\mathcal{A},\alpha),(\mathcal{A},\alpha)] = 0$$
Writing $(\mathcal{G},\chi) = (\mathcal{F},\varphi)\oplus(\mathcal{F},\varphi)\oplus(\mathcal{F},\varphi)$, $\mathcal{D} = \mathcal{A}\oplus\mathcal{B}\oplus\mathcal{C}$ and $\delta = \alpha\oplus\beta\oplus\gamma$, the left hand side of the former relation is the class of $((\mathcal{G},\chi),(\mathcal{D},\delta),(\mathcal{D},\sigma\circ\delta))$, where
$$\sigma = \begin{bmatrix} 0 & 1_{\mathcal{F}} & 0 \\ 0 & 0 & 1_{\mathcal{F}} \\ 1_{\mathcal{F}} & 0 & 0 \end{bmatrix}$$
In fact, σ is an automorphism of $(\mathcal{G},\chi)$. Furthermore, $\sigma^3 = 1$, hence $(\sigma+1)\circ(\sigma^2-\sigma+1) = 1+1$. As 2 is invertible, it follows that $\sigma+1$ is invertible. The left hand side of the second relation is of the same type with the automorphism of $(\mathcal{F},\varphi)$ being the identity.
From these considerations it follows that it now suffices to prove that if $\alpha : \mathcal{A}\to\mathcal{F}$ is a lagrangian of $(\mathcal{F},\varphi)$ and σ is an automorphism of $(\mathcal{F},\varphi)$ such that $\sigma+1$ is invertible then the split-formation
$$(\mathcal{F}, \mathcal{A}\oplus\mathcal{A}, \left[\tfrac{1}{2}\sigma\circ\alpha \;\; \tfrac{1}{2}\alpha\right], \left[\varphi\circ\sigma\circ\alpha \;\; -\varphi\circ\alpha\right])$$
is metabolic. Indeed, it is not too difficult to check that the elementary split-formation
$$\left(\mathcal{F}\oplus\mathcal{A}, \mathcal{F}\oplus\mathcal{A}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 2\varphi\circ(\sigma-1)\circ(\sigma+1)^{-1} & -\varphi\circ\alpha \\ \alpha^\top\circ\varphi & 0 \end{bmatrix}\right)$$
is an extension. The corresponding short exact sequences are
$$0 \to \mathcal{A} \xrightarrow{\begin{bmatrix} -\frac{1}{2}\alpha \\ 1 \end{bmatrix}} \mathcal{F}\oplus\mathcal{A} \xrightarrow{\begin{bmatrix} 1 & \frac{1}{2}\alpha \end{bmatrix}} \mathcal{F} \to 0$$
and
$$0 \to \mathcal{A}\oplus\mathcal{A} \xrightarrow{\begin{bmatrix} \frac{1}{2}(\sigma+1)\circ\alpha & 0 \\ -1 & 1 \end{bmatrix}} \mathcal{F}\oplus\mathcal{A} \xrightarrow{\begin{bmatrix} 2\alpha^\top\circ\varphi\circ(\sigma+1)^{-1} & 0 \end{bmatrix}} \mathcal{A}^\top \to 0$$

So we now also have a morphism $M(X) \to M^{\mathrm{spl}}(X)$. By construction, the composition $M(X) \to M^{\mathrm{spl}}(X) \to M(X)$ is the identity. To show that the other composition is also the identity it suffices to show that for any split-formation $(\mathcal{A}, \mathcal{C}, \alpha, \varepsilon)$ the split-formation

$$\left(\mathcal{A} \oplus \mathcal{A}^{\top}, \mathcal{C} \oplus \mathcal{A}, \begin{bmatrix} \frac{1}{2}\alpha & \frac{1}{2} \\ \frac{1}{2}\gamma & 0 \end{bmatrix}, \begin{bmatrix} \gamma & 0 \\ \alpha & -1 \end{bmatrix}\right)$$

is an extension. But that is easy.

$$0 \to \mathcal{A}^{\top} \xrightarrow{\begin{bmatrix} 0 \\ -1 \end{bmatrix}} \mathcal{A} \oplus \mathcal{A}^{\top} \xrightarrow{[1\ 0]} \mathcal{F} \to 0$$

and

$$0 \to \mathcal{C} \xrightarrow{\begin{bmatrix} 1 \\ \alpha \end{bmatrix}} \mathcal{C} \oplus \mathcal{A} \xrightarrow{[-\alpha\ 1]} \mathcal{A} \to 0$$

are corresponding short exact sequences.
This all proves that the natural morphism $M^{\mathrm{spl}}(X) \to M(X)$ of Witt groups is an isomorphism.

We saw that the formation $((\mathcal{F}, \varphi), (\mathcal{C}, \gamma), (\mathcal{A}, \alpha))$ has the same class as

$$\left((\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} \varphi & 0 \\ 0 & -\varphi \end{bmatrix}), (\mathcal{C} \oplus \mathcal{A}, \begin{bmatrix} \gamma & 0 \\ 0 & \alpha \end{bmatrix}), (\mathcal{F}, \begin{bmatrix} 1 \\ 1 \end{bmatrix})\right)$$

(and this did not depend on 2 being invertible). Changing the order of the summands in the first two components, we see that this formation is isomorphic to

$$\left((\mathcal{F} \oplus \mathcal{F}, \begin{bmatrix} -\varphi & 0 \\ 0 & \varphi \end{bmatrix}), (\mathcal{A} \oplus \mathcal{C}, \begin{bmatrix} \alpha & 0 \\ 0 & \gamma \end{bmatrix}), (\mathcal{F}, \begin{bmatrix} 1 \\ 1 \end{bmatrix})\right)$$

But, by the same argument as before, this last formation has the same class as $((\mathcal{F}, -\varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$. This shows that we can also describe the inverse of the class of $((\mathcal{F}, \varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$ as the class of $((\mathcal{F}, -\varphi), (\mathcal{A}, \alpha), (\mathcal{C}, \gamma))$.

Conclusion and Remarks

In this concluding section we prove the main result of the paper, the following theorem.

Theorem: There is a natural exact sequence

$$W(X) \to W(Y) \to M_{\top}(X)$$

of Witt groups.
Proof: Because of the results of Section 2.3 we may use $M_{\top}^{\mathrm{spl}}(X)$ instead of $M_{\top}(X)$.

Computations in Section 1.4 show that to any NN-pair $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$ there is associated a split-formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$. This clearly gives rise to a morphism from the Grothendieck group $K(NN)$ of isomorphism classes of NN-pairs to the Grothendieck group of isomorphism classes of split-formations. Composing with the natural projection we get a natural morphism $K(NN)\to M^{\mathrm{spl}}_{\top}(X)$. The results in Section 1.4 also show that extensions of NN-pairs map to extensions of split-formations (with the same vector bundle $\mathcal{Z}$).

Now let $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$ be an NN-pair such that the corresponding split-formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is metabolic. Then there is an extension $(\mathcal{A}_1,\mathcal{C}_1,\alpha_1,\varepsilon_1)$ of $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ with an isomorphism α_1. By the result of Section 1.5 there is a corresponding extension $((\mathcal{E}_1,\chi_1),(\mathcal{L}\otimes\mathcal{A}_1,\mu_1))$ of $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$. But α_1 being an isomorphism means exactly that $f_*(\mathcal{E}_1(-1))=0$ and $R^1f_*(\mathcal{E}_1(-1))=0$. So, by Section 1.1, the space $(\mathcal{E}_1,\chi_1)$ comes from X. As $(\mathcal{E}_1,\chi_1)$ and $(\mathcal{E},\chi)$ have the same class in $W(Y)$, it follows that the class of $(\mathcal{E},\chi)$ also lies in the image of $W(X)$ in $W(Y)$.

Assume now only that the split-formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ corresponding to $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$ has trivial class in $M^{\mathrm{spl}}_{\top}(X)$. Then there is a metabolic split-formation $(\mathcal{A}_0,\mathcal{C}_0,\alpha_0,\varepsilon_0)$ such that $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)\oplus(\mathcal{A}_0,\mathcal{C}_0,\alpha_0,\varepsilon_0)$ is metabolic. In an example at the end of Section 1.4 we saw, in the present parlance, that any elementary split-formation is the formation corresponding to an NN-pair. As quotients of split-formations correspond to quotients of NN-pairs, we conclude that any metabolic split-formation comes from an NN-pair. In particular, there is an NN-pair $((\mathcal{E}_0,\chi_0),(\mathcal{L}\otimes\mathcal{A}_0,\mu_0))$ such that the corresponding split-formation is $(\mathcal{A}_0,\mathcal{C}_0,\alpha_0,\varepsilon_0)$. Then our hypothesis says that the formation corresponding to $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))\oplus((\mathcal{E}_0,\chi_0),(\mathcal{L}\otimes\mathcal{A}_0,\mu_0))$ is metabolic. By the above, it follows that the classes of $(\mathcal{E}_0,\chi_0)$ and $(\mathcal{E},\chi)\oplus(\mathcal{E}_0,\chi_0)$ in $W(Y)$ both come from $W(X)$. We conclude that the class of $(\mathcal{E},\chi)$ in $W(Y)$ also comes from $W(X)$.

Now assume, conversely, that $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$ is an NN-pair such that the class of $(\mathcal{E},\chi)$ in $W(Y)$ comes from $W(X)$. Let $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ be the corresponding split-formation. From Theorem 2 in Section 1.2 and results in Section 1.4 it follows that there is an extension $((\mathcal{E}_1,\chi_1),(\mathcal{L}\otimes\mathcal{A}_1,\mu_1))$ of $((\mathcal{E},\chi),(\mathcal{L}\otimes\mathcal{A},\mu))$ such that the symmetric bilinear space $(\mathcal{E}_1,\chi_1)$ comes from X. Then the corresponding split-formation is elementary. But that split-formation then is an extension of $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ so it follows that $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ is metabolic.

We have now seen that the natural epimorphism $K(NN)\to W(Y)$ maps the kernel of our natural morphism $K(NN)\to M^{\mathrm{spl}}_{\top}(X)$ onto the image of $W(X)$ in $W(Y)$. This means that the morphisms $K(NN)\to W(Y)$ and $K(NN)\to M^{\mathrm{spl}}_{\top}(X)$ induce a morphism $W(Y)\to M^{\mathrm{spl}}_{\top}(X)$ making the sequence $W(X)\to W(Y)\to M^{\mathrm{spl}}_{\top}(X)$ exact.

This finishes the proof of the theorem. Note that we get, as a side result, that if the formation $(\mathcal{A},\mathcal{C},\alpha,\varepsilon)$ has trivial class in $M^{\mathrm{spl}}_{\top}(X)$ then it is metabolic.

If the rank 2 vector bundle $\mathcal{S}$ over X has a quotient bundle of rank 1 then there is a section $X \to Y$. It follows that $W(X) \to W(Y)$ is a monomorphism. In particular, this holds if $Y = \mathbf{P}^1_X$, the trivial projective line bundle over X. According to Walter, [W], the natural morphism $W(Y) \to M_\top(X)$ is an epimorphism in the case that $\mathcal{S}$ has a quotient bundle of rank 1. So in this case there is a natural short exact sequence

$$0 \to W(X) \to W(Y) \to M_\top(X) \to 0$$

We have not yet been able to prove that with the methods of this paper. But we can handle a special case, the case that X is affine and Y is the trivial projective line bundle over X. In fact, this was one of our original result, back in the early 1980's. As the terminology of that proof is different from what has been used here, we shall refrain from giving it.

References

[A] J.K. Arason, "Extensions of symmetric bilinear spaces", Science Institute, University of Iceland, 1981.

[B] P. Balmer, Derived Witt groups of a scheme, *J. pure Appl. Algebra* **141** (1999), 101-129.

[H] R. Hartshorne, "Algebraic Geometry", Springer-Verlag, New York Heidelberg Berlin, 1977.

[R] A. Ranicki, "Exact Sequences in the Algebraic Theory of Surgery", Princeton University Press, Princeton, 1981.

[W] C. Walter, Private communication.

Jón Kr. Arason
Science Institute,
University of Iceland
Reykjavik,
Iceland
jka@hi.is

Some Algebraic Aspects of Quadratic Forms over Fields of Characteristic Two

Ricardo Baeza[1]

Received: May 10, 2001

Communicated by Ulf Rehmann

Abstract. This paper is intended to give a survey in the algebraic theory of quadratic forms over fields of characteristic two. The relationship between differential forms and quadratics and bilinear forms over such fields discovered by Kato is used to reduced some problems on quadratics forms to concrete questions about differential forms, which in general are easier to handle.

1991 Mathematics Subject Classification: 11 E04, 11 E81, 12 E05, 12 F20
Keywords and Phrases: Keywords and Phrases: Bilinear forms, Quadratic forms, Differential forms, Witt-groups, Function fields.

1 Introduction.

In his historical account on the algebraic theory of quadratic forms (s [Sch]), Scharlau remarks that fields of characteristic two have remained the pariahs of the theory. Nevertheless, as he also mentions right before the above remark (s. loc. cit.), some aspects of the theory over these fields are more interesting and richer, because of the interplay of symmetric bilinear and quadratic forms, as well as both separable and purely inseparable quadratic extensions have to be considered. The purpose of this brief survey article is to show how these aspects work, and how some questions related to Milnor's conjecture for fields with $2 \neq 0$, can be answered in a more elementary way in the case of characteristic two.

[1]Partially supported by Fondecyt 1000392 and Programa Formas Cuadraticas, Universidad de Talca.

We will focus our attention on the $W(F)$-module structure of $W_q(F)$, where $W(F)$ is the Witt-ring of a field F with $2 = 0$ and $W_q(F)$ is the Witt-group of quadratic forms over F (s. [Mi]$_2$, [Sa] and section 2). If $I \subset W(F)$ is the maximal ideal of $W(F)$, then we have the graded Witt-ring

$$gr_I W(F) = \bigoplus_{n=0}^{\infty} I^n/I^{n+1}$$

and the graded $gr_I W(F)$- module

$$gr_I W_q(F) = \bigoplus_{n=0}^{\infty} I^n W_q(F)/I^{n+1}W_q(F).$$

The structure of this module is explained in sections 3 and 4. Section 3 deals with the relationship established by Kato between differential forms over F and symmetric bilinear and quadratic forms. If $k_*(F)$ denotes Milnor's graded k-ring of F, we introduce in section 4 a graded $k_*(F)$-module, defined by generators and relations, which describes the graded $gr_I W(F)$-module $gr_I W_q(F)$. In section 5 we examine the behaviour of this module under certain field extensions, particularly function field extensions of quadrics defined by Pfister-forms. As an application of these results we mention, how Knebusch's degree conjecture for fields with $2 = 0$ follows from them. The results of section 5, (c.f. (5.10), (5.11), (5.14), (5.16)), cited from [Ar-Ba]$_3$ and [Ar-Ba]$_4$ have not been published yet, but these manuscripts can be found at the server "Linear Algebraic Groups and Related Structures" http://www.mathematik.uni-bielefeld.de/LAG/.

2 Basic definitions.

Let F be a field of characteristic two. A symmetric bilinear form $b : V \times V \longrightarrow F$ defined on an n-dimensional F-vector space V is non-singular if $b(x, y) = 0$ for all $x \in V$ implies $y = 0$. (V, b) is anisotropic if $b(x, x) \neq 0$ for all $x \neq 0$, and in this case it is easy to see that (V, b) admits an orthogonal basis (s. [Mi]$_2$ for example). If $a \in F^* = F \setminus \{0\}$ we will denote by $< a >$ the one dimensional form axy, and by $< a_1, \cdots, a_n >$ $(a_i \in F^*)$ the orthogonal sum $< a_1 > \perp \cdots \perp < a_n >$ A non singular quadratic form on V is a map $q : V \longrightarrow F$ such that $q(\lambda x) = \lambda^2 q(x)$ and $b_q(x, y) = q(x + y) - q(x) - q(y)$ is a symmetric non singular bilinear form on V. Since $b_q(x, x) = 0$, n must be even. The most simple non singular quadratic forms over F are the forms $ax^2 + xy + by^2$ with $a, b \in F$ (i.e. $q : Fe \oplus Ff \longrightarrow F, \quad q(e) = a, \quad q(f) = b, \quad b_q(e, f) = b_q(f, e) = 1$), which we will denote by $[a, b]$. Any non singular quadratic form over F is of the form $[a_1, b_1] \perp \cdots \perp [a_m, b_m]$. Scaling a quadratic form q by $a \in F^*$ means $(aq)(x) = aq(x)$. This extends to an operation of bilinear forms on quadratic forms by $< a_1, \cdots a_n > \cdot q = a_1 q \perp \cdots \perp a_n q$. Besides the dimension, the most simple invariant of a symmetric bilinear form $b =< a_1, \cdots a_n >$ is its

discriminant $d(b) = a_1 \cdots a_n \in F^*/F^{*^2}$ If $q = [a_1, b_1] \perp \cdots \perp [a_n, b_n]$ is a quadratic form the analogue of the discriminant is its Arf-invariant $A(q) = a_1b_1 + \cdots + a_nb_n \in F/\wp F$, where $\wp F = \{a^2 - a \backslash a \in F\}$.
One can write $[a, b] =< a > [1, ab]$ if $a \neq 0$, so that in general one usually writes a quadratic form q as $q =< a_1 > [1, b_1] \perp \cdots \perp< a_n > [1, b_n]$, and hence its Arf-invariant is $A(q) = b_1 + \cdots + b_n \in F/\wp F$ (s. [A], $[\mathrm{Ba}]_1$, [Sa]). For quadratic forms (V, q) we have also the Clifford - algebra $C(q)$, which defines an element $w(q) \in Br(F) =$ Brauer group of F. If $q = \overset{m}{\underset{1}{\perp}} < a_i > [1, b_i]$, then $w(q) = \overset{m}{\underset{1}{\otimes}} (a_i, b_i] \in Br(F)$, where $(a, b]$ denotes the quaternion algebra $F \oplus Fe \oplus Ff \oplus Fef$ with $e^2 = a,\ f^2 + f = b,\ ef + fe = e$.
A symmetric bilinear form (V, b) is called metabolic if V contains a subspace $W \subseteq V$ with $W = W^\perp$ ($dim\ W = \frac{1}{2}\ dim\ V$). Two bilinear forms $b_1,\ b_2$ are Witt-equivalent if $b_1 \perp m_1 \cong b_2 \perp m_2$, where $m_1,\ m_2$ are metabolic. The set of classes $W(F)$ of symmetric non singular bilinear forms is a ring, additively generated by the classes $< a >,\ a \in F^*$ with relations $< a > + < b > = < a + b > + < ab(a + b) >$ if $a + b \neq 0$, $< a > + < a > = 0$ and $< a > \cdot < b > = < ab >$. We denote by $I_F \subset W(F)$ the maximal ideal of even dimensional forms (s. $[\mathrm{Mi}]_2$, [Sa] for basic facts on $W(F)$). A quadratic form (V, q) is hyperbolic if V contains a totally isotropic subspace $W \subset V$ with $dim\ W = \frac{1}{2}\ dim\ V$. The form $[0, 0] = \mathbb{H}$ is the hyperbolic plane and every hyperbolic space is of the form $\mathbb{H} \perp \ldots \perp \mathbb{H}$. The forms $q_1,\ q_2$ are Witt-equivalent if $q_1 \perp r \times \mathbb{H} \cong q_2 \perp s \times \mathbb{H}$ $(r, s \geq 0)$ and we denote by $W_q(F)$ the Witt-group of such classes. The action defined above of bilinear forms on quadratic forms induces a $W(F)$-module structure on $W_q(F)$.
I_F is additively generated by the 1-fold Pfister forms $< 1, a >,\quad a \in F^*$, so that for all $n \geq 1$, I_F^n is generated by the n-fold bilinear Pfister forms $\ll a_1, \cdots, a_n \gg = < 1, a_1 > \otimes \cdots \otimes < 1, a_n >$. These ideals define submodules $I_F^n \cdot W_q(F)$ of $W_q(F)$, which are additively generated by the n-fold quadratic Pfister forms $\ll a_1, \cdots a_n, a\,|] = \ll a_1, \cdots a_n \gg \otimes [1, a]$, $a_i \in F^*,\ a \in F$ (s. $[\mathrm{Ba}]_1$, [Sa] for details on these forms).
Thus we have now two filtrations

$$W(F) \supseteq I_F \supset I_F^2 \supset \cdots \supset I_F^n \cdots$$

$$W_q(F) \supseteq IW_q(F) \supset I^2W_q(F) \supset \cdots \supset I^nW_q(F) \supset \cdots$$

and we will be mainly concerned with the quotients I_F^n/I_F^{n+1} and $I^nW_q(F)/I^{n+1}W_q(F)$ which we denote by $\overline{I}_F^{\,n}$ and $\overline{I^nW_q}(F)$ respectively. One easily checks that $dim : \overline{I}_F^{\,0} \xrightarrow{\sim} \mathbb{Z}/2\mathbb{Z}$, $d : \overline{I}_F \xrightarrow{\sim} F^*/F^{*^2}$ and $A : \overline{I^0W_q}(F) \xrightarrow{\sim} F/\wp F$. The main result of [Sa] states that $w : \overline{IW_q}(F) \xrightarrow{\sim} Br(F)_2 =$ 2-torsion part of $Br(F)$. The surjectivity of w is a consequence of well-known results on p-algebras for $p = 2$ (s. [Al]), and the injectivity is shown in [Sa] by an elementary induction argument (notice

that the isomorphism $\overline{IW_q}(F) \xrightarrow{\sim} Br(F)_2$ is the analogue of Merkurjev's result $I_F^2/I_F^3 \xrightarrow{\sim} Br(F)_2$ for fields with $2 \neq 0$).
The higher groups $\overline{I}_F^n$ and $\overline{I^nW_q}(F)$ will be studied in the next section.

3 Differential forms and its relationship to quadratic and bilinear forms

The basic reference for what follows is Kato's fundamental paper [Ka]$_1$. Let Ω_F^1 be the F-vector space generated (over F) by the symbols da, $a \in F$, with the relations $d(ab) = bda + adb$. In particular $d(F^2) = 0$, and hence the map $d: F \longrightarrow \Omega_F^1$ is F^2-linear. Let $\Omega_F^n = \bigwedge^n \Omega_F^1$ be the F-space of n-differential forms over F. The map $d : F \longrightarrow \Omega_F^1$ extends to $d : \Omega_F^n \longrightarrow \Omega_F^{n+1}$ for all $n \geq 1$ by $d(xdx_1 \wedge \cdots \wedge dx_n) = dx \wedge dx_1 \wedge \cdots \wedge dx_n$. Recall that a 2-basis of F is a set $\{a_i\ ,\ i \in I\} \subset F$ such that the elements $\{a^\varepsilon = \prod_{i \in I} a_i^{\varepsilon_i}, \quad \varepsilon = (\varepsilon_i,\ i \in I), \quad \varepsilon_i \in \{0, 1\}$ and almost all $\varepsilon_i = 0\}$ form a F^2-basis of F. If $\{a_1, a_2, \dots\}$ is a 2-basis of F, then the forms $\dfrac{da_{i_1}}{a_{i_1}} \wedge \dots \wedge \dfrac{da_{i_n}}{a_{i_n}}$ $\quad 1 \leq i_1 < \cdots < i_n$ form a F-basis of Ω_F^n. Fixing such a 2-basis, we define

$$[\Omega_F^n]^2 = \{ \sum_{i_1 < \cdots < i_n} c_{i_1 \cdots i_n}^2 \frac{da_{i_1}}{a_{i_1}} \wedge \cdots \wedge \frac{da_{i_n}}{a_{i_n}}, \quad c_{i_1 \cdots i_n} \in F\}$$

which depends on the choice of the 2-basis. Then in [Ca] it is shown that the space $Z_F^n = ker(d : \Omega_F^n \longrightarrow \Omega_F^{n+1})$ has a direct-sum decomposition $Z_F^n = [\Omega_F^n]^2 \oplus d\Omega_F^{n-1}$.

One now defines a homomorphism

(3.1) $\qquad C : Z_F^n \longrightarrow \Omega_F^n$

by

$$C(\sum_{i_1 < \cdots < i_n} c_{i_1 \cdots i_n}^2 \frac{da_{i_1}}{a_{i_1}} \wedge \cdots \wedge \frac{da_{i_n}}{a_{i_n}} + d\eta) = \sum_{i_1 < \cdots < i_n} c_{i_1 \cdots i_n} \frac{da_{i_1}}{a_{i_1}} \wedge \cdots \wedge \frac{da_{i_n}}{a_{i_n}}$$

C obviously does not depend on the choice of the 2-basis and induces an isomorphism $\overline{C} : Z_F^n/d\Omega_F^{n-1} \xrightarrow{\sim} \Omega_F^n$ of abelian groups.
We will call C the Cartier-operator. Let us define now the homomorphism $\wp = \overline{C}^{\,-1} - 1 : \Omega_F^n \longrightarrow \Omega_F^n/d\Omega_F^{n-1}$, which is given on generators by $\wp(x\frac{dx_1}{x_1} \wedge \cdots \wedge \frac{dx_n}{x_n}) = (x^2 - x)\frac{dx_1}{x_1} \wedge \cdots \wedge \frac{dx_n}{x_n} \quad mod\ d\Omega_F^{n-1}$.
One can define a 2-basis dependent homomorphism $\wp : \Omega_F^n \to \Omega_F^n$ as follows. Fix a 2-basis $\mathcal{B} = \{a_1, a_2, \cdots\}$ of F. Then we set

$$\wp\left(\sum_{i_1<\cdots<i_n} c_{i_1\cdots i_n}\frac{da_{i_1}}{a_{i_1}}\wedge\cdots\wedge\frac{da_{i_n}}{a_{i_n}}\right)$$
$$=\sum_{i_1<\cdots<i_n}(c^2_{i_1\cdots i_n}-c_{i_1\cdots i_n})\frac{da_{i_1}}{a_{i_1}}\wedge\cdots\wedge\frac{da_{i_n}}{a_{i_n}}.$$

If for $\omega=\sum_{i_1<\cdots<i_n} c_{i_1\cdots i_n}\frac{da_{i_1}}{a_{i_1}}\wedge\cdots\wedge\frac{da_{i_n}}{a_{i_n}}$ we set

$$\omega^{[2]}=\sum_{i_1<\cdots<i_n} c^2_{i_1\cdots i_n}\frac{da_{i_1}}{a_{i_1}}\wedge\cdots\wedge\frac{da_{i_n}}{a_{i_n}},$$

then $\wp\omega=\omega^{[2]}-\omega$.
Obviously if we change the 2-basis, the image of $\omega\in\Omega^n_F$ under the new $\wp$-operator differs from $\wp\omega$ by an exact form. We will use this type of operator in section 5.
Let $\nu_F(n)=Ker(\wp)$ and $H^{n+1}(F)=Coker(\wp)$, so that $0\to\nu_F(n)\to\Omega^n_F\overset{\wp}{\to}\Omega^n_F/d\Omega^{n-1}_F\to H^{n+1}(F)\to 0$ is exact. An obvious characterization of $\nu_F(n)$ is the following

(3.2) Lemma. $\quad\nu_F(n)=\{\omega\in\Omega^n_F\backslash d\omega=0, C(\omega)=\omega\}$

In $[\mathrm{Ka}]_1$ it is shown that $\nu_F(n)$ is additively generated by the pure logarithmic differentials $\frac{dx_1}{x_1}\wedge\cdots\wedge\frac{dx_n}{x_n}$, which is a direct consequence of lemma 2 in $[\mathrm{Ka}]_2$. Since we will refer frequently to this lemma, we will state it explicitly. Let $\mathcal{B}=\{a_i,\ i\in I\}$ be a 2-basis of F and endow I with a totally ordering. For any $j\in I$ set F_i resp. $F_{\leq j}$ for the subfield of F generated over F^2 by the elements a_i with $i<j$ resp. $i\leq j$. Endow with the lexicographic ordering the set $\sum_n$ of functions $\alpha:\{1,\cdots n\}\to I$ with $\alpha(i)<\alpha(j)$ whenever $i<j$. Then $\{da_{\alpha(1)}\wedge\cdots\wedge da_{\alpha(n)},\alpha\in\sum_n\}$ is a F-basis of Ω^n_F and for any $\alpha\in\sum_n$ set $\Omega^n_{F,\alpha}$ resp. $\Omega^n_{F,<\alpha}$ for the subspace of Ω^n_F generated by the elements $da_{\beta(1)}\wedge\cdots\wedge da_{\beta(n)}$ with $\beta\leq\alpha$ resp. $\beta<\alpha$. Then Kato's lemma 2 in $[\mathrm{Ka}]_2$ asserts

(3.3) Lemma. *Let $y\in F,\ \alpha\in\sum_n$ and $\omega_\alpha=\frac{da_{\alpha(1)}}{a_{\alpha(1)}}\wedge\cdots\wedge\frac{da_{\alpha(n)}}{a_{\alpha(n)}}\in\Omega^n_F$, be such that*

$$(y^2-y)\omega_\alpha\in\Omega^n_{F,<\alpha}+d\Omega^{n-1}_F.$$

Then there exist $v\in\Omega^n_{F,<\alpha}$ and $a_i\in F^*_{\alpha(i)}$, $1\leq i\leq n$, with

$$y\omega_\alpha = v + \frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_n}{a_n}.$$

It is clear that the last remark above follows immediately from this result, which we will quote as Kato's lemma in what follows.

One of the main results of [Ka]$_1$ is the fact that there exist two natural isomorphisms

$$\text{(3.4)} \qquad \alpha_F : \nu_F(n) \longrightarrow \overline{I}_F^n$$

$$\text{(3.5)} \qquad \beta_F : H^{n+1}(F) \longrightarrow \overline{I^n W_q}(F)$$

given on generators by

$$\alpha_F\left(\frac{dx_1}{x_1} \wedge \cdots \wedge \frac{dx_n}{x_n}\right) = \overline{\ll x_1, \cdots x_n \gg}$$

$$\beta_F\left(\overline{x\frac{dx_1}{x_1} \wedge \cdots \wedge \frac{dx_n}{x_n}}\right) = \overline{\ll x_1, \cdots x_n, x \,]]}$$

Thus α and β translate many questions on bilinear and quadratic forms to corresponding problems in differential forms, which some times are easier to handle, in particular if one is able to choose a suitable 2-basis of the field F. Nevertheless the use of the isomorphism α can be some times difficult, since in order to compute $\alpha(\omega)$ one must first write $\omega \in \nu_F(n)$ as a sum of pure logarithmic differential forms.

4 Milnor's K-theory.

For any field F Milnor defined in [Mi]$_1$ its K-groups $K_n(F)$ in a purely algebraic manner as follows (s. also Pfister's survey [Pf] for more details). Let $K_1(F)$ be the multiplicative group of F written additively, i.e. $l : F^* \xrightarrow{\sim} K_1(F)$, $l(ab) = l(a) + l(b)$ for $a, b \in F^*$. Set $K_0(F) = \mathbb{Z}$ and $K_n(F) = K_1(F)^{\otimes n}/\mathfrak{I}_n \quad (n \geq 2)$, where $\mathfrak{I}_n$ is the subgroup of $K_1(F)^{\otimes n}$ generated by elements of the form $l(a_1) \otimes \cdots \otimes l(a_n)$ with $a_i + a_j = 1$ for some $i \neq j$. Denote by $l(x_1) \cdots l(x_n)$ the image of $l(x_1) \otimes \cdots \otimes l(x_n)$. Thus the main defining relation of these groups is $l(a)l(1-a) = 0$ in $K_2(F)$ for $a \neq 0, 1$.

Let $k_n(F) = K_n(F)/2K_n(F)$ and form the commutative ring $k_*(F) = k_0(F) \oplus k_1(F) \oplus \cdots$ with $k_0(F) = \mathbb{Z}/2\mathbb{Z}$, $k_1(F) \xrightarrow{\sim} F^*/F^{*^2}$. Milnor defines epimorphisms $s_n : k_n(F) \to \overline{I}_F^n$ by

$$s_n(l(a_1)\cdots l(a_n)) = \overline{\ll a_1, \cdots a_n \gg}$$

and conjectures that they are isomorphisms for all n. If 2=0 in F, then there are also natural homomorphisms (s. [Ka]$_1$)

$$d\log : k_n(F) \longrightarrow \nu_F(n)$$

given by $\quad d\log(l(a_1)\cdots l(a_n)) = \dfrac{da_1}{a_1} \wedge \cdots \wedge \dfrac{da_n}{a_n}.$

A consequence of Kato's lemma is that $d\log$ is an epimorphism. In [Ka]$_1$ it is shown that $d\log$ is an isomorphism, which combined with the isomorphism (3.3) gives us the following main result of [Ka]$_1$

(4.1) THEOREM (KATO) *For any field F with 2=0 there is a commutative diagram of isomorphisms*

$$\begin{tikzcd} k_n(F) \arrow[rr, "d\log"] \arrow[dr, "s_n"'] & & \nu_F(n) \arrow[dl, "\alpha_F"] \\ & \overline{I}_F^{\,n} & \end{tikzcd}$$

The defining relation $l(a)l(a-1) = 0 \quad (a \neq 0,\ 1)$ of the groups $k_n(F)$ corresponds in the case $2 \neq 0$ to the basic fact that the quaternion algebra $(a, 1-a)$ splits. Here (x, y) denotes the quaternion algebra $F \oplus Fe \oplus Ff \oplus Fef,\ e^2 = x,\ f^2 = y,\ ef = -fe.$

But if 2=0 we do not have such interpretation and the groups $k_n(F)$ are suitable only to describe symmetric bilinear forms and for quadratic forms, we need another universal object, which we introduce now. Thus in order to obtain groups which are appropriate to describe the quotients $\overline{I^n W_q}(F)$ by generators and relations one is led to alter Milnor's definition of k_n taking into account the basic relations of quaternion algebras over a field with 2=0. This has been done in [Ar-Ba]$_1$. Let $a \in F^*,\ b \in F$. The quaternion algebra $(a, b]$ is the algebra $F \oplus Fe \oplus Ff \oplus Fef$ with $e^2 = a,\ f^2 + f = b$ and $ef + fe = e$. It holds $(ax^2, b + y + y^2] \cong (a, b]$, and $(a, b]$ splits if and only if
$a \in D_F([1, b]) = \{x^2 + xy + by^2 /\ x, y \in F\}$, and $a \neq 0$. Thus the bilinear map

$$\phi : F^*/F^{*^2} \times F/\wp F \longrightarrow Br(F)_2, \qquad \phi(\bar{a}, \bar{b}) = (a, b]$$

satisfies $\phi(\bar{a}, \bar{b}) = 0$ iff $a \in D_F([1, b])$. The universal symbol for ϕ can be constructed as follows. Let $k_1(F) = F^*/F^{*^2}, \quad h_1(F) = F/\wp F$ and set

$$h_2(F) = \frac{k_1(F) \otimes h_1(F)}{< l(a) \otimes t(b) \quad a \in D_F[1,b],\ a \neq 0 >}$$

(here $t(b)$ is the image of b in $h_1(F) = F/\wp F$).
Thus one obtains a natural homomorphism

$$\phi_F :\ h_2(F) \longrightarrow Br(F)_2$$

which is in fact an isomorphism (s. [Ar-Ba]$_1$,[Sa]). On the other hand we also have a bilinear map

$$k_1(F) \times h_1(F) \longrightarrow H^2(F)$$

given by $\quad (l(a), t(b)) \longrightarrow b\frac{da}{a}$, which induce a natural homomorphism

$$d\log :\ h_2(F) \longrightarrow H^2(F).$$

This homomorphism is also an isomorphism (s. loc. cit), so that the group $h_2(F)$, $H^2(F)$, $Br(F)_2$, $\overline{IW}_q(F)$ are all isomorphic and we have a commutative diagram of isomorphisms

$$(4.2) \qquad \begin{array}{ccc} h_2(F) & \xrightarrow{\phi_F} & Br(F)_2 \\ {\scriptstyle d\log}\downarrow & & \uparrow{\scriptstyle \omega} \\ H^2(F) & \xrightarrow[\beta_F]{} & \overline{IW}_q(F) \end{array}$$

Let now

$$h_n(F) = k_1(F)^{\otimes(n-1)} \otimes h_1(F)/\mathcal{R}_n$$

where $\mathcal{R}_n$ is the subgroup generated by the elements $l(a_1) \otimes \cdots \otimes l(a_{n-1}) \otimes t(b)$ such that either $a_i + a_{i+1} = 1$ for some i or $a_i \in D_F[1,b]$. We denote by $l(a_1) \cdots l(a_{n-1}) t(b)$ in $h_n(F)$ the image of $l(a_1) \otimes \cdots \otimes l(a_{n-1}) \otimes t(b)$.
The natural product $k_r(F) \times h_s(F) \to h_{r+s}(F)$ induces a $k_*(F)$-module structure on $h_*(F) = h_1(F) \oplus h_2(F) \oplus \cdots$. There are natural epimorphisms

$$s_n : h_n(F) \longrightarrow \overline{I^{n-1}W}_q(F)$$

$$d\log : h_n(F) \longrightarrow H^n(F)$$

given by

$$s_n(l(a_1)\cdots l(a_{n-1})t(b)) = \overline{\ll a_1, \cdots a_{n-1}, b \,]]}$$

$$d\log(l(a_1)\cdots l(a_{n-1})t(b)) = \overline{b\frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_{n-1}}{a_{n-1}}}$$

In [Ar-Ba]$_1$ it is shown that $d\log$ is an isomorphism, and combining it with Kato's isomorphism β_F, we conclude also that s_n is an isomorphism. Thus we have (s. [Ar-Ba]$_1$ and [Ka]$_1$)

(4.3) Theorem. *For all n there is a commutative diagram of isomorphisms*

$$\begin{tikzcd} h_n(F) \arrow[rr, "s_n"] \arrow[dr, "d\log"'] & & \overline{I^{n-1}W_q}F(n) \arrow[dl, "\beta_F"] \\ & H^n(F) & \end{tikzcd}$$

Remark. The groups $k_n(F)$ and $h_n(F)$ are related through Galois cohomology. If F_s is a separable closure of F and $G_F = Gal(F_s/F)$ then $k_n(F_s)$ is a G_F-module and it holds (s. [Ar-Ba]$_1$)

$$H^0(G_F, k_n(F_s)) \;\cong\; k_n(F)$$

$$H^1(G_F, k_n(F_s)) \;\cong\; h_{n+1}(F)$$

(s. [Ar]).

5 Behaviour of quadratic and bilinear forms under field extensions.

A natural question is the behaviour of the groups Ω_F^n, $\nu_F(n)$, $H^{n+1}(F)$ resp $\overline{I}_F^n$, $\overline{I^nW_q}(F)$ under field extensions. Since the isomorphisms α_F, β_F (s. (3.4) and (3.5)) are functorial, we only need to study the behaviour of the groups $\nu_F(n)$, $H^{n+1}(F)$, to get information about $\overline{I}_F^n$ and $\overline{I^nW_q}(F)$ (but, as mentioned before, care must be taken with the use of α_F). If L/F is a field extension, we denote by $\Omega_{L/F}^n$ the kernel $Ker(\Omega_F^n \to \Omega_L^n)$, and similarly we define $\nu_{L/F}(n)$, $H^{n+1}(L/F)$, $\overline{I}_{L/F}^n$ and $\overline{I^nW_q}(L/F)$. By the remark above

$\alpha_F : \nu_{L/F}(n) \xrightarrow{\sim} \overline{I}^{\,n}_{L/F}$ and $\beta_F : H^{n+1}(L/F) \xrightarrow{\sim} \overline{I^nW_q}(L/F)$. The easiest group to handle is $\Omega^n_{L/F}$ because a suitable choice (if possible!) of a 2-basis of F and L gives quickly the answer. Since

$$\nu_{L/F}(n) = \nu_F(n) \cap \Omega^n_{L/F} \tag{5.1}$$

one also gets information about $\nu_{L/F}(n)$ knowing $\Omega^n_{L/F}$. Let us now review what we know about these kernels for some field extensions.

(i) Purely Transcendental extensions. If $L = F(X)$, X any set of variables over F, and $\mathcal{B}$ is a 2-basis of F, then $\mathcal{B} \cup \{X\}$ is a 2-basis of $F(X)$. In particular $\Omega^n_F \to \Omega^n_{F(X)}$ is injective and $\Omega^n_{F(X)/F} = 0$. Hence $\nu_{F(X)/F}(n) = 0$. Using Kato's lemma (3.3) one can also show $H^{n+1}(F(X)/F) = 0$ (s. [Ar-Ba]$_3$)

(ii) Quadratic extensions. Let $L = F(\sqrt{b})$, $b \in F\backslash F^2$ be a purely inseparable quadratic extension of F. Choose a 2-basis $\mathcal{B} = \{b_i,\ i \in I\}$ with $b = b_{i_0}$, some $i_0 \in I$. Then $\{b_i, i \in I - \{i_0\}, \sqrt{b}\ \}$ is a 2-basis of $F(\sqrt{b})$ and it is easy to check that

$$\Omega^n_{F(\sqrt{b})/F} = \Omega^{n-1}_F \wedge \frac{db}{b} \tag{5.2}$$

Hence $\nu_{F(\sqrt{b})/F}(n) = \{\omega \wedge \frac{db}{b}\ /\ \omega \in \Omega^{n-1}_F,\ \omega \wedge \frac{db}{b} \in \nu_F(n)\}$. It follows from (5.11) below that

$$\begin{aligned}\nu_{F(\sqrt{b})/F}(n) = \{\omega \wedge \frac{db}{b}\ /\ \omega \in \Omega^{n-1}_F \text{ and } \wp\omega \in a[\Omega^{n-1}_F]^2 + \\ d\Omega^{n-2}_F + \Omega^{n-2}_F \wedge da\}\end{aligned} \tag{5.3}$$

(s. section 3 for the definition of $\wp\omega$).

The corresponding result for $\overline{I}^{\,n}$ is now (s. (5.12) below for a more general statement)

$$\overline{I}^{\,n}_{F(\sqrt{b})/F} = \sum_{x \in F^2(b)^*} \overline{I}^{\,n-1}_F < 1, x > \tag{5.4}$$

Let us now examine the kernel $H^{n+1}(F(\sqrt{b})/F)$.
We have (s.[Ar-Ba]$_3$)

$$H^{n+1}(F(\sqrt{b})/F) = \overline{\Omega^{n-1}_F \wedge \frac{db}{b}} \tag{5.5}$$

The proof of this fact is again based on Kato's lemma and runs briefly as follows. Take $\mathcal{B} = \{b_1 = b,\ b_2, \cdots\}$ a 2-basis of F (one can assume w.l.o.g. that $\mathcal{B}$ is enumerable or even finite), so that $\mathcal{B}' = \{\sqrt{b_1},\ b_2, \cdots\}$ is a 2-basis of $F(\sqrt{b})$. $\overline{\omega} \in H^{n+1}(F(\sqrt{b})/F)$ means $\omega \in \Omega^n_F$ and $\omega = \wp u + dv$ with $u \in \Omega^n_{F(\sqrt{b})}$, $v \in \Omega^{n-1}_{F(\sqrt{b})}$. Order $\mathcal{B}'$ such that $\sqrt{b} > b_i, i = 2, 3 \cdots$. Since $\overline{\Omega^{n-1}_F \wedge \frac{db}{b}} \subseteq H^{n+1}(F(\sqrt{b})/F)$ we may assume that db does not appear in the 2-basis expansion of ω and let $\alpha \in \sum_n$ be the leading index of ω (notice $\alpha(i) > 1$ for all $i = 1, \cdots n$), and let $\beta \in \sum_n$ be the leading index of u. Using Kato's lemma one may assume $\beta \leq \alpha$, and we obtain

$$(\wp u_\alpha + \omega_\alpha)\frac{db_\alpha}{b_\alpha} \equiv dv \qquad \bmod \Omega^n_{F(\sqrt{b}),<\alpha}$$

(here $\frac{db_\alpha}{b_\alpha}$ means $\frac{db_{\alpha(1)}}{b_{\alpha(1)}} \wedge \cdots \wedge \frac{db_{\alpha(n)}}{b_{\alpha(n)}}$)
with $v \in \Omega^{n-1}_{F(\sqrt{b})}$. Since $b_{\alpha(i)} < \sqrt{b}$ for all i, we conclude comparing coefficients that the leading coefficient of dv is in F, so that u_α is defined over F. Thus v may be taken also in Ω^{n-1}_F. Since $\Omega^n_{F(\sqrt{b})/F} = \Omega^{n-1}_F \wedge \frac{db}{b}$, we conclude in Ω^n_F

$$\omega_\alpha \frac{db_\alpha}{b_\alpha} \equiv \wp(u_\alpha)\frac{db_\alpha}{b_\alpha} + dv \qquad \bmod \Omega^n_{F,<\alpha} + \Omega^{n-1}_F \wedge \frac{db}{b}$$

Inserting this relation in ω, we can lower the highest index in ω. This concludes the proof of the claim.
The corresponding kernel for $I^n W_q$ is now

$$\overline{I^n W_q}(F(\sqrt{b})/F) = \overline{\ll b \gg I^{n-1}W_q(F)} \tag{5.6}$$

For quadratic separable extensions of F the corresponding kernels are much easier to compute. Let $L = F(z)$, $z^2 + z = b$ $(b \notin \wp F)$ be a quadratic separable extension of F. Since we can alter b by elements of $\wp F$, we can assume $b \in F^2$. Thus $z \in L^2$ and we see that any 2-basis of F remains a 2-basis of L. In particular $\Omega^n_L = \Omega^n_F \oplus z \cdot \Omega^n_F$. Thus $\Omega^n_{L/F} = 0$ and also $\nu_{L/F} = 0$. The computation of $H^{n+1}(L/F)$ is in this case also very easy. We claim

$$H^{n+1}(L/F) = \overline{b\nu_F(n)} \tag{5.7}$$

For the proof, take $\overline{\omega} \in H^{n+1}(F)$ with $\omega = \wp u + dv$, $u \in \Omega^n_L$, $v \in \Omega^{n-1}_L$ and set $u = u_1 + zu_2$, $v = v_1 + zv_2$ with $u_i \in \Omega^n_F$, $v_i \in \Omega^{n-1}_F$. Inserting in the above equation it follows $\wp u_2 = dv_2 \in d\Omega^{n-1}_F$, and this means $u_2 \in \nu_F(n)$. Moreover $\omega = bu_2^{[2]} + \wp u_1 + dv_1$ in Ω^n_F. But $u_2 \in \nu_F(n)$ implies $u_2^{[2]} \equiv u_2$ ($\bmod\ d\Omega^{n-1}_F$) and since $b \in F^2$, it follows $\omega \equiv bu_2 \mod (\wp\Omega^n_F + d\Omega^{n-1}_F)$, ie $\overline{\omega} = \overline{bu_2}$. This proves (5.7). The corresponding result for quadratic forms is

$$\overline{I^n W_q}(L/F) = \overline{I^n_F \cdot [1, b]} \tag{5.8}$$

(iii) FUNCTION FIELDS OF PFISTER FORMS. Let us fix an anisotropic bilinear n-fold Pfister-form $\phi = \ll a_1, \cdots, a_n \gg$. This means that $\{a_1, \cdots, a_n\}$

are part of 2-basis of F. Let $L = F(\phi)$ be the function field of the quadric $\{\phi(x,x)=0\}$. Thus $L = F(X)(\sqrt{T})$, where $X = \{X_\mu,\ \mu \in S_n\}$ and $T = \sum_\mu a^\mu X_\mu^2$, $a^\mu = \prod_{i=1}^n a_i^{\mu(i)}$, for all $\mu \in S_n$ where S_n denotes the set of maps $\mu : \{1,\cdots,n\} \to \{0,1\}$ whith some $\mu(i)=1$
In [Ar-Ba]$_3$ it is shown that

$$\Omega^m_{L/F} = 0 \qquad \text{if } m < n \tag{5.9}$$

$$\Omega^m_{L/F} = \Omega_F^{m-n} \wedge \frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_n}{a_n} \qquad \text{if } m \geq n \tag{5.10}$$

In particular $\nu_{L/F}(m) = 0$ if $m < n$. The case $m \geq n$ has been considered in [Ar-Ba]$_4$ and the result is:

$$\nu_{L/F}(m) = \{\omega \wedge \tfrac{da_1}{a_1} \wedge \cdots \wedge \tfrac{da_n}{a_n} /\ \omega \in \Omega_F^{m-n},\ \wp\omega \in \sum_{\varepsilon \neq 0} a^\varepsilon [\Omega_F^{m-n}]^2 + d\Omega_F^{m-n-1} + \sum_{i=1}^n \Omega_F^{m-n-1} \wedge da_i\} \tag{5.11}$$

If $m = n$, this result looks nicer, namely

$$\nu_{L/F}(n) = \{a\frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_n}{a_n} /\ a^2 - a \in F^2(a_1,\cdots a_n)'\ \}$$

where $F^2(a_1,\cdots,a_n)' \subset F^2(a_1,\cdots a_n)$ is the subgroup consisting in the elements $\sum_{\varepsilon\neq 0} c_\varepsilon^2 a_1^{\varepsilon_1}\cdots a_n^{\varepsilon_n}$, $\varepsilon = (\varepsilon_1,\cdots\varepsilon_n) \in \{0,1\}^n$.

The corresponding result for bilinear forms is

$$I^m_{L/F} = \left\langle \overline{\psi \ll x_1,\cdots x_n \gg} /\ \psi \in \overline{I}_F^{\,m-n},\ \ x_1,\cdots,x_n \in F^2(a_1,\cdots a_n)^* \right\rangle \tag{5.12}$$

The case $m = n$ is particularly interesting, because

$$\overline{I}^{\,n}_{L/F} = \{\overline{\ll x_1,\cdots x_n \gg} /\ x_i \in F^2(a_1,\cdots a_n)^*\}$$

implies the following corollary

(5.13) Corollary. *Given $x_1,\cdots x_n, y_1,\cdots y_n \in F^2(a_1,\cdots a_n)^*$, then there exist $z_1,\cdots z_n \in F^2(a_1,\cdots a_n)^*$ such that*

$$\ll x_1,\cdots,x_n \gg + \ll y_1,\cdots,y_n \gg \equiv \ll z_1,\cdots z_n \gg \quad \mod\ I_F^{n+1}$$

This is a kind of relative n-linkage property of the subfields $F^2(a_1,\cdots,a_n)$ of F.

Let us now turn our attention to H^{n+1}. The main result of [Ar-Ba]$_3$ is

(5.14) Theorem. *If $\phi = \ll a_1, \cdots a_n \gg$ is anisotropic over F, then*

$$H^{n+1}(F(\phi)/F) = \overline{F\frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_n}{a_n}}$$

The proof of this fact, although elementary, is rather long. For $\overline{\omega} \in H^{n+1}(F(\phi)/F)$ we get an equation $\omega = \wp u + dv$ with $u \in \Omega^n_{F(\phi)}$ and $v \in \Omega^{n-1}_{F(\phi)}$. Writing $F(\phi) = L(y),\ L = F(X_\mu,\ \mu \in S_n),\ y^2 = T = \sum_{\mu \in S_n} a^\mu X_\mu^2,\quad a^\mu = a_1^{\mu(1)} \cdots a_n^{\mu(n)}$, we choose a 2-basis $\mathcal{B} = \{a_i,\ i \in I\}$ of F containing $a_1, \cdots a_n$, so that $\mathcal{B} \cup \{X_\mu,\ \mu \in S_n\}$ is a 2-basis of L and then we fix a 2-basis $\mathcal{B}' = \mathcal{B} \setminus \{a_1\} \cup \{X_\mu,\ \mu \in S_n\} \cup \{y\}$ of $F(\phi)$. We order the elements of this basis such that all $X_\mu > \mathcal{B} \setminus \{a_1\}$ and $y > X_\mu$ for all μ (i.e. y is maximal). Using these choices, and Kato's lemma, one sees that u and v can be chosen free of differentials of the form dX_μ or dy, and moreover that the scalar coefficients of u and v do not contain y in the 2-basis expansion. Thus u and v are defined over $L = F(X_\mu)$. But since $H^{n+1}(F(\phi)/L) = \overline{\Omega^{n-1}_L \wedge dT}$ by (5.5),we have

$$\omega = \wp u + dv + \lambda \wedge dT \tag{5.15}$$

in Ω^n_L, with some $\lambda \in \Omega^{n-1}_L$. Expanding with respect to the 2-basis $\mathcal{B} \cup \{X_\mu,\ \mu \in S_n\}$ and comparing coefficients, one can show that $u,\ v,\ \lambda$ can be taken in $\Omega^n_F \otimes M$ and $\Omega^{n-1}_F \otimes M$ respectively, where $M = F(X^2_\mu,\ \mu \in S_n)$. This is the start for long descent argument which leads to an equation $\omega = \wp u_0 + dv_0 + b da_1 \wedge \cdots \wedge da_n$ whith $b \in F$ and u_0, v_0 defined over F
The corresponding result for quadratic forms is

(5.16) Theorem

$$\overline{I^n W_q}(F(\phi)/F) = \{\overline{\ll a_1, \cdots, a_n, a\,]]}\ /\ a \in F\}$$

As it is shown in [Ar-Ba]$_2$, this result implies the following one. Let $p = \ll a_1, \cdots, a_n, a\,]]$ be now an anisotropic quadratic n-fold Pfister form and let $F(p)$ be the function field of the quadric $\{p(x) = 0\}$. Then

(5.17) Theorem

$$H^{n+1}(F(p)/F) = \{0,\ \bar{p}\}$$

Remark. One may expect that (5.14) generalizes to the following assertion

$$H^{m+1}(F(\phi)/F) = \overline{\Omega^{m-n}_F \wedge \frac{da_1}{a_1} \wedge \cdots \wedge \frac{da_n}{a_n}} \qquad, m \geq n.$$

6 AN APPLICATION:

generic splitting of quadratic forms.
One can develop a generic splitting theory for non singular quadratic forms over a field with $2 = 0$ in the same way as it has been done for the case $2 \neq 0$ in $[\mathrm{Kn}]_{1,2}$, because in the case $2 = 0$ one has:

(i) the analogue of Pfister's subform theorem (s. [Am], $[\mathrm{Ba}]_3$ and [Le])

(ii) The analogue of Knebusch's norm theorem (s. $[\mathrm{Ba}]_2$).

With these tools one defines a generic splitting tower of a non singular quadratic form q over F and obtains a leading form, which is similar to a Pfister form. The degree of this form is called the degree of q. Now define $\mathfrak{I}(n) = \{\bar{q} \in W_q(F) \ / deg\ q \geq n\}$. Then $\mathfrak{I}(n)$ is a $W(F)$-submodule of $W_q(F)$ and one easily sees that $I^n W_q(F) \subseteq \mathfrak{I}(n)$. In $[\text{Ar-Ba}]_3$ it is shown that the equality $\mathfrak{I}(n) = I^n W_q(F)$ for all n (over a field of any characteristic) is equivalent with the statement of theorem (5.17) above for any n. Thus we have

(6.1) THEOREM *For any field F with $2 = 0$, it holds*

$$\mathfrak{I}(n) = I^n W_q(F)$$

REMARK. The corresponding result for (5.17) over fields with $2 \neq 0$ has been announced by Orlov-Vishik-Voevodsky (s. [Pf]).

REFERENCES

[Al] Albert, A. Structure of Algebras. Amer. Math. Soc. Publ. 24, Providence, AMS, (1939).

[A] Arf, C. Untersuchungen über quadratische Formen in Körpern der Charakteristik 2. J. reine ang. Math. 183, 148-167 (1941).

[Am] Amer, M. Quadratische Formen über Funktionenkörpern. Unpublished dissertation, Johannes Gutenberg Universität, Mainz, (1976).

[Ar] Arason, J. Wittring und Galoiscohomologie bei Charakteristik 2 J. reine ang. Math. 307/308, 247-256, (1979).

$[\text{Ar-Ba}]_1$ Aravire, R., Baeza, R. Milnor's K-theory and quadratic forms over fields of characteristic two. Comm. Alg. 20(4) 1087-1107, (1992).

$[\text{Ar-Ba}]_2$ Aravire, R., Baeza, R. A note on generic splitting of quadratic forms. Comm. Alg., 27(7), 3473-3477, (1999).

$[\text{Ar-Ba}]_3$ Aravire, R., Baeza, R. The behaviour of quadratic and differential forms under function field extensions in characteristic two. Preprint 2000. (Submitted) http://www.mathematik.uni-bielefeld.de/LAG/

[Ar-Ba]$_4$ Aravire, R., Baeza, R. Linkage of fields in characteristic two. Preprint 2001. (Submitted) http://www.mathematik.uni-bielefeld.de/LAG/

[Ba]$_1$ Baeza, R. Quadratic Forms over semi-local Rings. LNM 655, Springer-Verlog, (1978).

[Ba]$_2$ Baeza, R. The norm theorem for quadratic forms over a field of characteristic two. Comm. Alg. 18(5), 1337-1348 (1990).

[Ba]$_3$ Baeza, R. Ein Teilformensatz für quadratische Formen in Charakteristik 2. Math. Zeit. 135, 175-184 (1974).

[Ca] Cartier, P. Questions de rationalité des diviseurs en géometrie algebrique. Bull. Soc. Math. France, 86, 177-251, 1958.

[Ka]$_1$ Kato, K. Symmetric bilinear forms, quadratic forms and Milnor's K-theory in characteristic two. Inv. Math. 66, 493 - 510, (1982).

[Ka]$_2$ Kato, K. Galois cohomology of complete discrete valuation fields. Algebraic K-theory, Proc. Oberwolfach, Part II (R. K. Dennis, Ed.) LNM vol. 967, Springer-Verlag, (1982).

[Kn]$_1$ Knebusch, M. Generic splitting of quadratic forms I. Proc. London Math. Soc. (3) 33, 65-93 (1976).

[Kn]$_2$ Knebusch, M. Generic splitting of quadratic forms II. Proc. London Math. Soc. (3) 34, 1-31 (1977).

[Le] Leep, D. The Amer-Brumer theorem over arbitrary fields. Preprint 2001.

[Mi]$_1$ Milnor, J. Algebraic K-theory and quadratic forms. Invent. Math. 9, 318-344 (1970).

[Mi]$_2$ Milnor, J. Symmetric inner products in characteristic 2. Prospects in Mathematics, Ann. of Math. Studies, Princeton Univ. Press. 59-75 (1971).

[Pf] Pfister, A. On the Milnor Conjectures History, Influence, Applications. Jber. d. Dt. Math. Verein. 102, 15-41 (2000).

[Sa] Sah, C-H. Symmetric bilinear forms and quadratic forms. J. of Algebra 20, 144-160 (1972).

[Sch] Scharlau, W. On the history of the algebraic theory of quadratic forms. Proc. of the Dublin Conference on Quadratic Forms and their Applications. Contemporary Mathematics 272 (E. Bayer-Fluckiger, D. Lewis, A. Ranicki, Ed.) A.M.S, 229-259, (2000).

Ricardo Baeza
Instituto de Matematica
Universidad de Talca
Casilla 721
Talca, Chile
rbaeza@inst-mat.utalca.cl

On the Number of Square Classes of a Field of Finite Level

Karim Johannes Becher

Received: May 29, 2001
Revised: October 31, 2001

Communicated by Ulf Rehmann

Abstract. The *level question* is, whether there exists a field F with finite *square class number* $q(F) := |F^\times/F^{\times 2}|$ and finite level $s(F)$ greater than four. While an answer to this question is still not known, one may ask for lower bounds for $q(F)$ when the level is given.

For a nonreal field F of level $s(F) = 2^n$, we consider the filtration of the groups $D_F(2^i)$, $0 \leq i \leq n$, consisting of all the nonzero sums of 2^i squares in F. Developing further ideas of A. Pfister, P. L. Chang and D. Z. Djoković and by the use of combinatorics, we obtain lower bounds for the invariants $\overline{q}_i := |D_F(2^i)/D_F(2^{i-1})|$, for $1 \leq i \leq n$, in terms of $s(F)$. As a consequence, a field with finite level ≥ 8 will have at least 512 square classes. Further we give lower bounds on the cardinalities of the Witt ring and of the 2-torsion part of the Brauer group of such a field.

1 Introduction

Let F be a field. The *level of* F, denoted by $s(F)$, is defined as the least positive integer m such that -1 is a sum of m squares in F whenever such an integer exists and ∞ otherwise. For fields of positive characteristic this invariant can take only the values 1 and 2, depending just on whether -1 is a square in F or not. Fields of level ∞, i.e. in which -1 is not a sum of squares, are called *real fields* and an equivalent condition to $s(F) = \infty$ is the existence of an ordering on F. Fields of finite level are also called *nonreal fields.*

For a long time it has been an open question which values exactly occur as the level of some field. The complete solution to this problem was given by A. Pfister in [10] and it inspired a big part of later advances in the theory of quadratic forms, e.g. the development of the theory of *Pfister forms* and the investigation of isotropy behaviors of quadratic forms under function field extensions.

Pfister proved that the level of a nonreal field is always a power of 2 [10, Satz 4] and further that, if F is any real field (e.g. $\mathbb{Q}$ or $\mathbb{R}$) and $n \geq 0$, then the function field of the projective quadric $X_0^2 + \cdots + X_{2^n}^2 = 0$ over F has level 2^n [10, Satz 5]. These were the first examples of nonreal fields of level greater than 4 and, actually, still no examples of an essentially different kind are known.
In general it remains a difficult problem to determine the level of a given field of characteristic zero. For an overview on what is known about levels of common types of fields we refer to [8, Chap. XI, Section 2]. In the same book T. Y. Lam also mentions the following question [8, p. 333]:

1.1. Level Question. *Does there exist a field F such that $4 < s(F) < \infty$ and such that $F^\times/F^{\times 2}$ is finite?*

Here and in the sequel we denote by $F^\times$ the multiplicative group of F and by $F^{\times 2}$ the subgroup of nonzero squares in F. The quotient $F^\times/F^{\times 2}$ is called the *square class group of F*. We call $q(F) := |F^\times/F^{\times 2}|$ the *square class number of F*. Another subgroup of $F^\times$ of importance is the group of nonzero sums of squares in F, denoted as $\sum F^{\times 2}$.
Further, for any $m \in \mathbb{N}$ we denote by $D_F(m)$ the set of elements of $F^\times$ which can be written as a sum of m squares over F. Pfister has shown that $D_F(m)$ is a group whenever m is a power of 2 [10, Satz 9]. We thus have the following group filtration for $\sum F^{\times 2}$:

$$F^{\times 2} \subsetneq D_F(2) \subsetneq D_F(4) \subsetneq \cdots \subsetneq D_F(2^{i-1}) \subsetneq D_F(2^i) \subsetneq \cdots \subset \sum F^{\times 2}. \quad (1.2)$$

If F is nonreal of level 2^n then we actually have $D_F(2^n+1) = \sum F^{\times 2} = F^\times$.
For $i \geq 1$ we define $\bar{q}_i(F) := |D_F(2^i)/D_F(2^{i-1})|$. Note that the quotients $F^\times/F^{\times 2}$ and $D_F(2^i)/D_F(2^{i-1})$ are 2-elementary abelian groups. So $q(F)$ and $\bar{q}_i(F)$ are each either a power of 2 or ∞.
From (1.2) we see that the inequality

$$q(F) \geq \bar{q}_1(F) \cdots \bar{q}_n(F) \quad (1.3)$$

holds for any $n \geq 1$. We will use this in particular when $s(F) = 2^n$.

While an answer to the level question is still not known, one may look for lower bounds on $|F^\times/F^{\times 2}|$ in terms of $s(F)$.
One approach is to search for lower bounds on the invariants $\bar{q}_i(F)$ and to use then (1.3) to obtain a bound for $q(F)$. Following this idea, A. Pfister obtained in [11, Satz 18.d] the following estimate for a field F of level 2^n:

$$q(F) \geq 2^{\frac{n(n+1)}{2}}. \quad (1.4)$$

His proof (see also [8, p. 325]) actually shows for $1 \leq i \leq n$ that

$$\bar{q}_i(F) \geq 2^{n+1-i}. \quad (1.5)$$

Our standard examples of fields of level $1, 2$ and 4, respectively, are the field of complex numbers $\mathbb{C}$, the finite field $\mathbb{F}_3$ and $\mathbb{Q}_2$, the field of dyadic numbers. These examples show that (1.4) is best possible for $n \leq 2$. For higher n, however, P. L. Chang has improved the bound using combinatorics. In [1] he shows that $q(F) \geq 128$ for a field F of level eight and further that $q(F) \geq 16 \cdot \frac{2^s}{s^2}$ for any nonreal field F of level $s \geq 16$. His approach has been refined by D. Ž. Djoković in [2], leading to the following estimate:

$$q(F) \geq 2 \cdot \sum_{i=1}^{s/2} \frac{1}{s+2-i} \binom{s+1}{i} > \frac{2^s}{s} . \tag{1.6}$$

Their method does not provide any information about the invariants $\bar{q}_i(F)$.

The aim of the present work is to extend this method and to get lower bounds for the invariants $\bar{q}_i(F)$ with respect to $s(F)$ which improve (1.5). The combinatorial aspect is postponed to the two appendices where a certain coloring problem for (hyper-)graphs is considered.

We use common notations and results from quadratic form theory; the standard references are [8] and [12]. (Note that the uncomfortable case of characteristic 2 is implicitly excluded whenever we deal with a field of level greater than 1.) For isometry of quadratic forms we use the symbol $\cong$. For a quadratic form φ over F we denote by $D_F(\varphi)$ the set of nonzero elements of F represented by φ. We sometimes say just "*form*" or "*quadratic form*" to mean "*non-degenerate quadratic form*".

A diagonalized quadratic form over F with coefficients $a_1, \ldots, a_m \in F^\times$ is denoted by $\langle a_1, \ldots, a_m \rangle$. An *m-fold Pfister form* is a quadratic form of the shape $\langle 1, a_1 \rangle \otimes \cdots \otimes \langle 1, a_m \rangle$ and shortly written as $\langle\langle a_1, \ldots, a_m \rangle\rangle$; its dimension is 2^m. A *neighbor* of an m-fold Pfister form π is a quadratic form φ which is similar to a subform of π and of dimension greater than 2^{n-1}. We know that in this situation φ is isotropic if and only of π is hyperbolic.

By $W(F)$ we denote the Witt ring of F, further by $\mathrm{Br}(F)$ the Brauer group and by $\mathrm{Br}_2(F)$ its 2-torsion part. In (3.1), (5.4) and (5.5) we shall use Milnor K-theory. For definitions and properties of the Milnor ring k_*F and its homgenous components k_mF ($m \geq 0$) we refer to [9] and [3]. However, we use the notation $\{a_1, \ldots, a_m\}$ instead of $\ell(a_1) \cdots \ell(a_m)$ for a symbol in k_mF. We recall that this symbol is zero in k_mF if and only if the corresponding m-fold Pfister form $\langle\langle -a_1, \ldots, -a_m \rangle\rangle$ over F is hyperbolic (see [3, Main Theorem 3.2]). In particular, $s(F) = 2^n$ is equivalent to $\{-1\}^n \neq 0$ and $\{-1\}^{n+1} = 0$ in k_*F. Everywhere else in the text, $\{x_1, \ldots, x_n\}$ stands simply for the set of elements $x_1, \ldots, x_1$.

Acknowledgments

This work is part of the author's Ph.D. thesis research, carried out at the *Université de Franche-Comté* under the supervision of Eva Bayer-Fluckiger and Detlev W. Hoffmann. The author expresses his gratitude to his supervisors as well as to Ján Mináč and to David B. Leep for encouragement and stimulating discussions on the subject.

2 Sums of squares in fields

Let F be a field. For an element $x \in F$ we define its *length (over F)* to be the least positive integer m such that x can be written as a sum of m nonzero squares over F if such an integer exists and ∞ otherwise (i.e. if x is not a nontrivial sum of nonzero squares over F). We denote this value in $\mathbb{N} \cup \{\infty\}$ by $\ell_F(x)$, or just by $\ell(x)$ whenever the context makes clear over which field F we are working. Obviously $\ell_F(x)$ depends on x only up to multiplication by a nonzero square in F; in other words, $\ell_F(x)$ is an invariant of the square class $xF^{\times 2}$ whenever $x \neq 0$.
For $m \geq 1$, $D_F(m)$ is by definition the set $\{x \in F^\times \mid \ell(x) \leq m\}$. Our investigation into lengths of field elements is based on the following famous result [10, Satz 2]:

2.1. Theorem (Pfister). *For any $i \geq 0$, $D_F(2^i)$ is a subgroup of $F^\times$.*

A simple proof within the theory of Pfister forms can be found in [12, 4.4.1. Lemma]. As a consequence of this theorem one gets an inequality linking the lengths of two elements to the length of their product. We include a proof of this result, which is [10, Satz 3].

2.2. Lemma. *For any $x, y \in F$ we have the inequalities $\ell(x+y) \leq \ell(x) + \ell(y)$ and $\ell(xy) \leq \ell(x) + \ell(y) - 1$.*

Proof: The first inequality is obvious from the definition of the length.
The second inequality is trivial if xy is zero or if x or y is not a sum of squares. So we may suppose that both x and y are nonzero sums of squares in F. Let then r be the least nonnegative integer such that $x, y \in D_F(2^r)$. We will prove $\ell(xy) < \ell(x) + \ell(y)$ by induction on r. If $r = 0$ then x, y and xy are squares in F and the inequality is clear. Suppose now that $r > 0$. Since $D_F(2^r)$ is a group we know that $\ell(xy) \leq 2^r$. So the inequality is clear if $2^r < \ell(x) + \ell(y)$. Otherwise, we may suppose that $\ell(y) \leq 2^{r-1}$. By the choice of r we then have $2^{r-1} < \ell(x) \leq 2^r$ and may therefore write $x = a + z$ with $a, z \in F^\times$ such that $\ell(a) = 2^{r-1}$ and $\ell(z) = \ell(x) - 2^{r-1} \leq 2^{r-1}$. By the induction hypothesis we have $\ell(zy) < \ell(y) + \ell(z)$. As $D_F(2^{r-1})$ is a group we have $\ell(ay) \leq 2^{r-1}$. Since $xy = ay + zy$, using the first inequality of the statement we obtain finally $\ell(xy) \leq \ell(ay) + \ell(zy) < 2^{r-1} + \ell(y) + \ell(z) = \ell(x) + \ell(y)$. □

According to the definition we gave in the introduction, the *level* of F is the length of -1 in F. We may also conclude that $\ell_F(0) = s(F) + 1$. Therefore, from any of the inequalities of the lemma we obtain immediately:

2.3. Corollary. *For any $x \in F$ we have $\ell(x) + \ell(-x) \geq s(F) + 1$.* □

2.4. Corollary. *Let $a_1, \dots, a_m \in F^\times$. If the quadratic form $\langle a_1, \dots, a_m \rangle$ over F represents the element $x \in F$ nontrivially then $\ell(a_1) + \cdots + \ell(a_m) \geq \ell(x)$.*

Proof: If the form $\langle a_1, \ldots, a_m\rangle$ represents $x \in F$ nontrivially, this means that there are $x_1, \ldots, x_m \in F$, not all zero, such that $a_1x_1^2 + \cdots + a_mx_m^2 = x$. We may suppose that x_i is nonzero for $1 \le i \le m'$ and zero for $m' < i \le m$. From the first inequality of the lemma we obtain $\ell(x) \le \ell(a_1x_1^2) + \cdots + \ell(a_{m'}x_{m'}^2) = \ell(a_1) + \cdots + \ell(a_{m'})$. $\square$

For $i \ge 0$, we say that the elements $a_1, \ldots, a_m \in F^\times$ are *independent modulo* $D_F(2^i)$ if in $F^\times/D_F(2^i)$, considered as an $\mathbb{F}_2$-vectorspace, the classes represented by $a_1, \ldots, a_m$ are $\mathbb{F}_2$-linear independent.

2.5. Proposition. *For $i \ge 2$, let $a, b \in D_F(3{\cdot}2^{i-2})\backslash D_F(2^{i-1})$ and $c \in D_F(2^i)$ such that $\ell(a+b+c) > 2^{i+1}$. Then the elements a, b and c of $D_F(2^i)$ are independent modulo $D_F(2^{i-1})$.*

Proof: We have to show that $a, b, c, ab, ac, bc, abc \notin D_F(2^{i-1})$. For a and b this is already given. We put $x := a+b+c$. Each of the quadratic forms $\langle a, b, c\rangle$, $\langle a, b, abc\rangle$, $\langle 1, ab, ac\rangle$ and $\langle ac, bc, 1\rangle$ over F represents one of the elements x, abx, ax and cx and neither of these elements lies in the group $D_F(2^{i+1})$. We obtain from (2.4) that each of the numbers $\ell(a) + \ell(b) + \ell(c)$, $\ell(a) + \ell(b) + \ell(abc)$, $1 + \ell(ab) + \ell(ac)$ and $\ell(ac) + \ell(bc) + 1$ is greater than 2^{i+1}. Since $\ell(a) + \ell(b) \le 3 \cdot 2^{i-1}$ and $ab, ac, bc \in D_F(2^i)$ we obtain $\ell(c), \ell(abc) \ge 2^{i-1}$ and further $\ell(ab) = \ell(ac) = \ell(bc) = 2^i$. $\square$

For the rest of this section we fix a sum of squares

$$x = x_1^2 + \cdots + x_l^2 \tag{2.6}$$

with $x_1, \ldots, x_l \in F^\times$, $x \in F$ and $l = \ell_F(x)$. For a subset $I \subset \{1, \ldots, l\}$ we denote $x_I := \sum_{i\in I} x_i^2$. If I is not empty then we have $\ell(x_I) = |I|$.
For a real number z we denote by $\lceil z\rceil$ the least integer $\ge z$.

2.7. Theorem. *Let I and J be nonempty proper subsets of $\{1, \ldots, l\}$. Let r be a nonnegative integer such that $x_Ix_J \in D_F(2^r)$. Then the following hold:*

(*i*) $\left\lceil \frac{|I|}{2^r} \right\rceil = \left\lceil \frac{|J|}{2^r} \right\rceil$, *in particular* $|\,|I| - |J|\,| < 2^r$,

(*ii*) $|I \setminus J|\ ,\ |J \setminus I| \ \le\ 2\ \ell(x_Ix_J) - 1 \ <\ 2^{r+1}$,

(*iii*) $|I \cup J| - |I \cap J| \ \le\ 2^{r+1} + \ell(x_Ix_J) - 1 \ \le\ 3 \cdot 2^r - 1$.

Proof: The hypothesis implies that x_I and x_J are nonzero elements of F. We set $m := \ell(x_Ix_J)$ and $a := \frac{x_J}{x_I}$. Then $\ell(a) = m \le 2^r$.
If ν is an integer such that $|I| \le \nu 2^r$ then we can write x_I as a sum of $\le \nu$ elements of $D_F(2^r)$. As $D_F(2^r)$ is a group, $x_J = ax_I$ can also be written as a sum of $\le \nu$ elements of $D_F(2^r)$ which means that $|J| = \ell(x_J) \le \nu 2^r$. By symmetry we obtain for any $\nu \in \mathbb{N}$ that $|I| \le \nu 2^r$ if and only if $|J| \le \nu 2^r$. This shows (*i*).

We compute $x_{I\cup J} = x_{I\setminus J} + x_J = (1+a)x_{I\setminus J} + ax_{I\cap J}$ and then substitute $y := (1+a)x_{I\setminus J}$ and $z := ax_{I\cap J}$ to have $x_{I\cup J} = y + z$.
If $y \neq 0$ then we have $\ell(y) \leq m + |I \setminus J|$ by (2.2), but also $\ell(y) \leq 2^{r+1}$ since $D_F(2^{r+1})$ is a group. If $z \neq 0$ then (2.2) yields $\ell(z) \leq m + |I \cap J| - 1$. Therefore, if at least one of y and z is nonzero then we obtain the inequalities $\ell(y+z) \leq |I| + 2m - 1$ and $\ell(y+z) \leq 2^{r+1} + m + |I \cap J| - 1$. Both inequalities remain valid in the case $y = z = 0$, since then necessarily $a = -1$, whence $\ell(y+z) = \ell(0) = m+1$. As $|I \cup J| = \ell(y+z)$ we obtain (*ii*) by symmetry from the first and (*iii*) from the second inequality. □

For $m = 1$ this leads to an observation made in the proof of [1, Theorem 1]:

2.8. Corollary (Chang). *Let I and J be as in the theorem. If x_I and x_J lie in the same square class then both sets have the same cardinality and differ by at most one element.* □

2.9. Corollary. *Let I and J be as in the theorem with $|I| = |J| = 2^i$, $i \geq 2$. If x_I and x_J represent the same class modulo $D_F(2^{i-1})$ then $|I \cap J| \geq 2^{i-2} + 1$.*

Proof: If x_I and x_J lie in the same class modulo $D_F(2^{i-1})$ then $\ell(x_I x_J) \leq 2^{i-1}$. Applying part (*iii*) of the theorem for $r = i - 1$ we obtain $|I \cup J| - |I \cap J| \leq 3 \cdot 2^{i-1} - 1$. But our hypothesis here gives $|I \cup J| = 2 \cdot 2^i - |I \cap J|$. This together implies $|I \cap J| > 2^{i-2}$. □

3 The invariants $\bar{q}_i$

For a nonreal field F of level 2^n we are going to study the invariants $\bar{q}_i(F) = |D_F(2^i)/D_F(2^{i-1})|$ for $1 \leq i \leq n$. In particular, we are interested to know whether Pfister's bounds (1.5) can be improved.

First we note that the bound $\bar{q}_n(F) \geq 2$, obtained from (1.5) for $i = n$, just takes into account that -1 represents a nontrivial class in the group $D_F(2^n)/D_F(2^{n-1})$. In spite of the simple argument, this bound is optimal for every $n \geq 1$. More precisely, for any $n \geq 1$ there is a field F of level 2^n such that $F^\times = D_F(2^{n-1}) \cup -D_F(2^{n-1})$. The construction of such an example will be included in a forthcoming paper of the author.
We now turn to consider $\bar{q}_{n-1}(F)$. For $i = n-1$, (1.5) gives $\bar{q}_{n-1}(F) \geq 4$. The example $F = \mathbb{Q}_2$ shows that this bound is optimal for $n = 2$.

3.1. Theorem. *Let F be a field of level 2^n with $n \geq 3$. Then $\bar{q}_{n-1}(F) \geq 16$.*

Proof: Since $\ell(0) = 2^n + 1$ and $n \geq 3$, we may choose elements $a_1, a_2, a_3 \in F^\times$ such that $a_1 + a_2 + a_3 = 0$ and $2^{n-2} + 1 \leq \ell(a_i) \leq 3 \cdot 2^{n-3}$ for $i = 1, 2, 3$. Then by (2.5), a_1, a_2 and a_3 are independent modulo $D_F(2^{n-2})$. Let H be the subgroup of $D_F(2^{n-1})$ generated by $D_F(2^{n-2})$ and the elements a_1, a_2 and a_3. Since $|H/D_F(2^{n-2})| = 8$ it remains to show that $H \neq D_F(2^{n-1})$.

To this aim, we will calculate in the Milnor ring k_*F. For $i = 1,2,3$ we fix the symbols $\beta_i := \{a_1a_2a_3, a_i\}$ and $\gamma_i := \{-a_1a_2a_3, -a_i\}$ in k_2F. Let ε denote the element $\{-1\}$ in k_1F. Since $s(F) = 2^n$ we have $\varepsilon^n \neq 0$. As $a_1, a_2, a_3 \in D_F(2^{n-1})$ we observe that $\beta_1+\beta_2+\beta_3 = \{-1, a_1a_2a_3\}$ is annihilated by ε^{n-2} and that $\varepsilon^{n-2}(\beta_i+\gamma_i) = \varepsilon^{n-2}(\{a_1a_2a_3,-1\}+\{-1,a_i\}+\{-1,-1\}) = \varepsilon^n$ for $i = 1,2,3$.

If $\varepsilon^{n-2}\beta_i \neq 0$ in k_nF for some i then by the above relations we may suppose that $\varepsilon^{n-2}\beta_i \neq 0$ for $i = 1,2$ and $\varepsilon^{n-2}\beta_3 \neq \varepsilon^n$, i.e. $\varepsilon^{n-2}\gamma_3 \neq 0$. Using that $a_1 + a_2 + a_3 = 0$ we compute $\{-a_2, -a_3\} = \{a_1, -a_2a_3\} = \beta_1$ and equally $\{-a_1, -a_3\} = \beta_2$. Since none of β_1, β_2 and γ_3 is annihilated by ε^{n-2}, the symbols $\varepsilon^{n-2}\{-a_2, -a_3\}$, $\varepsilon^{n-2}\{-a_1, -a_3\}$ and $\varepsilon^{n-2}\{-a_1a_2, -a_3\}$ in k_nF are all nonzero. Therefore the Pfister forms $2^{n-2}\times\langle\langle a_2, a_3\rangle\rangle$, $2^{n-2}\times\langle\langle a_1, a_3\rangle\rangle$ and $2^{n-2}\times\langle\langle a_1a_2, a_3\rangle\rangle$ are anisotropic. Further, $2^{n-2}\times\langle\langle 1, a_3\rangle\rangle \cong 2^n\times\langle 1\rangle$ is anisotropic since $s(F) = 2^n$. This shows that $-1, -a_1, -a_2, -a_1a_2 \notin D_F(2^{n-2}\times\langle\langle a_3\rangle\rangle)$. As the group $D_F(2^{n-2}\times\langle\langle a_3\rangle\rangle)$ contains the subgroup $D_F(2^{n-2})$ and the element a_3 we conclude that $D_F(2^{n-2}\times\langle\langle a_3\rangle\rangle) \cap -H = \emptyset$. On the other hand, since $\ell(-a_3) \leq \ell(a_1)+\ell(a_2) \leq 3\cdot 2^{n-2}$ we can write $-a_3 = x+y$ with $x \in D_F(2^{n-1})$, $y \in D_F(2^{n-2})$ and obtain $-x = y + a_3 \in D_F(2^{n-2}\times\langle\langle a_3\rangle\rangle) \cap -D_F(2^{n-1})$.

Now we study the case where $\varepsilon^{n-2}\beta_i = 0$ for $i = 1,2,3$. As $\varepsilon^{n-2}\beta_i = \varepsilon^{n-2}\{-a_1a_2a_3, a_i\}$, this means that the Pfister form $2^{n-2} \times \langle\langle a_1a_2a_3, -a_i\rangle\rangle$ is hyperbolic for $i = 1,2,3$. We conclude that $H \subset D_F(2^{n-2} \times \langle\langle a_1a_2a_3\rangle\rangle)$. As the Pfister form $2^{n-1} \times \langle\langle a_1a_2a_3\rangle\rangle \cong 2^n \times \langle 1\rangle$ is anisotropic we have $-1 \notin D_F(2^{n-2} \times \langle\langle a_1a_2a_3\rangle\rangle)$ and therefore $D_F(2^{n-2} \times \langle\langle a_1a_2a_3\rangle\rangle) \cap -H = \emptyset$. Since $-a_1a_2a_3 = a_1^2a_2 + a_2^2a_1$ we have $\ell(-a_1a_2a_3) \leq \ell(a_2) + \ell(a_1) \leq 3 \cdot 2^{n-2}$ and may therefore write $-a_1a_2a_3 = x+y$ with $x \in D_F(2^{n-1})$ and $y \in D_F(2^{n-2})$ to obtain this time $-x = y + a_1a_2a_3 \in D_F(2^{n-2} \times \langle\langle a_1a_2a_3\rangle\rangle) \cap -D_F(2^{n-1})$.

In both cases we have found an element $x \in D_F(2^{n-1}) \setminus H$. $\square$

While the lower bound on $\bar{q}_{n-1}$ of the last theorem is based upon several algebraic arguments, the improvement (with respect to (1.5)) for the lower bounds on $\bar{q}_i(F)$ for $2 \leq i \leq n-2$ which we present now, is obtained by combinatorial reasoning, developed in appendix A.

For integers $0 \leq k \leq l$ we denote by $\mathcal{P}_k^l$ the set of subsets of $\{1,\ldots,l\}$ with exactly k elements.

3.2. Theorem. *Let F be a field of level 2^n. Then*

$$\bar{q}_i(F) \;\geq\; \begin{cases} 2^7 & \text{for } \; i = n-2 \geq 3\,, \\ 2^{(n-i)(2^{n-i}+1)+1} & \text{for } \; \frac{n+1}{2} < i \leq n-3\,, \\ 2^{(n-i)(2^{i-2}+1)+1} & \text{for } \; 2 \leq i \leq \frac{n+1}{2}\,. \end{cases}$$

Proof: We fix elements $x_1, \ldots, x_{2^n} \in F^\times$ such that $x_1^2 + \cdots + x_{2^n}^2 = -1$. For a subset $J \subset \{1,\ldots,2^n\}$ we denote $x_J := \sum_{j\in J} x_j^2$.

Let $2 \leq i \leq \frac{n+1}{2}$. We consider the map $f : \mathcal{P}_{2^i}^{2^n} \longrightarrow D_F(2^i)/D_F(2^{i-1})$ which sends a 2^i-subset $J \subset \{1,\ldots,2^n\}$ to the class $x_JD_F(2^{i-1})$. By (2.9), if $J_1, J_2 \in$

$\mathcal{P}_{2^i}^{2^n}$ are such that $f(J_1) = f(J_2)$ then $|J_1 \cap J_2| \geq 2^{i-2}+1$. Therefore (A.8) in appendix A shows $|D_F(2^i)/D_F(2^{i-1})| \geq |Im(f)| > 2^r$ for $r := (n-i)(2^{i-2}+1)$. Since $D_F(2^i)/D_F(2^{i-1})$ is a 2-elementary abelian group it must then have at least 2^{r+1} elements. This establishes the third case in the statement.
In the remaining cases we cannot apply (A.8) directly for i and $m := n$. In the case $\frac{n+1}{2} < i \leq n-3$ we have $n \geq 8$ and $i \geq 5$ and define $n' := 2(n-i+1)$ and $i' := n-i+2 = \frac{n'}{2}+1$. In the case $i = n-2$ and $n \geq 5$ we set instead $n' := 5$ and $i' := 3 = \frac{n'+1}{2}$. Note that in both cases $n'-i' = n-i$.
For $1 \leq \nu \leq 2^{n'}$ let $J_\nu := \{(\nu-1)\cdot 2^{n-n'}+1, \ldots, \nu\cdot 2^{n-n'}\}$ and $y_\nu := x_{J_\nu}$. This yields $y_1 + \cdots + y_{2^{n'}} = -1$ and $\ell(y_\nu) = |J_\nu| = 2^{n-n'}$ for $1 \leq \nu \leq 2^{n'}$. Now we consider the map $f' : \mathcal{P}_{2^{i'}}^{2^{n'}} \longrightarrow D_F(2^i)/D_F(2^{i-1})$ which sends a $2^{i'}$-subset $N \subset \{1, \ldots, 2^{n'}\}$ to the class $(\sum_{\nu\in N} y_\nu) D_F(2^{i-1})$.
Suppose that $f'(N_1) = f'(N_2)$ for $N_1, N_2 \in \mathcal{P}_{2^{i'}}^{2^{n'}}$. For $k = 1, 2$ let $I_k := \bigcup_{\nu\in N_k} J_\nu \in \mathcal{P}_{2^i}^{2^n}$. Since by hypothesis $\sum_{\nu\in N_1} y_\nu = x_{I_1}$ and $\sum_{\nu\in N_2} y_\nu = x_{I_2}$ lie in the same class of $D_F(2^i)/D_F(2^{i-1})$, (2.9) shows that $|I_1 \cap I_2| \geq 2^{i-2}+1$ and it follows that $|N_1 \cap N_2| \geq 2^{i-2-(n-n')}+1 = 2^{i'-2}+1$.
Having established this intersection property of f', we obtain from (A.8) that $|D_F(2^i)/D_F(2^{i-1})| \geq |Im(f')| > 2^{r'}$ holds for $r' := (n'-i')(2^{i'-2}+1)$. As before, we conclude that $|D_F(2^i)/D_F(2^{i-1})| \geq 2^{r'+1}$. This finishes the proof since $r' = 6$ in case $i = n-2$ and $r' = (n-i)(2^{n-i}+1)$ otherwise. □

4 Nonreal fields with $\bar{q}_1$ equal to the level

From (1.5) we know that $\bar{q}_1(F) \geq s(F)$ holds for any nonreal field F. This bound is optimal for fields of level $1, 2$ and 4 as the standard examples show (see introduction). For nonreal fields of higher level, however, there is still no known example where $\bar{q}_1(F) < \infty$.
We show that $\bar{q}_1(F) = s(F) < \infty$ is a rather strong condition, with several consequences on the quadratic form structure of F. In particular, for $s(F) \geq 8$ it implies that $\bar{q}_2(F) \geq \frac{s(F)^2}{2}$ (4.9).

Let ξ be an element of length $l \geq 3$ of F. We fix a representation of ξ as a sum of l squares

$$\xi = x_1^2 + \cdots + x_l^2 \tag{4.1}$$

with $x_1, \ldots, x_l \in F^\times$. Let $f : \mathcal{P}_2^l \to D_F(2)/F^{\times 2}$ be the function which sends a (nonordered) pair of distinct $i, j \leq l$ to the square class of $x_i^2 + x_j^2$. Considering the elements of $D_F(2)/F^{\times 2}$ as a set of colors, we can interprete f as an edge-coloring of a complete graph in l vertices $v_1, \ldots, v_l$. We denote this graph together with its edge-coloring f by $\mathcal{G}$. If in this graph two edges $[v_i, v_j]$ and $[v_{i'}, v_{j'}]$ are of the same color (with $\{i,j\}, \{i',j'\} \in \mathcal{P}_2^l$) this means that $x_i^2 + x_j^2$

and $x_{i'}^2 + x_{j'}^2$ lie in the same square class of F, which by (2.8) implies that the sets $\{i, j\}$ and $\{i', j'\}$ intersect. In other words, two edges of the same color in $\mathcal{G}$ need to have a vertex in common, i.e. $\mathcal{G}$ is a *CC-graph* in the terminology of appendix B.
We get from (B.1) that at least $l - 2$ colors appear in $\mathcal{G}$. Furthermore, since $x_1^2 + \cdots + x_l^2$ is of length l, no sum $x_i^2 + x_j^2$ with $i \neq j$ can be a square. This gives a proof of [13, Theorem 1]:

4.2. Proposition (Tort). *In* (4.1), *the partial sums* $x_i^2 + x_j^2$ *with* $1 \leq i < j \leq l$ *represent at least* $l - 2$ *different nontrivial classes of* $D_F(2)/F^{\times 2}$. □

Let now F be a nonreal field of level $s = 2^n$. We then can choose $\xi := 0$, which is of length $s + 1$ over F, and write (4.1) as

$$0 = x_1^2 + \cdots + x_{s+1}^2 \,. \tag{4.3}$$

By the above proposition the partial sums $x_i^2 + x_j^2$ (with $1 \leq i < j \leq s + 1$) represent at least $s - 1$ nontrivial classes of $D_F(2)/F^{\times 2}$. This shows:

4.4. Corollary. *Let* F *be a nonreal field of level* s. *Then* $\bar{q}_1(F) \geq s$. *Moreover, if* $\bar{q}_1(F) = s$ *then, given any representation* (4.3) *of zero as a sum of* $s+1$ *nonzero squares over* F, *every nontrivial class of* $D_F(2)/F^{\times 2}$ *is represented by a partial sum* $x_i^2 + x_j^2$ *with* $1 \leq i < j \leq s+1$. □

Given a subgroup $G \subset F^\times/F^{\times 2}$ of finite order 2^m we may choose an irredundant set of representatives $a_1, \ldots, a_{2^m} \subset F^\times$ of the square classes in G and define the quadratic form $\pi_G := \langle a_1, \ldots, a_{2^m} \rangle$. Up to isometry, this form does only depend on G and not on the particular choice of the a_i. If we choose the a_i such that $a_1, \ldots a_m$ are independent modulo $F^{\times 2}$ then π_G is equal to $\langle\langle a_1, \ldots, a_m \rangle\rangle$, hence π_G is an m-fold Pfister form. If $\bar{q}_1(F)$ is finite we write $\pi_{D(2)}$ for π_G with $G := D_F(2)/F^{\times 2}$.

4.5. Proposition. *Let* F *be a nonreal field with* $s(F) > 1$ *and* $\bar{q}_1(F) < \infty$. *Then* $\pi_{D(2)}$ *is hyperbolic.*

Proof: Let $s := s(F)$. Given a representation (4.3) of zero as sum of $s + 1$ squares over F we define $a_i := x_{2i-1}^2 + x_{2i}^2$ for $1 \leq i \leq s/2$. By (2.8) the a_i lie in distinct nontrivial square classes. Since $a_1 + \cdots + a_{s/2} + x_{s+1}^2 = 0$ the form $\langle 1, a_1, \ldots, a_{s/2} \rangle$ is isotropic. On the other hand, this is a subform of the Pfister form $\pi_{D(2)}$, which then must be hyperbolic. □

4.6. Lemma. *Let* H *be a subgroup of* $F^\times$ *containing* $F^{\times 2}$ *such that* $H/F^{\times 2}$ *is of order* 2^m *with* $m \geq 2$. *If* $a, b, c, d \in H$, *lie in distinct square classes then there are* $a_3, \ldots, a_m \in H$ *such that* $\pi_H = \langle a, b, c, d \rangle \otimes \langle\langle a_3, \ldots, a_m \rangle\rangle$.

Proof: It is easy to verify that, given four distinct elements t, u, v, w in a 2-elementary abelian group G there exists a subgroup K of index 4 in G such that t, u, v, w represent the four classes of G/K.
We apply this fact to the square classes $aF^{\times 2}, bF^{\times 2}, cF^{\times 2}$ and $dF^{\times 2}$ in $G := H/F^{\times 2}$. A subgroup K with the stated property must have order 2^{m-2}. We choose elements $a_3, \ldots, a_m \in F^\times$ such that their square classes form an $\mathbb{F}_2$-basis of K. The rest is clear. $\square$

4.7. PROPOSITION. *Let F be a field with $\bar{q}_1(F) = s(F) = 2^n$, $n \geq 2$, and let a, b, c, d be elements of $D_F(2)$ which lie in distinct square classes.*

(a) If $a \notin F^{\times 2}$ then $D_F(\langle 1,1\rangle) \cap D_F(\langle 1,a\rangle) = \{1,a\}F^{\times 2}$.

(b) If $x \in D_F(\langle 1,a\rangle) \cap D_F(\langle 1,b\rangle) \cap D_F(\langle 1,c\rangle)$ then $\ell(-x) = 2^n$.

(c) $D_F(\langle a,b\rangle) \cap D_F(\langle c,d\rangle) = \emptyset$.

(d) If $n \geq 3$ then $D_F(\langle a,b\rangle) \cap D_F(\langle a,c\rangle) \cap D_F(\langle b,c\rangle) = \emptyset$.

(e) If $x \in D_F(\langle 1,a\rangle) \cap D_F(\langle 1,b\rangle)$ then $\ell(cx) = 4$ or $\ell(-x) \geq 2^n - 1$.

Proof: (a) Given a and b lying in distinct nontrivial classes of $D_F(2)/F^{\times 2}$ we may choose $a_3, \ldots, a_{2^{n-1}} \in D_F(2)$ such that $\varphi := \langle 1, a, b, a_3 \ldots, a_{2^{n-1}}\rangle$ is a neighbor of the Pfister form $\pi_{D(2)}$ which is hyperbolic by the last proposition. So φ is isotropic. Now $b \in D_F(\langle 1,a\rangle)$ would imply that φ is isometric to $\langle 1, 1, ab, a_3 \ldots, a_{2^{n-1}}\rangle$ which is a subform of $2^n \times \langle 1\rangle$. This is impossible since the latter form is anisotropic by the hypothesis that $s(F) = 2^n$.
(b) Let $x \in D_F(\langle 1,a\rangle) \cap D_F(\langle 1,b\rangle) \cap D_F(\langle 1,c\rangle)$ where $a, b, c \in D_F(2)$ are distinct modulo squares. Then clearly $\ell(x) \leq 3$ and we have also $x \in D_F(\langle 1, abc\rangle)$ (with $-a, -b$ and $-c$ also $-abc$ lies in $D_F(\langle 1, -x\rangle)$). It follows from (a) that $\ell(x) \neq 2$. If x is a square then $\ell(-x) = \ell(-1) = 2^n$. Otherwise we must have $\ell(x) = 3$. Then none of a, b, c, abc can be a square. Further $\ell(-x) \geq 2^n - 2$ by (2.3). Thus (4.2) shows that, in a representation of $-x$ as sum of $\ell(-x)$ squares over F, the partial sums of length two lie in at least $2^n - 4$ distinct nontrivial square classes. As $|D_F(2)/F^{\times 2}| = 2^n$ by hypothesis, at least one of these square classes must also be represented by one of a, b, c or abc. Without loss of generality we may suppose that $-x = y + at^2$ with $\ell(y) = \ell(-x) - 2$. Writing $x = u^2 + av^2$ yields $0 = x - x = y + u^2 + a(t^2 + v^2)$. Thus $2^n + 1 \leq \ell(y) + 3$ and $2^n \leq \ell(y) + 2 = \ell(-x)$. Then $-x = (-1) \cdot x \in D_F(2^n)$ implies $\ell(-x) = 2^n$.
(c) By the above lemma there are $a_3, \ldots, a_n \in D_F(2)$ such that $\pi_{D(2)}$ is equal to $\langle a, b, c, d\rangle \otimes \langle\langle a_3, \ldots, a_n\rangle\rangle$.
Suppose now that there exists an $x \in D_F(\langle a,b\rangle) \cap D_F(\langle c,d\rangle)$. Then $\langle a,b,c,d\rangle \cong \langle x, abx, x, cdx\rangle$, which is similar to $\langle 1,1,1,abcd\rangle$. Hence $\pi_{D(2)}$ is similar to $\langle 1,1,1,abcd\rangle \otimes \langle\langle a_3, \ldots, a_n\rangle\rangle \cong 2^{n-1} \times \langle 1\rangle \perp \langle\langle abcd, a_3, \ldots, a_n\rangle\rangle$. It follows that the form $(2^{n-1} + 1) \times \langle 1\rangle$ is a Pfister neighbor of $\pi_{D(2)}$, hence isotropic since $\pi_{D(2)}$ is hyperbolic. This is a contradiction to $s(F) = 2^n$.

(d) After multiplying by a in the statement we may suppose that $a = 1$. Suppose that there exists $x \in D_F(\langle 1, b\rangle) \cap D_F(\langle 1, c\rangle) \cap D_F(\langle b, c\rangle)$. It follows $-b, -c \in D_F(\langle 1, -x\rangle)$, thus $bc \in D_F(\langle 1, -x\rangle) \cap D_F(\langle 1, 1\rangle) \subset D_F(\langle 1, x\rangle)$. Therefore we have $\langle 1, b, c, bc\rangle \cong \langle 1, x, bcx, bc\rangle \cong \langle bc, bcx, bcx, bc\rangle$, whence $\langle 1, b, c, bc\rangle \cong \langle 1, 1, x, x\rangle$. Next we choose $a_3, \ldots, a_n \in D_F(2)$ such that $\pi_{D(2)} \cong \langle 1, b, c, bc\rangle \otimes \langle\langle a_3, \ldots, a_n\rangle\rangle$ and obtain $\pi_{D(2)} \cong \langle\langle 1, x, a_3, \ldots, a_n\rangle\rangle \cong 2^{n-1} \times \langle\langle x\rangle\rangle \cong 2^n \times \langle 1\rangle$, since $a_3, \ldots, a_n \in D_F(2)$, $n \geq 3$ and $x \in D_F(4)$. This is contradictory since $\pi_{D(2)}$ is hyperbolic but $s(F) = 2^n$.
(e) Let $x \in D_F(\langle 1, a\rangle) \cap D_F(\langle 1, b\rangle)$. Then, certainly, x and cx belong to $D_F(4)$. If $\ell(cx) \leq 2$ then $\ell(x) \leq 2$ and (2.3) yields $\ell(-x) \geq 2^n - 1$. Suppose now $\ell(cx) = 3$ and write $cx = e + t^2$ with $t \in F^\times$ and $e \in D_F(2)$. We have $cx \in D_F(\langle c, ac\rangle) \cap D_F(\langle c, bc\rangle) \cap D_F(\langle 1, e\rangle)$. Since $1, c, ac$ and bc represent distinct square classes, we conclude with (c) that e and c lie in the same square class. Therefore $x \in D_F(\langle 1, a\rangle) \cap D_F(\langle 1, b\rangle) \cap D_F(\langle 1, c\rangle)$, which by (b) implies $\ell(-x) = 2^n$. □

4.8. Theorem. *Let F be a nonreal field of level s, equal to $\bar{q}_1(F)$. Any representation* (4.3) *of zero as a nontrivial sum of $s+1$ squares over F may be reordered in such way that the following holds: for $\{i, j\}, \{i', j'\} \in \mathcal{P}_2^{s+1}$ the partial sums $x_i^2 + x_j^2$ and $x_{i'}^2 + x_{j'}^2$ lie in the same square class if and only if* $\max\{i, j, 3\} = \max\{i', j', 3\}$.

Proof: Let $\mathcal{G}$ be a complete graph in $s+1$ vertices $v_1, \ldots, v_{s+1}$ and with the edge-coloring given by $f : \mathcal{P}_2^{s+1} \to D_F(2)/F^{\times 2}, \{i, j\} \mapsto (x_i^2 + x_j^2)F^{\times 2}$ (see at the beginning of this section). We know from (4.4) that exactly $s-1$ colors appear in $\mathcal{G}$. Further, $\mathcal{G}$ does not contain any triangle with three different colors; indeed, such a triangle would correspond to a partial sum of three squares $x := x_i^2 + x_j^2 + x_k^2$ with $1 \leq i < j < k \leq s+1$ where $a := x_i^2 + x_j^2$, $b := x_i^2 + x_k^2$ and $c := x_j^2 + x_k^2$ lie in three distinct square classes which is impossible by part (b) of the last proposition since $\ell(-x) = s - 2$. Therefore by (B.3), $\mathcal{G}$ is a total CC-graph.
Since $\mathcal{G}$ has precisely $(s+1) - 2$ colors we obtain from the definition of a total CC-graph in appendix B and the subsequent remarks: the vertices in $\mathcal{G}$ (and at the same time the x_i) may be renumbered in such way that for $\{i, j\} \in \mathcal{P}_2^{s+1}$ the color of the edge between v_i and v_j (i.e. the square class of $x_i^2 + x_j^2$) depends precisely on $\max\{i, j, 3\}$. □

4.9. Corollary. *Let F be a nonreal field of level $s = \bar{q}_1(F) \geq 8$. Then $\bar{q}_2(F) \geq \frac{s^2}{2}$.*

Proof: Let $0 = x_1^2 + \cdots + x_{s+1}^2$ be a representation of zero as a nontrivial sum of $s+1$ squares over F. By the theorem we may, after reordering the indices, suppose that for $\{i, j\} \in \mathcal{P}_2^{s+1}$ the square class of $x_i^2 + x_j^2$ depends precisely on $\max\{i, j, 3\}$.

Defining $a_i := x_{i+1}^2 + x_{i+2}^2$ for $1 \leq i \leq s-1$, we get a system of representatives $a_1, \ldots, a_{s-1}$ of the $s-1$ nontrivial classes of $D_F(2)/F^{\times 2}$. Further we set $c_{jk} := x_1^2 + x_{j+2}^2 + x_{k+2}^2$ for $1 \leq j < k \leq s-1$.
Suppose now that $b\, c_{jk} = c_{j'k'}$ for $b \in D_F(2)$ and $1 \leq j' < k' \leq s-1$. Then $c_{j'k'} \in D_F(\langle 1, a_{j'}\rangle) \cap D_F(\langle 1, a_{k'}\rangle) \cap D_F(\langle b, b\, a_j\rangle) \cap D_F(\langle b, b\, a_k\rangle)$. In view of (b), (c) and (d) of the proposition this is only possible if $b \in F^{\times 2}$, $j = j'$ and $k = k'$.
This shows that the elements c_{jk} for $1 \leq j < k \leq s-1$ represent distinct nontrivial classes of $D_F(4)/D_F(2)$. Therefore $\bar{q}_2(F) > \binom{s-1}{2}$. Since s is a power of 2, at least 8, and $\bar{q}_2(F)$ is a power of 2 or infinite we obtain $\bar{q}_2(F) \geq \frac{s^2}{2}$. □

5 Lower bounds for the square class number

We start this section with Djoković's proof of his bound (1.6), rephrased in the terminology of appendix A.

5.1. Theorem (Djoković). *If F is a nonreal field of level $s \geq 8$ then*

$$q(F) \;\geq\; 2 \cdot |D_F(s/2)/F^{\times 2}| \;\geq\; 2 \cdot \sum_{i=1}^{s/2} \frac{1}{s+2-i} \binom{s+1}{i} .$$

Proof: The first inequality is clear since $|F^\times / D_F(s/2)| \geq 2$.
Next we consider a representation $0 = x_1^2 + \cdots + x_{s+1}^2$ of zero as a sum of $s+1$ nonzero squares over F. We denote by $\mathcal{P}$ the set of nonempty subsets of $\{1, \ldots, s+1\}$ of cardinality not greater than $s/2$. We define $f : \mathcal{P} \to D_F(s/2)/F^{\times 2}, J \mapsto (\sum_{j \in J} x_j^2) F^{\times 2}$. For $1 \leq k \leq s/2$ we write f_k for the restriction of f to $\mathcal{P}_k^{s+1}$. By (2.8), for $k \neq k'$ the images of f_k and $f_{k'}$ are disjoint. Also by (2.8), f_k is $(k-1)$-connected for any $k \leq s/2$ and therefore $|Im(f_k)| \geq \frac{1}{(s+1)-k+1} \binom{s+1}{k}$ by (A.4, c). All together we obtain

$$|D_F(s/2)/F^{\times 2}| \;\geq\; \sum_{k=1}^{s/2} |Im(f_k)| \;\geq\; \sum_{k=1}^{s/2} \frac{1}{s-k+2} \binom{s+1}{k}$$

which shows the second inequality. □

5.2. Remark. For an integer $s \geq 8$, let $\sum(s)$ denote the term on the right hand side in the inequality of the above theorem. Djoković showed by an elementary counting argument that $\sum(s) > \frac{2^s}{s}$ [2]. As was pointed out by David B. Leep, the argument may be improved to obtain the bound $\sum(s) > \frac{2^{s+1}}{s}$ for every even $s \geq 8$. Under the hypothesis of the last theorem one has thus $q(F) > \frac{2^{s+1}}{s}$; further, since $s = s(F)$ is a power of 2 and $q(F)$ is also a power of 2 or infinite, it follows that $q(F) \geq \frac{2^{s+2}}{s}$.
Our calculations have shown that, at least for s a power of 2 in the range between 8 and 2^{13}, actually one has $\frac{2^{s+1}}{s} < \sum(s) \leq \frac{2^{s+2}}{s}$.

However, for level 8 and 16 we get stronger bounds on $q(F)$.

5.3. Theorem. *Let F be a field. If $s(F) = 8$ then $q(F) \geq 512$. If $s(F) = 16$ then $q(F) \geq 2^{15}$.*

Proof: Under the hypothesis $s(F) = 8$ we have $\bar{q}_3(F) \geq 2$, $\bar{q}_2(F) \geq 16$ (3.1) and $\bar{q}_1(F) \geq 8$ (1.5). Moreover, by (4.9) one of the last two inequalities must be proper. From $|F^\times/{F^\times}^2| \geq \bar{q}_1(F) \cdot \bar{q}_2(F) \cdot \bar{q}_3(F)$ we get therefore $q(F) \geq 512$, since $F^\times/{F^\times}^2$ is an elementary abelian 2-group.
For $s(F) = 16$ we have by the previous sections $\bar{q}_4(F) \geq 2$, $\bar{q}_3(F) \geq 16$, $\bar{q}_2(F) \geq 32$ and $\bar{q}_1(F) \geq 16$ and one of the last two inequalities must be proper. As $|F^\times/{F^\times}^2| \geq \bar{q}_1(F) \cdots \bar{q}_4(F)$ this leads to $q(F) \geq 2^{15}$. □

For $s(F) = 2^n$ with $n \geq 5$ the analogous arguments are not sufficient to improve Djoković's result. For $s(F) = 32$, for example, we may get in this way $q(F) \geq 2^{25}$ while (5.1) yields $q(F) \geq 2^{29}$.

5.4. Theorem. *Let F be a field of level 2^n with $n \geq 3$. Then $|k_{n-1}F| \geq 128$. More precisely, the subgroup $\{-1\}^{n-2}k_1F$ of $k_{n-1}F$ is of index at least 4 and order at least 32.*

Proof: Again, we use the notation $\varepsilon := \{-1\} \in k_1F$. The homomorphism $F^\times \to \{-1\}^{n-2}k_1F$ which maps $x \in F^\times$ to the symbol $\varepsilon^{n-2} \cdot \{x\}$, has kernel $D_F(2^{n-2})$. Since $\bar{q}_n(F) \geq 2$ and $\bar{q}_{n-1}(F) \geq 16$ by (3.1), we have $|F^\times/D_F(2^{n-2})| \geq \bar{q}_n(F) \cdot \bar{q}_{n-1}(F) \geq 32$. Therefore $\{-1\}^{n-2}k_1F$ has at least 32 elements.
To show that the index of this group in $k_{n-1}F$ is at least 4 we just need to find $\alpha, \beta, \gamma \in k_{n-1}F \setminus \{-1\}^{n-2}k_1F$ such that $\alpha + \beta + \gamma \in \{-1\}^{n-2}k_1F$.
By the hypothesis there are $a, b, c \in D_F(3 \cdot 2^{n-3}) \setminus D_F(2^{n-2})$ such that $a + b + c = 0$. In k_2F we compute $\{-a, -b\} + \{-a, -c\} + \{-b, -c\} = \{-a, bc\} + \{a, -bc\} = \{-1, abc\}$. Therefore we are finished if we show that none of the symbols $\varepsilon^{n-3}\{-a, -b\}, \varepsilon^{n-3}\{-a, -c\}$ and $\varepsilon^{n-3}\{-b, -c\}$ in $k_{n-1}F$ lies actually in $\{-1\}^{n-2}k_1F$.
If this is not true we may by case symmetry suppose that $\varepsilon^{n-3}\{-a, -b\} = \varepsilon^{n-2}\{-x\}$ for some $x \in F^\times$. Then the $(n-1)$-fold Pfister forms $2^{n-3} \times \langle\langle a, b\rangle\rangle$ and $2^{n-2} \times \langle\langle x\rangle\rangle$ over F are isometric, i.e. the quadratic form $\varphi := 2^{n-3} \times \langle 1, x, x, -a, -b, -ab\rangle$ over F is hyperbolic. It follows that any subform of φ of dimension greater than $\frac{1}{2}\dim(\varphi) = 3 \cdot 2^{n-3}$ is isotropic. In particular, the form $2^{n-2} \times \langle -ax\rangle \perp 2^{n-3} \times \langle 1\rangle \perp \langle b\rangle$, similar to a subform of φ, must be isotropic. It follows that $ax \in D_F(2^{n-2}) \cdot D_F(2^{n-3} \times \langle 1\rangle \perp \langle b\rangle) \subset D_F(2^{n-1})$ whence $x \in D_F(2^{n-1})$. On the other hand, $\varphi \cong 2^{n-3} \times \langle 1, x, x, c, abc, -ab\rangle$ shows that $2^{n-2} \times \langle x\rangle \perp 2^{n-3} \times \langle 1\rangle \perp \langle c\rangle$ is isotropic. This in turn implies that $-x \in D_F(2^{n-2}) \cdot D_F(2^{n-3} \times \langle 1\rangle \perp \langle c\rangle) \subset D_F(2^{n-1})$. Together this leads to $-1 \in D_F(2^{n-1})$ which contradicts $s(F) = 2^n$. □

5.5. Corollary. *Let F be a nonreal field with $s(F) \geq 8$. Then $|\mathrm{Br}_2(F)| \geq 128$ and $|W(F)| \geq 2^{18}$.*

Proof: If $s(F) = 8$ then the theorem shows $|k_2F| \geq 128$. But this is also true if $s(F) = 2^n > 8$ since then already the subgroup $\{-1\}k_1F$, isomorphic to $F^\times/D_F(2)$, has order at least $\bar{q}_n(F) \cdot \bar{q}_{n-1}(F) \cdot \bar{q}_{n-2}(F)$ which is sufficiently large by the results of section 3. By Merkuriev's theorem, $\mathrm{Br}_2(F)$ is isomorphic to k_2F, so in particular we have $|\mathrm{Br}_2(F)| \geq 128$. (In fact, the arguments to estimate the size of k_2F work similarly for $\mathrm{Br}_2(F)$, so it is not necessary to invoke Merkuriev's theorem here.)
Let I denote the fundamental ideal of $W(F)$ and let $\bar{I}^i := I^i/I^{i+1}$ for $i \geq 0$. For $i = 0, 1, 2$ it follows from [9] that $\bar{I}^i \cong k_iF$. Thus $|\bar{I}^0| = 2$, $|\bar{I}^1| = q(F) \geq 512$ and $|\bar{I}^2| \geq 128$. Moreover, $s(F) \geq 8$ implies $|\bar{I}^3| \geq 2$. Therefore $|W(F)| \geq |\bar{I}^0| \cdot |\bar{I}^1| \cdot |\bar{I}^2| \cdot |\bar{I}^3| \geq 2^{18}$. □

A Hypergraphs with connected colorings

In this appendix t, k and n denote nonnegative integers with $t \leq k \leq n$. We briefly say *k-set* for a set of cardinality k. A *k-hypergraph* is a system $\mathcal{H} = (V, \mathcal{E})$ where V is a set whose elements are called *vertices* and $\mathcal{E}$ a collection of distinct k-subsets of V called *edges*. A graph in the usual sense is then just a 2-hypergraph.
Let $\mathcal{H} = (V, \mathcal{E})$ be a k-hypergraph. Its number of vertices $|V|$ is called the *order of* $\mathcal{H}$. We say that $\mathcal{H}$ is *complete* if each k-subset of V is actually an edge, i.e. if $\mathcal{E} = \{E \subset V \mid |E| = k\}$. By an *edge-coloring* of $\mathcal{H}$ we mean a function $f : \mathcal{E} \to C$. We consider the elements of C as *colors* and for $E \in \mathcal{E}$ we call $f(E)$ the *color of* E. For $t > 0$ we say that the edge-coloring f is *t-connected* if any two edges of the same color meet in at least t vertices, i.e. if for any $E, E' \in \mathcal{E}$ with $f(E) = f(E')$ we have $|E \cap E'| \geq t$.

A.1. Problem. *Let t, k, n be nonnegative integers with $t \leq k \leq n$. Let $\mathcal{H} = (V, \mathcal{E})$ be a complete k-hypergraph of order n. What is the least integer m such that there exists a t-connected edge-coloring $f : \mathcal{E} \to C$ on $\mathcal{H}$ with $|C| = m$?*

The integer m which meets the condition in the problem depends only on the values of t, k and n and will be denoted by $M(t, k, n)$. We recall our notation $\mathcal{P}_k^n$ for the set of all k-subsets of $\{1, \ldots, n\}$. A complete k-hypergraph of order n is then given by $\mathcal{K}_k^n := (\{1, \ldots, n\}, \mathcal{P}_k^n)$. So $M(t, k, n)$ is just the least integer m such that there exists a function $f : \mathcal{P}_k^n \to C$ where $|C| = m$ and such that $f(X) = f(X')$ implies $|X \cap X'| \geq t$ for any $X, X' \in \mathcal{P}_k^n$. To study $M(t, k, n)$ as a function in t, k and n we use the theory of *intersecting families* in combinatorics.
Let $\mathcal{F}$ be a family of sets. We write $\bigcup \mathcal{F}$ (resp. $\bigcap \mathcal{F}$) for the union (resp. the intersection) of all sets belonging to $\mathcal{F}$. If $|U \cap V| \geq t$ holds for every $U, V \in \mathcal{F}$ then we say that the family $\mathcal{F}$ is *t-intersecting* (just *intersecting* for $t = 1$). A

coloring $f : \mathcal{E} \to C$ of a k-hypergraph $\mathcal{H} = (V, \mathcal{E})$ is thus t-connected if and only if $f^{-1}(\{c\})$ is a t-intersecting family for every $c \in C$.
The crucial result on intersecting families is the Erdös-Ko-Rado theorem [4] which we state in the slightly stronger version of [14]:

A.2. Theorem (Erdös-Ko-Rado). *Let $n \geq (k-t+1)(t+1)$. If $\mathcal{F}$ is a t-intersecting family of k-subsets of an n-set then $|\mathcal{F}| \leq \binom{n-t}{k-t}$.*

This theorem gives the optimal bound. Indeed, if N is an n-set and T a t-subset then $\mathcal{F} := \{U \subset N \mid |U| = k, T \subset U\}$ is a t-intersecting family with precisely $\binom{n-t}{k-t}$ elements. However, under the additional condition $|\bigcap \mathcal{F}| < t$, better bounds on $|\mathcal{F}|$ can be given. In the case $t = 1$ this is the following main result of [6]. (A short proof of this can be found in [5] where the case $t > 1$ is also treated.)

A.3. Theorem (Hilton-Milner). *Let $\mathcal{F}$ be a family of pairwise intersecting k-subsets of an n-set such that $\bigcap \mathcal{F} = \emptyset$. Then $|\mathcal{F}| \leq \binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1$.*

Now we begin with the investigation $M(t,k,n)$ as a function in t, k and n with $0 < t \leq k \leq n$. We first treat the easy cases when t and k take extremal values. Part (c) is implicitly shown in [2].

A.4. Proposition. *(a) $M(t,k,n) = 1$ is equivalent to $n \leq 2k - t$.*

(b) $M(t,k,n) = \binom{n}{k}$ is equivalent to $k = t$.

(c) $M(k-1,k,n) = M(n-k-1,n-k,n) \geq \frac{1}{n-k+1}\binom{n}{k}$ for $1 \leq k \leq n/2$.

Proof: (a) $M(t,k,n)$ is equal to 1 if and only if $\mathcal{P}_k^n$ is t-intersecting; this is the case if and only if $n \leq 2k - t$.
(b) Each condition holds if and only if any nonempty t-intersecting family of k-subsets of $\{1,\ldots,n\}$ consists of just one k-set.
(c) It is quite obvious that a family $\mathcal{F} \subset \mathcal{P}_k^n$ is $(k-1)$-intersecting if and only if the family of complement sets $\{\{1,\ldots,n\} \setminus U \mid U \in \mathcal{F}\}$ is $(n-k-1)$-intersecting. So $f : \mathcal{P}_k^n \to C$ is $(k-1)$-connected if and only if $f' : \mathcal{P}_{n-k}^n \to C, V \mapsto f(\{1,\ldots,n\} \setminus V)$ is $(n-k-1)$-connected. This shows in particular $M(k-1,k,n) = M(n-k-1,n-k,n)$.
For a $(k-1)$-intersecting family $\mathcal{F} \subset \mathcal{P}_k^n$ it is easy to check that either $|\bigcap \mathcal{F}| \geq k-1$ or $|\bigcup \mathcal{F}| \leq k+1$. In the first case we conclude $|\mathcal{F}| \leq n-k+1$ and in the second case $|\mathcal{F}| \leq k+1 \leq n-k+1$. If now $f : \mathcal{P}_k^n \to C$ is $(k-1)$-connected then $\mathcal{P}_k^n$ is covered by the $(k-1)$-intersecting families $f^{-1}(\{c\})$ for $c \in C$, which implies that $\binom{n}{k} = |\mathcal{P}_k^n| \leq (n-k+1) \cdot |C|$. □

A.5. Examples. (1) The function $f : \mathcal{P}_k^n \to \mathcal{P}_t^{n-k+t}$ which associates to $X \in \mathcal{P}_k^n$ the set of the t smallest numbers in X is a t-connected edge-coloring of $\mathcal{K}_k^n$.

(2) If $n\geq 2k-1$ then a 1-connected edge-coloring of $\mathcal{K}_k^n$ is given by

$$f\,:\,\mathcal{P}_k^n\longrightarrow\{1,\ldots,n-2k+2\}\,,\;\;X\longmapsto \max\,(X\cup\{2k-1\})-2k+2\,.$$

(3) Let $t<k<n$. If $f:\mathcal{P}_k^n\to C$ be a t-connected edge-coloring of $\mathcal{K}_k^n$ and $g:\mathcal{P}_{k+1}^n\to C'$ is a $(t+1)$-connected edge-coloring of $\mathcal{K}_{k+1}^n$, where C and C' are disjoint sets, then a $(t+1)$-connected edge-coloring of $\mathcal{K}_{k+1}^{n+1}$ is defined by

$$h\,:\,\mathcal{P}_{k+1}^{n+1}\longrightarrow C\cup C'\,,\;\;X\longmapsto\begin{cases} f\,(X\backslash\{n+1\}) & \text{if } n+1\in X,\\ \quad g\,(X) & \text{otherwise.}\end{cases}$$

From these examples we conclude:

A.6. Proposition. *(a)* $M(t,k,n)\leq\binom{n-k+t}{t}$.

(b) If $n\geq 2k-1$ *then* $M(1,k,n)\leq n-2k+2$.

(c) If $t<k<n$ *then* $M(t+1,k+1,n+1)\leq M(t,k,n)+M(t+1,k+1,n)$. □

For lower bounds on $M(t,k,n)$ we first consider the case $t\geq 2$.

A.7. Theorem. *Let* $2\leq t<k$. *Then for* $n\geq(k-t+1)(t+1)$ *we have*

$$M(t,k,n)\;\geq\;\prod_{i=0}^{t-1}\frac{n-i}{k-i}\;>\;\left(\frac{n}{k}\right)^t.$$

Proof: Let $f:\mathcal{P}_k^n\to C$ be a t-connected edge-coloring of $\mathcal{K}_k^n$ with $n\geq(k-t+1)(t+1)$. For each $c\in C$ we have then by the Erdös-Ko-Rado theorem $|f^{-1}(\{c\})|\leq\binom{n-t}{k-t}$. As $\mathcal{P}_k^n=\bigcup_{c\in C}f^{-1}(\{c\})$ we get $\binom{n}{k}\leq|C|\cdot\binom{n-t}{k-t}$. Therefore $|C|\geq\frac{n}{k}\cdot\frac{n-1}{k-1}\cdots\frac{n-t+1}{k-t+1}$ and an easy computation shows the second inequality. □

For the purposes of section 3 we state the following particular case:

A.8. Corollary. *Let i and m be positive integers satisfying either* $2\leq i\leq\frac{m}{2}$ *or* $3\leq i=\frac{m+1}{2}$ *or* $5\leq i=\frac{m}{2}+1$. *Then* $M(2^{i-2}+1,2^i,2^m)>2^{(m-i)(2^{i-2}+1)}$. □

Now we come to the case $t=1$.

A.9. Lemma. *For $k>1$ we define the polynomial*

$$F_k(X)\;:=\;\prod_{i=0}^{k-1}(X-i)-k\,(X-2k+1)\left(\prod_{i=1}^{k-1}(X-i)-\prod_{i=1}^{k-1}(X-k-i)+(k-1)!\right).$$

If $k\leq n$ *and* $f:\mathcal{P}_k^n\to C$ *is such that* $\bigcap f^{-1}(\{c\})=\emptyset$ *for every* $c\in C$ *then either* $|C|\geq n-2k+2$ *or* $F_k(n)\leq 0$.

Proof: Suppose that f has the stated property. Then the Hilton-Milner theorem implies $\binom{n}{k} \le |C| \cdot [\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1]$. On the other hand, $(k!)^{-1} \cdot F_k(n) = \binom{n}{k} - (n-2k+1) \cdot [\binom{n-1}{k-1} - \binom{n-k-1}{k-1} + 1]$. Thus $F_k(n) > 0$ implies $|C| > (n-2k+1)$. $\square$

A.10. Remark. The polynomial F_k defined in the lemma is monic of degree k. In particular, we have $F_k(n) > 0$ for all n sufficiently large. Computation for small values of k yields: $F_2(X) = X^2 - 7X + 18$, $F_3(X) = X^3 - 21X^2 + 140X - 240$ and $F_4(X) = X^4 - 54X^3 + 731X^2 - 3534X + 5880$. Thus we have $F_2(n) > 0$ for any $n \in \mathbb{N}$, $F_3(n) > 0$ for $n \ge 3$ and $F_4(n) > 0$ for $n \ge 37$ whereas $F_4(36) < 0$.

A.11. Theorem. *For any $k \ge 1$ there is a constant $c_k \ge 2k-2$ such that for all $n \in \mathbb{N}$ sufficiently large we have*

$$M(1,k,n) = n - c_k.$$

For $k \le 3$ we have, more precisely, $M(1,k,n) = n - 2k + 2$ for $n \ge 2k-1$.

Proof: For $k = 1$ there is nothing to show since $M(1,1,n) = n$. For $k \ge 2$ let $F_k(X)$ be defined as in the lemma. By the above remark we may choose the least integer $n_k \ge 2k-1$ such that $F_k(n) > 0$ for all $n \ge n_k - 1$. In particular we have $n_2 = 3$ and $n_3 = 5$. Let $c_k := n_k - M(1,k,n_k)$. Then (A.6, b) implies $c_k \ge 2k-2$ and we check that equality holds for $k = 2, 3$.
We want to prove by induction that $M(1,k,n) = n - c_k$ for $n \ge n_k$. For $n = n_k$ this is trivial statement. Suppose it is true for $n - 1 \ge n_k$. Let $f : \mathcal{P}_k^n \to C$ be a 1-connected edge-coloring of $\mathcal{K}_k^n$. If $\bigcap f^{-1}(\{c\}) = \emptyset$ for each $c \in C$ then by the lemma we have $|C| \ge n - 2k + 2 \ge n - c_k$. On the other hand, if there is $c \in C$ such that the intersection $\bigcap f^{-1}(\{c\})$ is not empty then we may suppose that it contains the element n. Then the restriction $f' : \mathcal{P}_k^{n-1} \to C \setminus \{c\}$ of f to $\mathcal{P}_k^{n-1}$ is a 1-connected edge-coloring of $\mathcal{K}_k^{n-1}$. By the induction hypothesis we have $|C \setminus \{c\}| \ge M(1,k,n-1) = (n-1) - c_k$ and thus $|C| \ge n - c_k$. This implies $M(1,k,n) \ge n - c_k$. But (A.6, c) shows $M(1,k,n) \le M(1,k,n-1) + M(0,k-1,n-1) = n - c_k$ since $M(0,k-1,n-1) = 1$. Hence $M(1,k,n) \ge n - c_k$ which finishes the induction step. $\square$

A.12. Question. *Does $M(1,k,n) = n - 2k + 2$ hold for all $n \ge 2k-1$, even if $k > 3$?*

B CC-Graphs

In this appendix we study connected edge-colorings for usual complete graphs. Here we are not only interested in the minimal number of colors but also in the distribution of the colors in the graph.

Let $\mathcal{G}$ denote a complete graph of order n with vertices $v_1, \ldots, v_n$ and colored edges. The distribution of colors in $\mathcal{G}$ can be equivalently represented by an edge-coloring of $\mathcal{K}_2^n$ (see appendix A), i.e. by a function $f : \mathcal{P}_2^n \to C$, where C stands for the set of colors in $\mathcal{G}$ and f associates to $\{i,j\} \in \mathcal{P}_2^n$ the color of the edge between the vertices v_i and v_j.
A set of all the edges of a certain color shall be called a *color-component.* If such a color-component consists of $r \geq 3$ edges all together having a vertex x in common we call it an *r-star* and x its *center.* By a *triangle* in $\mathcal{G}$ we mean a complete subgraph of order 3 of $\mathcal{G}$. A *triangle* is said to be *monochrome* (resp. *three-colored*) if the three edges are of the same color (resp. of three different colors). A second complete colored graph $\mathcal{G}'$ of order n is said to be *equivalent to* $\mathcal{G}$ if there is a bijection between the sets of vertices of $\mathcal{G}$ and $\mathcal{G}'$ such that the induced bijection on the sets of edges preserves the color-components (in both directions).
We call $\mathcal{G}$ *color-connected* or a *CC-graph* if in $\mathcal{G}$ any two edges of the same color are adjacent. This is equivalent to the edge-coloring f being 1-connected. The only possible color-components in $\mathcal{G}$ are then single edges, pairs of edges with a vertex in common, stars and monochrome triangles.
Theorem (A.11) says that $M(1,2,n) = n-2$ for $n \geq 3$. This corresponds to a result of [13]. We rephrase it as follows and give a direct proof.

B.1. Proposition (Tort). *A CC-graph of order $n \geq 3$ has at least $n-2$ colors.*

Proof: For $n = 3$ the statement is trivial. If $n > 3$ and $\mathcal{G}$ has less than n colors then one of its color-components must be a star. Deleting the center of this star yields a CC-graph $\mathcal{G}'$ of order $n-1$ with less colors. By induction hypothesis $\mathcal{G}'$ has at least $n-3$ and therefore $\mathcal{G}$ at least $n-2$ colors. □

For any $n \geq 3$ the complete graph $\mathcal{K}_2^n$, whose vertices are the integers $1, \ldots, n$, together with the 1-connected coloring $f_n : \mathcal{P}_2^n \to \{1, \ldots, n-2\}$, $\{i,j\} \mapsto \max\{i,j,3\} - 2$ defines a particular CC-graph $\mathcal{G}_n$ of order n with $n-2$ colors (compare with example (A.5, 2)). The color-components of $\mathcal{G}_n$ are one monochrome triangle and one i-star for each $3 \leq i \leq n-1$. For $3 \leq n \leq 5$, every CC-graph with $n-2$ colors is equivalent to $\mathcal{G}_n$. This is not true for $n = 6$, since there is a CC-graph of order 6 with color-components a triangle and three 4-stars.

B.2. Proposition. *Let $\mathcal{G}$ be a CC-graph with $n \geq 3$ vertices and $n-2$ colors. Then $\mathcal{G}$ has as color-components one monochrome triangle and $n-3$ stars. Moreover, each vertex of $\mathcal{G}$ lies either on the monochrome triangle or is the center of exactly one star.*

Proof: Let $\mathcal{G}'$ be the complete subgraph spanned by all vertices of $\mathcal{G}$ which are not the center of a star in $\mathcal{G}$. We want to show that $\mathcal{G}'$ is a monochrome triangle. Then the vertices of $\mathcal{G}$ outside of $\mathcal{G}'$ will be the centers of $n-3$ stars and as $\mathcal{G}$ has just $n-2$ colors the entire statement follows.

Let n' be the order of $\mathcal{G}'$. The $n - n'$ vertices of $\mathcal{G}$ outside of $\mathcal{G}'$ are all centers of stars whose colors do not appear in $\mathcal{G}'$. As a consequence, $\mathcal{G}'$ has at least $n - n'$ colors less than $\mathcal{G}$. Then by (B.1), $\mathcal{G}'$ has exactly $n' - 2$ colors. Since $\mathcal{G}'$ is a graph without stars each color appears at most three times, counting the edges yields $3(n' - 2) \geq \frac{n'(n'-1)}{2}$ whence $n' \leq 5$. As $\mathcal{G}'$ has $n' - 2$ colors and contains no star, we have $n' = 3$ and $\mathcal{G}'$ is a monochrome triangle. $\square$

A CC-graph $\mathcal{G}$ will be called *total* if there is a permutation $\sigma \in \mathcal{S}_n$ such that for any $\{i, j\} \in \mathcal{P}_2^n$ the color of the edge between v_i and v_j depends only on $\max\{\sigma(i), \sigma(j)\}$. After renumbering the vertices $\mathcal{G}$ we may then suppose that the permutation σ is the identity on $\{1, \ldots, n\}$.
Let $\mathcal{G}$ be a total CC-graph of order n with vertices $v_1, \ldots, v_n$ enumerated in such a way that the color of any edge linking v_i and v_j depends only on $\max\{i, j\}$. Then $\mathcal{G}$ has at most $n-1$ different colors. From (B.1) it follows that the number of colors in $\mathcal{G}$ is either $n - 2$ or $n - 1$. Further, by (B.2) the number of colors is $n - 2$ if and only if v_1, v_2 and v_3 form a monochrome triangle and then the color of the edge between v_i and v_j depends precisely on $\max\{i, j, 3\}$. In both cases the enumeration of the vertices is unique up to changing the first three respectively the first two indices. Moreover, $\mathcal{G}$ contains exactly $n - 3$ stars. More precisely, for each $4 \leq i \leq n$ there is exactly one $(i-1)$-star in $\mathcal{G}$ whose center is v_i. It is clear from the definition that a complete subgraph of a total CC-graph is also a total CC-graph.

B.3. Proposition. *A CC-graph $\mathcal{G}$ is total if and only if it contains no three-colored triangle.*

Proof: The necessity of the condition follows from the definition of a total CC-graph. Suppose now that $\mathcal{G}$ is a CC-graph with n vertices with no three-colored triangle. We show by induction on n that $\mathcal{G}$ is total. For $n \leq 3$ this is evident. If $n \geq 4$ then any complete subgraph with 4 vertices contains a star since otherwise it would contain a three-colored triangle. So we can choose an r-star in $\mathcal{G}$ where r is as large as possible. For the ease of imagination say, it is of red color. We may suppose that v_n is the center of this star. Let $\mathcal{G}'$ be the complete subgraph of $\mathcal{G}$ with all the vertices of $\mathcal{G}$ except v_n. Then $\mathcal{G}'$ is also a CC-graph with $n-1$ vertices and contains no three-colored triangle. So, by the induction hypothesis, $\mathcal{G}'$ is total, i.e. its vertices can be enumerated as $v_1, \ldots, v_{n-1}$ in such a way that the color of an edge connecting vertices v_i and v_j depends just on $\max\{i, j\}$. This would still be true for the enumeration of the vertices $v_1, \ldots, v_n$ of $\mathcal{G}$, if v_n is connected with each of the $v_1, \ldots, v_{n-1}$ by an edge of red color. So we just have to show that $r = n - 1$. Suppose that $r < n - 1$. Then certainly $n > 4$ since $r \geq 3$ by the definition of an r-star. But v_{n-1} is the center of an $n-2$-star in $\mathcal{G}'$, say of blue color. By the maximality of r we see that the edge between v_{n-1} and v_n cannot be blue and that $r = n-2$. So there must be exactly one vertex v_k with $1 \leq k \leq n - 1$ which is connected with v_n with an edge of color different from red. It cannot be of blue color either so say that its color is green. Now we see that there is a triangle of colors red, blue and green contained in $\mathcal{G}$, formed by v_k, v_{n-1}, v_n if $k < n-1$ and by v_1, v_{n-1}, v_n if $k = n-1$, which gives the desired contradiction. $\square$

REFERENCES

[1] P. L. Chang. On the number of square classes of a field with finite Stufe. *J. Number Theory*, 6:360–368, 1974.

[2] Dragomir Ž. Djoković. Level of a field and the number of square classes. *Math. Z.*, 135:267–269, 1973/74.

[3] Richard Elman and T. Y. Lam. Pfister forms and K-theory of fields. *J. Algebra*, 23:181–213, 1972.

[4] P. Erdős, Chao Ko, and R. Rado. Intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 12:313–320, 1961.

[5] P. Frankl and Z. Füredi. Nontrivial intersecting families. *J. Combin. Theory Ser. A*, 41(1):150–153, 1986.

[6] A. J. W. Hilton and E. C. Milner. Some intersection theorems for systems of finite sets. *Quart. J. Math. Oxford Ser. (2)*, 18:369–384, 1967.

[7] Irving Kaplansky. Quadratic forms. *J. Math. Soc. Japan*, 5:200–207, 1953.

[8] T. Y. Lam. *The algebraic theory of quadratic forms*. W. A. Benjamin, Inc., Reading, Mass., 1973. Mathematics Lecture Note Series.

[9] John Milnor. Algebraic K-theory and quadratic forms. *Invent. Math.*, 9:318–344, 1969/1970.

[10] Albrecht Pfister. Zur Darstellung von -1 als Summe von Quadraten in einem Körper. *J. London Math. Soc.*, 40:159–165, 1965.

[11] Albrecht Pfister. Quadratische Formen in beliebigen Körpern. *Invent. Math.*, 1:116–132, 1966.

[12] Winfried Scharlau. *Quadratic and Hermitian forms*. Springer-Verlag, Berlin, 1985.

[13] Jean-René Tort. Sur deux invariants d'un corps de niveau fini. *Bull. Soc. Math. France*, 106(2):161–168, 1978.

[14] Richard M. Wilson. The exact bound in the Erdős-Ko-Rado theorem. *Combinatorica*, 4(2-3):247–257, 1984.

Karim Johannes Becher
Laboratoire de Mathématiques
U.F.R. Sciences et Techniques
Université de Franche-Comté
16, Route de Gray
25030 Besançon Cedex
France
becher@math.univ-fcomte.fr

2-Torsion of the Brauer Group of an Elliptic Curve: Generators and Relations

V. Chernousov, V. Guletskiĭ

Received: May 1, 2001

Communicated by Ulf Rehmann

Abstract. In this paper we describe the 2-torsion part of the Brauer group $\mathrm{Br}\, E$ of an elliptic curve E defined over an arbitrary field of characteristic $\neq 2$ in terms of generators and relations.

2000 Mathematics Subject Classification: 14H52, 16H05, 16K20
Keywords and Phrases: Elliptic curves, Brauer groups.

1 *Introduction*

Let E be an elliptic curve defined over a field K of characteristic different from 2 and given by an affine equation

$$y^2 = f(x),$$

where $f(x)$ is a unitary cubic polynomial over K without multiple roots. We will say that E is *split*, *semisplit* or *non-split* if $f(x)$ has 3, 1 or no roots in K respectively.

Let $\mathrm{Br}\, E$ be the Brauer group of the curve E. The group $\mathrm{Br}\, E$ plays an important role in arithmetic and algebraic geometry. For example, it can be used to study arithmetical properties of elliptic surfaces and some other algebraic varieties (cf. [AM72], [CEP71], [CSS98], [S99]). Another important application is the construction of unirational varieties which are not rational. Let us describe the last point in some more details. We follow the famous paper of Artin and Mumford [AM72] slightly modifying their examples.

Let S be a smooth projective surface defined over an algebraically closed field of characteristic $\neq 2$, say $\mathbb{C}$ for simplicity. Assume that S is a rational elliptic

surface defined by a regular map $\pi : S \to \mathbb{P}^1$ such that the generic fiber $E_\xi = \pi^{-1}(\xi)$ is an elliptic curve.
Given a quaternion algebra $D = (d_1, d_2)$ over the function field $L = \mathbb{C}(S)$ of the surface S, whose ramification curve has nonsingular components, one can associate a smooth S-scheme $\phi : V_D \to S$ in a natural way, all of whose geometric fibres are isomorphic to $\mathbb{P}^1$ or to $\mathbb{P}^1 \vee \mathbb{P}^1$ (the so-called Brauer-Severi scheme). Let C be the ramification curve of D and let $C = C_1 \cup \ldots \cup C_n$ be its decomposition into irreducible components. The remarkable thing about V_D is that V_D viewed as a variety over $\mathbb{C}$ is not rational if all components $C_1, \ldots, C_n$ are disjoint. Namely, Artin and Mumford [AM72] proved that under these conditions V_D has 2-torsion in $H^3(V_D, \mathbb{Z})$. Since the torsion in H^3 is a birational invariant for complete smooth 3-dimensional varieties, V_D is not rational.
On the other hand, it turns out that for many quaternion algebras D the variety V_D is unirational. To prove it we first remark that if we want to have the ramification curve C of D with disjoint irreducible components it is natural to take D such that C has vertical components (with respect to π) only. It easily follows that all candidates for such D are among quaternion algebras in the Brauer group of the generic fiber E_ξ. As we show in this paper, there are lots of non-trivial quaternion algebras in $\mathrm{Br}\, E_\xi$. Taking the appropriate D we may assume that C has ≥ 2 irreducible components. As it was said, this implies that the corresponding V_D is not rational.
Now let η be a generic point of S. Then $V_\eta = \phi^{-1}(\eta)$ is a conic over $\mathbb{C}(\eta) = \mathbb{C}(S) = L$. Consider the extension F/L of degree 4 corresponding to the Kummer map $E_\xi \xrightarrow{2} E_\xi$. It kills D, hence the conic V_η has an F-point. In particular V_η is rational over F, i.e. the function field $F(V_\eta)$ is isomorphic to $F(z)$ over F, where z is a transcendental variable over F. Furthermore, since F/L corresponds to the Kummer map, we have $F \overset{\mathbb{C}(t)}{\simeq} \mathbb{C}(t)(E_\xi)$, hence

$$F(V_\eta) \overset{F}{\simeq} F(z) \overset{\mathbb{C}(t)}{\simeq} \mathbb{C}(t)(E_\xi)(z) = \mathbb{C}(S)(z)$$

is a purely transcendental extension of $\mathbb{C}$. Here we used the fact that S is a rational surface. Finally, since $\mathbb{C}(V_D) = L(V_\eta)$ is a subfield of $F(V_\eta)$, V_D viewed as a 3-fold variety over $\mathbb{C}$ is unirational.

Our construction shows that if we want to produce an explicit example of an unirational variety which is not rational, one needs to know the structure of 2-torsion of $\mathrm{Br}\, E_\xi$. So it makes sense to get an explicit description of 2-torsion ${}_2\mathrm{Br}\, E$ of the Brauer group of an elliptic curve E defined over an arbitrary field K. One of the main goals of this paper is to accomplish (to some extent) a description of ${}_2\mathrm{Br}\, E$ in terms of generators and relations. The initial results in this direction were obtained in [Pu98] where a description of quaternion algebras over E is presented and in [GMY97] where an explicit description of generators of ${}_2\mathrm{Br}\, E$ for a split elliptic curve is given. The second-named author [G99] generalized the results of [GMY97] for semisplit elliptic curves. Our paper, in fact, grew out of his preprint [G99] and here we go further and

obtain more complete results that concern generators as well as relations for arbitrary elliptic curves. Our arguments are elementary and based only on using standard properties of restriction and corestriction maps for H^1 with coefficients in certain finite modules.
After this paper was released as a preprint [CG00] we learnt of the nice paper [S99] of Skorobogatov where he gave, among other things, a description of generators of the Brauer groups of algebraic varieties X defined over a field K of characteristic 0 satisfying the condition $H^0(\overline{K}, G_m) = \overline{K}[X]^\times = \overline{K}^\times$ where $\overline{K}$ is an algebraical closure of K. In that paper the generators of $\mathrm{Br}\, X$ are given in the form of the cup product of certain torsors over X and cocycles in H^1 with coefficients in finitely generated submodules of $\mathrm{Pic}\,(\overline{X})$. The proofs in [S99] are based on the heavy machinery of homological algebra. However, it seems worth while to have elementary constructions and proofs for elliptic curves as well.

We proceed to describe our results. Let $\overline{K}$ be a separable closure of K and $\overline{E} = E(\overline{K})$. The starting point of our consideration is the following exact sequence:

$$0 \to \mathrm{Br}\, K \longrightarrow \mathrm{Br}\, E \xrightarrow{\kappa} H^1(K, \overline{E}) \to 0\,. \tag{1}$$

Since $E(K) \neq \emptyset$, the homomorphism κ has a section, so that (1) induces the exact sequence

$$0 \to {}_2\mathrm{Br}\, K \longrightarrow {}_2\mathrm{Br}\, E \xrightarrow{\kappa} {}_2H^1(K, \overline{E}) \to 0\,,$$

where the subscript 2 means the 2-torsion part.
The main result of the paper is formulated in Theorems 3.6, 4.12, 5.2 and 5.3. After some preliminaries given in Section 2 we construct a section for κ in an explicit form. This eventually enables us to give an explicit description of ${}_2\mathrm{Br}\, E$ in terms of generators and relations.
More exactly, let M be the 2-torsion part of $\overline{E}$ and let $\Gamma = \mathrm{Gal}\,(\overline{K}/K)$. The Kummer sequence

$$0 \to M \longrightarrow \overline{E} \xrightarrow{2} \overline{E} \to 0,$$

where the symbol 2 over the arrow means multiplication by 2, yields the exact sequence

$$0 \to E(K)/2 \xrightarrow{\delta} H^1(\Gamma, M) \xrightarrow{\zeta} {}_2H^1(\Gamma, \overline{E}) \to 0\,.$$

Here $\delta : E(K)/2 \hookrightarrow H^1(\Gamma, M)$ is a connecting homomorphism. In Sections 3 through 5 we show that there exists a homomorphism $\epsilon : H^1(\Gamma, M) \to {}_2\mathrm{Br}\, E$ with the properties

$$\kappa \circ \epsilon = \zeta, \qquad \epsilon\,(\ker(\zeta)) = 0. \tag{2}$$

The second property implies that ϵ factors through ${}_2H^1(\Gamma, \overline{E})$, i.e. there is a unique homomorphism $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$ such that $\varepsilon \circ \zeta = \epsilon$, and the first one shows that ε is a required section.

If $f(x) = (x-a)(x-b)(x-c)$ with a, b, $c \in K$, then $M \simeq \mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$; hence

$$H^1(\Gamma, M) \simeq K^*/(K^*)^2 \times K^*/(K^*)^2.$$

It turns out that the map

$$\epsilon : \; K^*/(K^*)^2 \times K^*/(K^*)^2 \to {}_2\mathrm{Br}\, E$$

which takes a pair $(r,\, s) \in K^* \times K^*$ into the product $(r, x-b) \otimes (s, x-c)$ of quaternion algebras over $K(E)$ satisfies (2). Thus letting $I = \mathrm{Im}\, \epsilon$, we obtain the natural isomorphism ${}_2\mathrm{Br}\, E \simeq {}_2\mathrm{Br}\, K \oplus I$ where, by construction, the second summand I is generated by quaternion algebras over $K(E)$ of the form $(r, x-b)$ and $(s, x-c)$ with $r, s \in K^*$.

Assume that $f(x)$ does not split over K. We denote the minimal extension of K over which a section ϵ is already constructed by L. Then using standard properties of restriction and corestriction maps we show that for a special map $\tau : H^1(K, M) \to H^1(L, M)$ the composition $\epsilon = \mathrm{cor} \circ \epsilon_L \circ \tau$ satisfies (2). As a corollary of our construction, we again obtain the decomposition

$$ {}_2\mathrm{Br}\, E \simeq {}_2\mathrm{Br}\, K \,\oplus\, \mathrm{cor}\; (\mathrm{Im}\, \epsilon_L)\,. \tag{3}$$

Note that in all cases the degree of L/K is either two or three. This fact enables us to present generators of the second summand in (3) in an explicit form. It turns out that all of them are tensor products of quaternion algebras over $K(E)$ of a very specific form.

It follows from the construction that all relations between our generators are given by algebras from $(\epsilon \,\circ\, \delta)(E(K)/2)$. These algebras are also presented in an explicit form in Theorems 3.6, 4.12, 5.2 and 5.3 and all of them are parametrized by K-points of the elliptic curve E. This result shows that the two problems of an explicit description of the 2-torsion part of $\mathrm{Br}\, E$ (of course, modulo *numerical algebras*, i.e. algebras from ${}_2\mathrm{Br}\, K$) and the group $E(K)/2$ are, in fact, equivalent. So, every time information about $E(K)/2$ is available we can effectively describe ${}_2\mathrm{Br}\, E$ and vice versa.

In the second part of the paper we apply our results to the computation of ${}_2\mathrm{Br}\, E$ for an elliptic curve E over a local non-dyadic field K. In this case the structure of the group $E(K)$ is well understood. Applying known results we easily construct generators of $E(K)/2$ in Sections 7 and 8. This, in turn, yields an explicit description of ${}_2\mathrm{Br}\, E$ in the concluding Sections 8 and 9 very quickly. Thus, we reopen a result of Margolin and Yanchevskii [YM96]. It seems that in this part our argument is more natural and shorter (cp. loc. cit.).

Finally, we remark that by repeating almost verbatim our argument one can describe in a similar way the 2-torsion part of $\mathrm{Br}\, X$ for a hyperelliptic curve X defined over a field K such that $X(K) \neq \emptyset$. However, in order to keep the volume reasonable we do not consider hyperelliptic curves in the present paper.

If A is an abelian group, $A \xrightarrow{2} A$ denotes the homomorphism of multiplication by 2 and ${}_2A$, $A/2$ are its kernel and cokernel respectively.

$|S|$ denotes the number of elements in a finite set S.
Throughout this paper all fields under consideration are of characteristic $\neq 2$. For a field K denote by $\overline{K}$ a separable closure of K, K^* its multiplicative group and K^{*2} the subgroup of squares. By abuse of language, we will write s for a coset sK^{*2}, whenever there is no danger of confusion.
A variety is always a smooth projective and geometrically integral scheme over a field K. For a variety X over K, we write $K(X)$ for the function field of X and $X(K)$ for the set of its K-points. If L/K is a field extension, we put $X_L = X \times_{Spec\,K} Spec\,L$. We also write $\overline{X} = X \times_{Spec\,K} Spec\,\overline{K}$ and for brevity $\overline{K}$-points of $\overline{X}$ will be denoted by the same symbol $\overline{X}$.
In the paper we will consider quaternion algebras and their tensor products only. If A is a central simple algebra over a field K then $[A]$ means its class in the Brauer group $\operatorname{Br} K$. If $a, b \in K^*$ and (a, b) is a quaternion algebra, then, for short, we write $[a, b]$ instead of $[(a, b)]$. The group law in a Brauer group we always write additively: if $a, b, c, d \in F^*$, then $[(a, b) \otimes (c, d)] = [a, b] + [c, d]$.
If Γ is a profinite group, then $H^*(\Gamma, -)$ is a Galois cohomology functor. Let Λ be a subgroup of finite index in Γ. Then $\operatorname{res} : H^*(\Gamma, -) \to H^*(\Lambda, -)$ and $\operatorname{cor} : H^*(\Lambda, -) \to H^*(\Gamma, -)$ are the restriction and corestriction homomorphisms respectively. In particular, if $\Gamma = \operatorname{Gal}(\overline{K}/K)$ and Λ corresponds to a finite exension F/K then (using the cohomological description of Brauer groups) we have the homomorphism of a scalar extension $\operatorname{Br} K \to \operatorname{Br} F$ and the corestriction homomorphism $\operatorname{cor}_{F/K} : \operatorname{Br} F \to \operatorname{Br} K$. Thus, $\operatorname{cor}_{F/K}[A]$ means the value of the homomorphism $\operatorname{cor}_{F/K}$ on the class $[A] \in \operatorname{Br} F$.
If E is an elliptic curve over K, then its Brauer group is naturally isomorphic to the unramified Brauer group $\operatorname{Br}_{nr}(K(E)/K)$ (see [Lich69], [Co88]). So we will always identify $\operatorname{Br} E$ with $\operatorname{Br}_{nr}(K(E)/K)$.

Acknowledgements. The authors gratefully acknowledge the support of SFB 343 "Diskrete Strukturen in der Mathematik", TMR ERB FMRX CT-97-0107 and the hospitality of the University of Bielefeld. We would like also to express our thanks to H. Abels and U. Rehmann for support and encouragement during the preparation of this paper and O. Izhboldin for useful discussions.

2 *Preliminaries*

Let E be an elliptic curve over a field K defined by an affine equation

$$y^2 = f(x),$$

where $f(x)$ is a unitary cubic polynomial over K without multiple roots. Let O be the infinite point on E. On the set of K-points $E(K)$ there is a natural structure of an abelian group, such that O is a zero element. Throughout the paper we denote the 2-torsion subgroup in $\overline{E}$ by M. Let $\Gamma = \operatorname{Gal}(\overline{K}/K)$ be the absolute Galois group of the ground field K. If

$$f(x) = (x - a)(x - b)(x - c)$$

is the decomposition of $f(x)$ over $\overline{K}$, then

$$M = \{\, O,\ (a,0),\ (b,0)\,(c,0)\,\}.$$

We say that E is *split* if $a, b, c \in K$. In this case $M \subset E(K)$; hence M is a trivial Γ-module. We say that E is *semisplit* if $f(x)$ has one root in K only. If $f(x)$ is irreducible over K, then we say that E is *non-split*.

A starting point of our explicit description of ${}_2\mathrm{Br}\, E$ is the following exact sequence:

$$0 \to \mathrm{Br}\, K \xrightarrow{\iota} \mathrm{Br}\, E \xrightarrow{\kappa} H^1(\Gamma, \overline{E}) \to 0\ . \tag{4}$$

Here the maps ι and κ are defined as follows (see details in [Fadd51], [Lich69], [Mi81] or [Sch69]). Recall that we identify $\mathrm{Br}\, E$ with the unramified Brauer group $\mathrm{Br}_{nr}(K(E)/K)$. Then ι is induced by the scalar extension functor: if A is a central simple algebra over K, then $\iota([A]) = [A \otimes_K K(E)]$.

Next let $h \in \mathrm{Br}\, E$. By Tsen's theorem (see [P82]), we have $\mathrm{Br}\, K(E) \cong H^2(\Gamma, \overline{K}(E)^*)$. Hence h can be viewed as an element in $H^2(\Gamma, \overline{K}(E)^*)$. Let $\mathrm{Div}\, \overline{E}$ be the group of divisors on $\overline{E}$ and let $\mathrm{P}\,(\overline{E})$ be the group of principal divisors on $\overline{E}$. Let h' be the image of h under the homomorphism

$$H^2(\Gamma, \overline{K}(E)^*) \longrightarrow H^2(\Gamma, \mathrm{P}\,(\overline{E}))$$

induced by the map $\overline{K}(E)^* \to \mathrm{P}\,(\overline{E})$ that takes a rational function f to its divisor $\mathrm{div}(f)$. Since h belongs to the unramified subgroup of $\mathrm{Br}\, K(E) \cong H^2(\Gamma, \overline{K}(E)^*)$, it follows that h' lies in the kernel of the homomorphism

$$H^2(\Gamma, \mathrm{P}\,(\overline{E})) \longrightarrow H^2(\Gamma, \mathrm{Div}(\overline{E})) \tag{5}$$

induced by the embedding $\mathrm{P}\,(\overline{E}) \to \mathrm{Div}(\overline{E})$.

Let $\mathrm{Div}^0(\overline{E})$ be the group of degree zero divisors on $\overline{E}$. Clearly, $H^1(\Gamma, \mathbb{Z}) = 0$, so that a natural homomorphism $H^2(\Gamma, \mathrm{Div}^0(\overline{E})) \to H^2(\Gamma, \mathrm{Div}(\overline{E}))$ is injective. Therefore, the kernel of (5) coincides with the kernel of

$$H^2(\Gamma, \mathrm{P}\,(\overline{E})) \longrightarrow H^2(\Gamma, \mathrm{Div}^0(\overline{E}))$$

and the last one coincides with the image of the connecting homomorphism

$$\partial : H^1(\Gamma, \overline{E}) \longrightarrow H^2(\Gamma, \mathrm{P}\,(\overline{E}))$$

induced by the exact sequence

$$0 \to \mathrm{P}\,(\overline{E}) \longrightarrow \mathrm{Div}^0(\overline{E}) \longrightarrow \overline{E} \to 0\ .$$

Since $E(K) \neq \emptyset$ and $H^1(\Gamma, \mathbb{Z}) = 0$, we easily get

$$H^1(\Gamma, \mathrm{Div}^0(\overline{E})) = H^1(\Gamma, \mathrm{Div}(\overline{E})) = 1,$$

so that ∂ is injective. It follows that there exists a unique element $h'' \in H^1(\Gamma, \overline{E})$ such that $\partial(h'') = h'$. Then, by definition, $\kappa(h) = h''$.

We claim that sequence (4) splits. Indeed, if $x \in E(K)$ and $K(E)_x$ is the completion of $K(E)$ at x, then $\mathrm{Br}\, K(E)_x \cong \mathrm{Br}\, K \oplus \mathrm{Hom}_{cont}(\Gamma, \mathbb{Q}/\mathbb{Z})$. Let

$$\varsigma : \mathrm{Br}\, E \longrightarrow \mathrm{Br}\, K$$

be the composition

$$\mathrm{Br}\, E \hookrightarrow \mathrm{Br}\, K(E) \to \mathrm{Br}\, K(E)_x \cong \mathrm{Br}\, K \oplus \mathrm{Hom}_{cont}(\Gamma, \mathbb{Q}/\mathbb{Z}) \to \mathrm{Br}\, K$$

where the last homomorphism is the projection on the first summand. It is easy to check that the composition $\varsigma \circ \iota$ is an identical map and the claim follows. In view of splitness, (4) induces the exact sequence

$$0 \to {}_2\mathrm{Br}\, K \xrightarrow{\iota} {}_2\mathrm{Br}\, E \xrightarrow{\kappa} {}_2H^1(\Gamma, \overline{E}) \to 0, \tag{6}$$

which also splits. Since ${}_2H^1(\Gamma, \overline{E})$ can be easily computed, we obtain that for an explicit description of ${}_2\mathrm{Br}\, E$ it suffices to construct a section for κ. To do it, we first consider the Kummer sequence

$$0 \to M \longrightarrow \overline{E} \xrightarrow{2} \overline{E} \to 0. \tag{7}$$

It yields the exact sequence

$$0 \to E(K)/2 \xrightarrow{\delta} H^1(\Gamma, M) \xrightarrow{\zeta} {}_2H^1(\Gamma, \overline{E}) \to 0 \tag{8}$$

where $\delta : E(K)/2 \hookrightarrow H^1(\Gamma, M)$ is a connecting homomorphism. In the next three sections we will construct a homomorphism $\epsilon : H^1(\Gamma, M) \to {}_2\mathrm{Br}\, E$ with the properties

$$\kappa \circ \epsilon = \zeta, \qquad \epsilon\,(\ker(\zeta)) = 0.$$

The second property implies that ϵ induces a unique homomorphism ε : ${}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$ such that $\varepsilon \circ \zeta = \epsilon$. Then it follows that $\kappa \circ \varepsilon \circ \zeta = \kappa \circ \epsilon = \zeta$. Since ζ is surjective, we conclude that $\kappa \circ \varepsilon = 1$, i.e. ε is a required section for κ.

Letting $I = \mathrm{Im}\, \varepsilon$, we have ${}_2\mathrm{Br}\, E \cong I \oplus \mathrm{Im}\, \iota \cong I \oplus {}_2\mathrm{Br}\, K$. As we see in Sections 3, 4 and 5, elements in I are tensor product of quaternion algebras over $K(E)$ of a very specific form. So our construction eventually gives a simple system of generators of ${}_2\mathrm{Br}\, E$ modulo *numerical algebras* (i.e. algebras from $\mathrm{Im}\, \iota$) and according to the construction of the maps ϵ and ε all relations between the generators are given by algebras from $\epsilon(\ker(\zeta))$. Thus, to find all relations explicitly, we first have to describe the subset $\mathrm{Im}\, \delta \subset H^1(\Gamma, M)$ and then apply ϵ to its elements.

Since the structure of the group $H^1(\Gamma, M)$ (and hence the construction of ϵ) depends on splitting properties of the polynomial $f(x)$, to realize our program we consider split, semisplit and non-split cases in the next three sections separately.

3 *Split elliptic case*

Let E be a split elliptic curve. Then M is a trivial Γ-module; hence we have

$$H^1(\Gamma, M) = \mathrm{Hom}\,(\Gamma, M)\ .$$

Fix two non-zero points in M, say $(b, 0)$ and $(c, 0)$. Considering them as generators of M we have an isomorphism

$$M \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2\ .$$

It induces the isomorphism

$$H^1(\Gamma, M) = \mathrm{Hom}(G, M) \cong K^*/{K^*}^2 \oplus K^*/{K^*}^2\ .$$

Consider a map

$$\epsilon_b : K^*/{K^*}^2 \longrightarrow {}_2\mathrm{Br}\,E$$

which takes $s \in K^*$ into the class $[s, x - b]$. Here and below, for an element $r \in K$ the polynomial $x - r$ is considered as a rational function on E. Clearly, the quaternion algebra $(s, x - b)$ is unramified and ϵ_b is a homomorphism. Analogously, consider a homomorphism

$$\epsilon_c : K^*/{K^*}^2 \longrightarrow {}_2\mathrm{Br}\,E$$

which takes $s \in K^*$ into the class $[s, x - c]$. Let now

$$\epsilon = \epsilon_b \oplus \epsilon_c : K^*/{K^*}^2 \oplus K^*/{K^*}^2 = \mathrm{Hom}(\Gamma, M) \longrightarrow {}_2\mathrm{Br}\,E\ . \tag{9}$$

Using the description of κ given in Section 2 it is easy to show that $\kappa \circ \epsilon = \zeta$.

Lemma 3.1 $\kappa \circ \epsilon = \zeta$.

Proof. Let P be a non zero point in M. For any $s \in K^* \backslash {K^*}^2$ let $\phi_{P,s}$ be a homomorphism from Γ into M, such that $\phi_{P,s}(g) = P$ if $g \notin U_s = \mathrm{Gal}(\overline{K}/K(\sqrt{s}))$ and $\phi_{s,b}(g) = O$ otherwise. The group $H^1(\Gamma, M) = Hom(\Gamma, M)$ is generated by the homomorphisms of type $\phi_{P,s}$. Therefore it is sufficient to show that $(\kappa \circ \epsilon)(\phi_{P,s}) = \zeta(\phi_{P,s})$ for any P and s.
Let $\Phi_{P,s}$ be a homomorphism from Γ into $\mathrm{Div}^0(\overline{E})$, such that $\Phi_{P,s}(g) = (P) - (O)$ if $g \notin U_s$ and $\Phi_{s,b}(g) = 0$ otherwise. Let $d\Phi_{P,s} : \Gamma \times \Gamma \to \mathrm{Div}^0(\overline{E})$ be a codifferential of $\Phi_{P,s}$, that is

$$(d\Phi_{P,s})(g_1, g_2) = g_1\Phi_{P,s}(q_2) - \Phi_{P,s}(g_1g_2) + \Phi_{P,s}(g_2)$$

for any $g_1, g_2 \in \Gamma$. Then $d\Phi_{P,s}$ takes its values in $\mathrm{P}(E)$ and $\partial(\mathrm{cls}(\phi_{P,s})) = \mathrm{cls}(d\Phi_{P,s})$ where cls denotes a cohomology class of a cocycle. Using the above formula for $d\Phi_{P,s}$ it is easy to compute that $d\Phi_{P,s}(g_1, g_2) = 2(P) - 2(O)$ if g_1 and g_2 lie in $\Gamma \backslash U_s$ and $d\Phi_{P,s}(g_1, g_2) = 0$ otherwise. Let $x(P)$ be the x-coordinate of P and let $\psi_{P,s} : \Gamma \times \Gamma \to \overline{K}(E)^*$ be a map, such that

$\Psi_{P,s}(g_1, g_2) = x - x(P)$ if g_1 and g_2 lie in $\Gamma \backslash U_s$ and $d\Phi_{P,s}(g_1, g_2) = 1$ otherwise. Then we see that the composition of $\Psi_{P,s}$ with the natural homomorphism div : $\overline{K}(E)^* \to \mathrm{P}(\overline{E})$ coincides with the homomorphism $d\Phi_{P,s}$. Therefore $\partial(\mathrm{cls}(\phi_{P,s})) = \eta(\mathrm{cls}(\Psi_{P,s}))$. Since $\Psi_{P,s}$ is a cocycle of the unramified quaternion algebra $(s, x - x(P))$, we see that $\kappa([s, x - x(P)]) = \mathrm{cls}(\phi_{P,s})$. But $[s, x - x(P)]$ is equal to $\epsilon(\phi_{P,s})$. So we have $\kappa(\epsilon(\phi_{P,s})) = \mathrm{cls}(\phi_{P,s}) = \zeta(\phi_{P,s})$. □

According to our plan we also need to make sure that $\epsilon(\mathrm{Im}\,\delta) = 0$. The description of $\mathrm{Im}(\delta)$ in the split case is well known. However for the reader's convenience we describe this image in details.

To ease notation, for a point $(u, v) \in E(K)$ the coset $(u, v) + 2E(K)$ will be denoted by the same symbol (u, v). We start with a simple lemma which gives a formula for dividing a point $(u, v) \in E(K)$ in the group $\overline{E}$ by 2. Let

$$r = \sqrt{u-a}\ , \quad s = \sqrt{u-b}\ , \quad t = \sqrt{u-c} \quad \text{and} \quad w = r + s - t\ .$$

Let also

$$p = \frac{1}{2}(w^2 - (r^2 + s^2 + t^2)) + u = rs - rt - st + u \quad \text{and} \quad q = w(p - u) + v\ .$$

Lemma 3.2 *We have $(p, q) \in \overline{E}$ and $2(p, q) = (u, v)$.*

Proof. This is a straightforward calculation (see also the proof of Theorem 4.1 on page 38 in [Hu87]) and we omit the details to the reader. □

Proposition 3.3 *Let $(u, v) \in E(K)$. Then*

$$\delta(u,v) = \begin{cases} (u-c, u-b) & \text{if } u \neq b \text{ and } u \neq c, \\ (b-c, (b-c)(b-a)) & \text{if } u = b, \\ ((c-a)(c-b), c-b) & \text{if } u = c, \\ (1,1) & \text{if } u = \infty. \end{cases}$$

Proof. If $u = b$, then $u \neq a$ and $u \neq c$ and, analogously, if $u = c$, then $u \neq a$ and $u \neq b$. Therefore, by the symmetry argument, it suffices to prove the statement in the case $u \neq b$ and $u \neq c$. Moreover, we consider only "a generic case" where $u - b$ and $u - c$ generate a subgroup in K^*/K^{*2} of order 4, i.e. $u - b$ and $u - c$ are nontrivial and different modulo squares. The other cases can be handled in a similar way.

We keep the notation of Lemma 3.2. Since $2(p, q) = (u, v)$, the cocycle $\delta(u, v)$ corresponds to the homomorphism $\phi_{(u,v)} : \Gamma \to M$ that takes γ to the point $(p, q)^\gamma - (p, q)$. Let $U = \mathrm{Gal}(\overline{K}/K(s))$ and $V = \mathrm{Gal}(\overline{K}/K(t))$. We fix arbitrary automorphisms

$$\sigma \in U \backslash V \quad \text{and} \quad \tau \in V \backslash U\ .$$

Let $\psi_{(u,v)} \in \mathrm{Hom}\,(\Gamma, M)$ be the homomorphism corresponding to the pair $(u - c, u - b)$. Clearly, $\phi_{(u,v)}(\gamma) = \psi_{(u,v)}(\gamma) = 0$ for all $\gamma \in \mathrm{Gal}(\overline{K}/K(s, t))$

and $\psi_{(u,v)}(\sigma) = b$, $\psi_{(u,v)}(\tau) = c$. So it suffices to show that the abscissas of the points $(p,q)^\sigma - (p,q)$ and $(p,q)^\tau - (p,q)$ are b and c respectively.
Note that, by construction, we have

$$\sigma(r) = -r \ , \quad \sigma(s) = s \quad \text{and} \quad \sigma(t) = -t \ .$$

Then it easily follows that $(p,q)^\sigma \neq \pm(p,q)$. Denoting by m the abscissa of the point $(p,q)^\sigma - (p,q)$ and taking into account the group law algorithm given on p. 58 in [Sil85], we have

$$\begin{aligned} m &= \left(\tfrac{q+\sigma(q)}{\sigma(p)-p}\right)^2 + a + b + c - \sigma(p) - p \\ &= \left(\tfrac{q+\sigma(q)}{\sigma(p)-p}\right)^2 + 3u - r^2 - s^2 - t^2 - \sigma(p) - p \ . \end{aligned}$$

Since $q = w(p-u) + v$ and $p = rs - rt - st + u$, we can write

$$\begin{aligned} q + \sigma(q) &= w(p-u) + v + \sigma(w)\sigma(p-u) + v \\ &= w(p-u) + \sigma(w)\sigma(p-u) + 2v \\ &= (r+s-t)(rs-rt-st) + (-r+s+t)(-rs-rt+st) + 2rst \\ &= 2r^2s - 4rst + 2st^2 \\ &= 2s(r-t)^2 \ , \end{aligned}$$

and

$$\sigma(p) - p = -rs - rt + st - rs + rt + st = 2s(t-r) \ .$$

Thus, we obtain

$$\begin{aligned} m &= \left(\tfrac{(2s(r-t)^2}{2s(t-r)}\right)^2 + 3u - r^2 - s^2 - t^2 + 2rt - 2u \\ &= -s^2 + u \\ &= b \ . \end{aligned}$$

The equality $(p,q)^\tau - (p,q) = (c,0)$ is proved in exactly the same fashion. □

Proposition 3.4 $\epsilon(\operatorname{Im}\delta) = 0$.

Proof. Let $(u,v) \in E(K)$. Since $\kappa \circ \epsilon = \zeta$, we have $(\kappa \circ \epsilon)(\delta(u,v)) = 0$, i.e. the algebra $\epsilon(\delta(u,v))$ is numerical. We claim that this algebra is trivial. Indeed, we may assume that (u,v) is a point in $E(K)$ such that $u-b \neq 0$ and $u-c \neq 0$. Then the evaluation of the algebra

$$\epsilon(\delta(u,v)) = [u-c, x-b] + [u-b, x-c]$$

at the point (u,v) yields

$$[u-c, u-b] + [u-b, u-c] = 2[u-c, u-b] = 0 \ .$$

This implies that the algebra $\epsilon(\delta(u,v))$ is itself trivial, as required. □

Summarizing the above results, we obtain the following

Proposition 3.5 *Let E/K be a split elliptic curve over K, char $K \neq 2$. Let $\kappa : {}_2\mathrm{Br}\, E \to {}_2H^1(\Gamma, \overline{E})$ be the homomorphism described in Section 2 and let $\zeta : H^1(\Gamma, M) \to {}_2H^1(\Gamma, \overline{E})$ be the homomorphism induced by the embedding $M \subset \overline{E}$. Let also*

$$\epsilon : H^1(\Gamma, M) \longrightarrow {}_2\mathrm{Br}\, E$$

be the homomorphism defined by (9). Then

(i) $\kappa \circ \epsilon = \zeta$.

(ii) *There exists a unique homomorphism*

$$\varepsilon : {}_2H^1(\Gamma, \overline{E}) \longrightarrow {}_2\mathrm{Br}\, E$$

such that $\varepsilon \circ \zeta = \epsilon$ and $\kappa \circ \varepsilon = 1_{{}_2H^1(\Gamma, \overline{E})}$ is an identical map.

Proof. The equality $\kappa \circ \epsilon = \zeta$ is proved in Lemma 3.1. Since ζ is the cokernel of the homomorphism δ and, by Proposition 3.4, $\epsilon(\mathrm{Im}\, \delta) = 0$, there exists a unique homomorphism $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$, such that $\varepsilon \circ \zeta = \epsilon$. Since $\kappa \circ \varepsilon \circ \zeta = \kappa \circ \epsilon = \zeta$, we obtain that $\kappa \circ \varepsilon = 1_{{}_2H^1(\Gamma, \overline{E})}$ because ζ is an epimorphism. □

Reformulating the results of Proposition 3.5 in terms of central simple algebras and using Proposition 3.3, we obtain

Theorem 3.6 *Let E/K be a split elliptic curve defined by an affine equation*

$$y^2 = (x-a)(x-b)(x-c),$$

where $a, b, c \in K$ and char $K \neq 2$. Let $\epsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$ be the section for the homomorphism $\kappa : {}_2\mathrm{Br}\, E \to {}_2H^1(\Gamma, \overline{E})$ constructed in Proposition 3.5 and let $I = \mathrm{Im}\, \varepsilon$. Then

$${}_2\mathrm{Br}\, E = {}_2\mathrm{Br}\, K \oplus I$$

and every element in I is represented by a biquaternion algebra

$$(r, x-b) \otimes (s, x-c)$$

with $r, s \in K^*$. Conversely, every algebra of such a type is unramified over E. An algebra $A = (r, x-b) \otimes (s, x-c)$ is trivial in $I = \mathrm{Im}\, (\varepsilon)$ if and only if A is similar to an algebra of one of the three following types:

(i) *an algebra*

$$(u-c, x-b) \otimes (u-b, x-c),$$

where u is the abscissa of a point in $E(K)$ such that $u - b \neq 0$ and $u - c \neq 0$;

(ii) *an algebra*

$$(b-c, x-b) \otimes ((b-c)(b-a), x-c) \; ;$$

(iii) *an algebra*

$$((c-a)(c-b), x-b) \otimes (c-b, x-c) \; .$$

4 *Semisplit elliptic case*

Let E be a semisplit elliptic curve given by an affine equation

$$y^2 = (x-w)(x^2-d),$$

where $w, d \in K$, $char\,K \neq 2$ and d is not a square in K^*. Let $L = K(\sqrt{d})$, $\Gamma = \mathrm{Gal}\,(\overline{K}/K)$ and $\Lambda = \mathrm{Gal}\,(\overline{K}/L)$. Clearly, Λ is a subgroup of index two in Γ and

$$M \cong M_\Gamma^\Lambda(\mathbb{Z}/2),$$

where $M_\Gamma^\Lambda(\mathbb{Z}/2)$ is an induced Γ-module. Therefore, by Shapiro's lemma (see, for example, [Serre64]), we have

$$H^1(\Gamma, M) = H^1(\Gamma, M_\Gamma^\Lambda(\mathbb{Z}/2)) \cong H^1(\Lambda, \mathbb{Z}/2) \cong L^*/L^{*2}\ .$$

Let us consider the split elliptic curve $E_L = E \times_K L$ over L. Fixing its points $(b, 0)$, $(c, 0)$, where $b = \sqrt{d}$, $c = -\sqrt{d}$, we get the isomorphisms over L

$$M \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2, \qquad H^1(\Lambda, M) \cong L^*/L^{*2} \oplus L^*/L^{*2}\ .$$

Under these identifications the restriction map is given by the formula

$$\mathrm{res} : H^1(\Gamma, M) \to H^1(\Lambda, M), \quad l \in L^*/L^{*2} \to (l^\sigma, l) \in L^*/L^{*2} \oplus L^*/L^{*2} \tag{10}$$

where σ is the nontrivial automorphism L/K.

We denote the homomorphisms constructed in the previous section for the split curve E_L by the same symbols but equipped with the subscript L. Thus, we have the homomorphisms

$$\epsilon_L : H^1(\Lambda, M) \longrightarrow {}_2\mathrm{Br}\,(E_L)\ ,$$

$$\zeta_L : H^1(\Lambda, M) \longrightarrow {}_2H^1(\Lambda, \overline{E})$$

and

$$\varepsilon_L :\ {}_2H^1(\Lambda, \overline{E}) \longrightarrow {}_2\mathrm{Br}\,(E_L)\ .$$

Let

$$H^1(\Gamma, M) \cong L^*/L^{*2} \xrightarrow{\tau} L^*/L^{*2} \oplus L^*/L^{*2} \cong H^1(\Lambda, M)$$

be the homomorphism which takes l into the pair $(1, l)$. We define the homomorphism

$$\epsilon : H^1(\Gamma, M) \longrightarrow {}_2\mathrm{Br}\,E$$

by means of the following commutative diagram

$$\begin{tikzcd}
H^1(\Lambda, M) \arrow[r, "\epsilon_L"] & {}_2\mathrm{Br}\,(E_L) \arrow[d, "\mathrm{cor}"] \\
H^1(\Gamma, M) \arrow[u, "\tau"] \arrow[r, "\epsilon"] & {}_2\mathrm{Br}\,E
\end{tikzcd}$$

PROPOSITION 4.1 *Let E/K be a semisplit elliptic curve. Let $\zeta : H^1(\Gamma, M) \to {}_2H^1(\Gamma, \overline{E})$ be the homomorphism induced by the embedding $M \subset \overline{E}$ and let ϵ be the above homomorphism. Then there exists a homomorphism*

$$\varepsilon : {}_2H^1(\Gamma, \overline{E}) \longrightarrow {}_2\mathrm{Br}\, E$$

such that $\kappa \circ \varepsilon = 1_{{}_2H^1(\Gamma,\overline{E})}$ (i.e. ε is a section for the homomorphism κ) and $\varepsilon \circ \zeta = \epsilon$.

Proof. The proof is based on a diagram chase. We divide it into a sequence of simple observations.

LEMMA 4.2 *The restriction homomorphism*

$$H^1(\Gamma, M) \xrightarrow{\mathrm{res}} H^1(\Lambda, M)$$

is injective.

Proof. This easily follows from (10). □

LEMMA 4.3 *The composition*

$$H^1(\Gamma, M) \xrightarrow{\tau} H^1(\Lambda, M) \xrightarrow{\mathrm{cor}} H^1(\Gamma, M)$$

coincides with the identical map $1_{H^1(\Gamma,M)}$.

Proof. By Lemma 4.2, the homomorphism res : $H^1(\Gamma, M) \to H^1(\Lambda, M)$ is injective. Therefore, it is sufficient to prove that $\mathrm{res} \circ \mathrm{cor} \circ \tau = \mathrm{res}$. Let $l \in L^*$. Using (10) we have

$$(\mathrm{res} \circ \mathrm{cor} \circ \tau)(l) = (\mathrm{res} \circ \mathrm{cor})(1, l) = (1, l) + (1, l)^\sigma =$$
$$(1, l) + (l^\sigma, 1) = (l^\sigma, l) = \mathrm{res}(l) \ .$$

□

LEMMA 4.4 $\kappa \circ \epsilon = \zeta$.

Proof. The commutative diagram

$$\begin{array}{ccccc}
H^1(\Lambda, M) & \xrightarrow{\zeta_L} & {}_2H^1(\Lambda, \overline{E}) & \xleftarrow{\kappa_L} & {}_2\mathrm{Br}\,(E_L) \\
\downarrow{\scriptstyle \mathrm{cor}} & & \downarrow{\scriptstyle \mathrm{cor}} & & \downarrow{\scriptstyle \mathrm{cor}} \\
H^1(\Gamma, M) & \xrightarrow{\zeta} & {}_2H^1(\Gamma, \overline{E}) & \xleftarrow{\kappa} & {}_2\mathrm{Br}\, E
\end{array}$$

and Lemma 4.3 imply

$$\kappa \circ \epsilon = \kappa \circ \mathrm{cor} \circ \epsilon_L \circ \tau = \mathrm{cor} \circ \kappa_L \circ \varepsilon_L \circ \zeta_L \circ \tau = \mathrm{cor} \circ \zeta_L \circ \tau = \zeta \circ \mathrm{cor} \circ \tau = \zeta \ .$$

□

Lemma 4.5 $\mathrm{cor}\circ\zeta_L\circ\tau=\zeta\,.$

Proof. Clearly, we have $\mathrm{cor}\circ\zeta_L=\zeta\circ\mathrm{cor}$. Multiplying from the right hand by τ we obtain that $\mathrm{cor}\circ\zeta_L\circ\tau=\zeta\circ\mathrm{cor}\circ\tau=\zeta$ (the last equality holds by Lemma 4.3). □

Lemma 4.6 $\epsilon\,(\mathrm{Im}\,\delta)\subset\mathrm{Im}\,\iota$.

Proof. By Lemma 4.4, we have $\kappa\circ\epsilon=\zeta$, hence

$$\epsilon\,(\mathrm{Im}\,\delta)=\epsilon\,(\ker\,\zeta)\subset\ker\,\kappa=\mathrm{Im}\,\iota\,.$$

□

Lemma 4.7 $\mathrm{Im}\,\epsilon\cap\mathrm{Im}\,\iota=0$.

Proof. Our computations are illustrated by the following commutative diagram

$$\begin{tikzcd}
 & {}_2H^1(\Lambda,\overline{E}) \arrow[dr, "\varepsilon_L"] & & \\
H^1(\Lambda,M) \arrow[ur, "\zeta_L"] \arrow[rr, "\epsilon_L"] & & {}_2\mathrm{Br}\,E_L \arrow[r, bend left, "\varsigma_L"] \arrow[d, "\mathrm{cor}"] & {}_2\mathrm{Br}\,L \arrow[l, "\iota_L"] \arrow[d, "\mathrm{cor}"] \\
H^1(\Gamma,M) \arrow[u, "\tau"] \arrow[rr, "\epsilon"] & & {}_2\mathrm{Br}\,E \arrow[r, bend left, "\varsigma"] & {}_2\mathrm{Br}\,K \arrow[l, "\iota"]
\end{tikzcd}$$

Let $b\in{}_2\mathrm{Br}\,E$ be such that $b=\epsilon(h)=\iota(a)$ for some $h\in H^1(\Gamma,M)$ and some $a\in{}_2\mathrm{Br}\,K$. Let $c=\zeta_L(\tau(h))$. Then

$$a=(\varsigma\circ\iota)(a)=\varsigma(b)=(\varsigma\circ\mathrm{cor}\circ\varepsilon_L)(c)=(\mathrm{cor}\circ\varsigma_L\circ\varepsilon_L)(c)=0,$$

because $\varsigma_L\circ\varepsilon_L=0$. □

Lemma 4.8 $\epsilon\,(\mathrm{Im}\,\delta)=0$.

Proof. By Lemmas 4.6 and 4.7, we have $\epsilon\,(\mathrm{Im}\,\delta)\subset\mathrm{Im}\,\epsilon\cap\mathrm{Im}\,\iota=0$. □

We are now in the position to finish the proof of Proposition 4.1. Since $\epsilon\,(\mathrm{Im}\,\delta)=\epsilon\,(\ker\,\zeta)=0$, it follows that there exists a unique homomorphism $\varepsilon:\,{}_2H^1(\Gamma,\overline{E})\to{}_2\mathrm{Br}\,E$ such that $\epsilon=\varepsilon\circ\zeta$. Furthermore,

$$\kappa\circ\varepsilon\circ\zeta=\kappa\circ\epsilon=\kappa\circ\mathrm{cor}\circ\epsilon_L\circ\tau=\kappa\circ\mathrm{cor}\circ\varepsilon_L\circ\zeta_L\circ\tau=$$

$$\mathrm{cor}\circ\kappa_L\circ\,\varepsilon_L\circ\zeta_L\circ\tau=\mathrm{cor}\circ\zeta_L\circ\tau=\zeta\circ\mathrm{cor}\circ\tau=\zeta\,.$$

Since ζ is an epimorphism, it follows that $\kappa \circ \varepsilon = 1_{\,{}_2H^1(\Gamma,\overline{E})}$. Proposition 4.1 is proved. $\square$

To reformulate the results of Proposition 4.1 in terms of central simple algebras we need three well-known lemmas which describe images of quaternion algebras under corestriction homomorphisms.

LEMMA 4.9 *Let F be a field and let P be a finite separable extension of F. Then for elements $a \in F$ and $b \in P$ we have*

$$\operatorname{cor}_{P/F}[a,b] = [a, N_{P/F}(b)]$$

in the Brauer group $\operatorname{Br} F$.

Proof. This is a well-known fact (see, for instance, [Serre79], p. 209). $\square$

LEMMA 4.10 *Let F be a field and let P be a quadratic extension of F. Suppose that $P = F(\sqrt{s})$, where $s \in F$. Then for elements $a, b \in F$ with the property $a + b \neq 0$ we have*

$$\operatorname{cor}_{P/F}[a+\sqrt{s}, b-\sqrt{s}] = [a+b, (a^2-s)(b^2-s)] \ .$$

Proof. Let

$$t = \frac{a+\sqrt{s}}{a+b} \quad \text{and} \quad l = \frac{b-\sqrt{s}}{a+b} \ .$$

Then $t + l = 1$, whence $[t,l] = [t, 1-t] = 0$ in $\operatorname{Br} P$. Substituting t and l, we have

$$0 = [t,l] = \left[\frac{a+\sqrt{s}}{a+b}, \frac{b-\sqrt{s}}{a+b}\right] =$$

$$= [a+\sqrt{s}, b-\sqrt{s}] + [a+b, b-\sqrt{s}] + [a+\sqrt{s}, a+b] + [a+b, a+b].$$

Taking $\operatorname{cor}_{P/F}$ and using Lemma 4.9 we obtain that

$$0 = \operatorname{cor}_{P/F}[a+\sqrt{s}, b-\sqrt{s}] + [a+b, b^2-s] + [a^2-s, a+b] + [a+b, (a+b)^2] \ .$$

Therefore,

$$\operatorname{cor}_{P/F}[a+\sqrt{s}, b-\sqrt{s}] = [a+b, b^2-s] + [a^2-s, a+b] \ .$$

$\square$

LEMMA 4.11 *Let F be a field and let $P = F(\sqrt{s})$ be a quadratic extension of F. Let $u_1, v_1, u_2, v_2 \in F$ be such that $v_1 \neq 0$, $v_2 \neq 0$ and $v_1u_2 \neq u_1v_2$. Then*

$$\operatorname{cor}_{P/F}[u_1 + v_1\sqrt{s}, u_2 + v_2\sqrt{s}] =$$

$$[v_1, u_1^2 - v_1^2 s] + [-v_2, u_2^2 - v_2^2 s] + [v_1u_2 - u_1v_2, (u_1^2 - v_1^2 s)(u_2^2 - v_2^2 s)] \ .$$

Proof. Let
$$a = \frac{u_1}{v_1} \quad \text{and} \quad b = -\frac{u_2}{v_2}\ .$$
Then
$$[u_1 + v_1\sqrt{s}, u_2 + v_2\sqrt{s}] = [v_1(a + \sqrt{s}), -v_2(b - \sqrt{s})] =$$
$$= [v_1, -v_2] + [a + \sqrt{s}, b - \sqrt{s}] + [v_1, b - \sqrt{s}] + [a + \sqrt{s}, -v_2]\ .$$
Lemmas 4.10 and 4.9 give
$$\mathrm{cor}_{D/F}[u_1 + v_1\sqrt{s}, u_2 + v_2\sqrt{s}] =$$
$$[a + b, (a^2 - s)(b^2 - s)] + [v_1, b^2 - s] + [-v_2, a^2 - s]$$
and it remains to substitute $a = u_1/v_1$, $b = -u_2/v_2$. □

Theorem 4.12 *Let E be a semisplit elliptic curve over K, char $K \neq 2$, given by an affine equation $y^2 = (x - w)(x^2 - d)$, where $w, d \in K$ and d is not a square in K. Let $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$ be the section for the homomorphism $\kappa : {}_2\mathrm{Br}\, E \to {}_2H^1(\Gamma, \overline{E})$ constructed in Proposition 4.1 and let $I = \mathrm{Im}\, \varepsilon$. Then*
$${}_2\mathrm{Br}\, E \cong {}_2\mathrm{Br}\, K \oplus I$$
and every element in I is represented by either a quaternion algebra
$$(r, x - w),$$
where $r \in K^*$, or a biquaternion algebra
$$(t, r^2 - t^2d) \otimes (tx + r, (r^2 - t^2d)(x^2 - d))$$
where $r, t \in K$ and $t \neq 0$. Conversely, every algebra of the above types is unramified over E. It is trivial in I if and only if it is similar to a quaternion algebra
$$(x + u, (u - w)(x - w)),$$
where u is the abscissa of a point in $E(K)$.

Proof. The first statement is trivial because ε is a section for the homomorphism κ. To prove the second one we have to compute $\epsilon(h)$ in terms of quaternion algebras for all $h \in H^1(\Gamma, M)$.
By definition, $\epsilon = \mathrm{cor} \circ \epsilon_L \circ \tau$, where $L = K(\sqrt{d})$. Recall that we identify $L^*/L^{*2} \cong H^1(\Gamma, M)$ and $L^*/L^{*2} \oplus L^*/L^{*2} \cong H^1(\Lambda, M)$ and that $\tau : L^*/L^{*2} \to L^*/L^{*2} \oplus L^*/L^{*2}$ takes $l \in L^*/L^{*2}$ into $(1, l)$. Let $l \in L^*$. Then we have
$$(\mathrm{cor} \circ \epsilon_L \circ \tau)\,(l) = (\mathrm{cor} \circ \epsilon_L)\,(1, l) = \mathrm{cor}_{L(E)/K(E)}\,[l, x - \sqrt{d}]\ .$$
Let $l = r + t\sqrt{d}$. If $t = 0$, then, by Lemma 4.9, we have
$$\mathrm{cor}_{L(E)/K(E)}\,[r, x - \sqrt{d}] = [r, x^2 - d] = [r, x - w]\ .$$

If $t \neq 0$, then, by Lemma 4.11, we have

$$\operatorname{cor}_{L(E)/K(E)}[r + t\sqrt{d}, x - \sqrt{d}] =$$

$$[t, r^2 - t^2 d] + [1, x^2 - d] + [tx + r, (r^2 - t^2 d)(x^2 - d)] =$$
$$[t, r^2 - t^2 d] + [tx + r, (r^2 - t^2 d)(x^2 - d)]\ .$$

It remains to find out when an algebra $b \in I = \operatorname{Im} \epsilon$ is trivial. Let $b = \epsilon\,(l)$. By Proposition 4.1, we have $\epsilon = \varepsilon \circ \zeta$ and $\ker \varepsilon = 0$. So b is trivial if and only if $l \in \ker \zeta = \operatorname{Im} \delta$.

Let $(u, v) \in E(K)$ and $l = \delta(u, v)$. The commutative square

$$\begin{tikzcd} E(L)/2 \arrow[r, hook, "\delta_L"] & L^*/L^{*2} \oplus L^*/L^{*2} \\ E(K)/2 \arrow[u, "\mathrm{res}"] \arrow[r, hook, "\delta"] & L^*/L^{*2} \arrow[u, "\mathrm{res}"'] \end{tikzcd}$$

shows that

$$(l^\sigma, l) = \operatorname{res}(l) = (\operatorname{res} \circ \delta)(u, v) = (\delta_L \circ \operatorname{res})\,(u, v) = \delta_L(u, v),$$

where σ is a unique nontrivial automorphism L/K. Proposition 3.3 gives

$$\delta_L(u, v) = (u + \sqrt{d}, u - \sqrt{d}).$$

Thus, $l = u - \sqrt{d}$ and finally we get

$$\begin{array}{rcl} (\epsilon \circ\ \delta)\,(u, v) & = & (\operatorname{cor}_{L/K} \circ \epsilon_L \circ\ \tau)\,(l) \\ & = & (\operatorname{cor}_{L/K} \circ \epsilon_L)\,(1, l) \\ & = & \operatorname{cor}_{L/K}[u - \sqrt{d}, x + \sqrt{d}] \\ & = & [x + u, (u^2 - d)(x^2 - d)] \\ & = & [x + u, (u - w)(x - w)]\ . \end{array}$$

The theorem is proved. $\square$

To consider the non-split case it is convenient to have a reformulation of the last theorem without conditions on the equation of E. Let E be a semisplit elliptic curve given by an affine equation

$$y^2 = (x - a)g(x),$$

where $a \in K$ and $g(x)$ is a unitary irreducible polynomial over K. Denote the roots of $g(x)$ by b and c. Let also E' be a semisplit elliptic curve given by an equation

$$y^2 = (x - w)(x^2 - d),$$

where

$$w = a - \frac{b+c}{2} \quad \text{and} \quad d = \frac{(b-c)^2}{4} \ .$$

Clearly, the map

$$E \longrightarrow E'$$
$$(u,v) \mapsto (u - \frac{b+c}{2}, v)$$

is an isomorphism of elliptic curves. It induces the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & {}_2\mathrm{Br}\,K \arrow[r] \arrow[d, equal] & {}_2\mathrm{Br}\,E \arrow[r, "\kappa"] & {}_2H^1(\Gamma, \overline{E}) \arrow[r] & 0 \\
0 \arrow[r] & {}_2\mathrm{Br}\,K \arrow[r] & {}_2\mathrm{Br}\,E' \arrow[r, "\kappa'"] \arrow[u, "\cong"] & {}_2H^1(\Gamma, \overline{E}') \arrow[r] \arrow[u, "\cong"] & 0
\end{tikzcd}$$

Let $\varepsilon' : {}_2H^1(\Gamma, \overline{E}') \to {}_2\mathrm{Br}\,E'$ be the section for the homomorphism $\kappa' : {}_2\mathrm{Br}\,E' \to {}_2H^1(\Gamma, \overline{E}')$ described in Proposition 4.1. Let $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\,E$ be the section for the homomorphism $\kappa : {}_2\mathrm{Br}\,E \to {}_2H^1(\Gamma, \overline{E})$ defined by the following commutative square

$$\begin{tikzcd}
{}_2\mathrm{Br}\,E & {}_2H^1(\Gamma, \overline{E}) \arrow[l, "\varepsilon"'] \\
{}_2\mathrm{Br}\,E' \arrow[u, "\cong"] & {}_2H^1(\Gamma, \overline{E}') \arrow[l, "\varepsilon'"] \arrow[u, "\cong"]
\end{tikzcd}$$

Theorem 4.13 *Let E be a semisplit elliptic curve defined by an equation*

$$y^2 = (x-a)g(x),$$

where $a \in K$, $g(x)$ is a unitary irreducible quadratic polynomial over K and $g(x) = (x-b)(x-c)$ over $\overline{K}$. Let $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\,E$ be the section for the homomorphism $\kappa : {}_2\mathrm{Br}\,E \to {}_2H^1(\Gamma, \overline{E})$ defined above and let $I = \mathrm{Im}\,\varepsilon$. Then

$${}_2\mathrm{Br}\,E \cong {}_2\mathrm{Br}\,K \oplus I$$

and every element in I is represented by either a quaternion algebra of the form

$$(r, x-a),$$

where $r \in K^*$, or a biquaternion algebra of the form

$$(t, r^2 - h^2t^2) \otimes (t(x-h) + r, (r^2 - t^2h^2)g(x)),$$

where $h = (b+c)/2 \in K$, $r, t \in K$ and $t \neq 0$. Conversely, every algebra of the above types is unramified over E. It is trivial in I if and only if it is similar to a quaternion algebra

$$(x - h + u, (u + h - a)(x - a)) ,$$

where u is the abscissa of a point in $E(K)$.

Proof. All statements follow from Theorem 4.12. □

5 *Non-split elliptic case*

In this section we consider a non-split elliptic curve E given by an affine equation

$$y^2 = f(x),$$

where $f(x)$ is an irreducible unitary polynomial without multiple roots. Let a be a root of $f(x)$. We define $L = K(a)$ and $\Theta = \mathrm{Gal}(\overline{K}/L)$.
By construction, the curve $E_L = E \times_K L$ is either split or semisplit over L. Let

$$\zeta_L : H^1(\Theta, M) \longrightarrow {}_2H^1(\Theta, \overline{E})$$

be the homomorphism induced by the embedding $M \subset \overline{E}$ and let

$$\kappa_L : {}_2\mathrm{Br}\, E_L \longrightarrow {}_2H^1(\Theta, \overline{E})$$

be the homomorphism defined either in Section 3 or 4. Let also

$$\epsilon_L : H^1(\Theta, M) \longrightarrow {}_2\mathrm{Br}\, E_L$$

be the homomorphism defined either by formula (9) in the split case or by means of the homomorphism τ in the semisplit case (see Section 4).
According to Propositions 3.5 and 4.1 there exists a section

$$\varepsilon_L : {}_2H^1(\Theta, \overline{E}) \longrightarrow {}_2\mathrm{Br}\, E_L$$

for the homomorphism κ_L, such that the composition $\varepsilon_L \circ \zeta_L$ coincides with ϵ_L. We are now in the position to prove the existence of ϵ and ε with the same properties for the curve E/K in the non-split case.

Proposition 5.1 *Let E be a non-split elliptic curve over K, char $K \neq 2$. Let $\kappa : {}_2\mathrm{Br}\, E \to {}_2H^1(\Gamma, \overline{E})$ be the homomorphism defined in Section 2 and let $\zeta : H^1(\Gamma, M) \to {}_2H^1(\Gamma, \overline{E})$ be the homomorphism induced by the embedding $M \subset \overline{E}$. Let also ϵ be the composition*

$$\epsilon : H^1(\Gamma, M) \xrightarrow{\mathrm{res}} H^1(\Theta, M) \xrightarrow{\epsilon_L} {}_2\mathrm{Br}\, E_L \xrightarrow{\mathrm{cor}} {}_2\mathrm{Br}\, E$$

where ϵ_L is as above. Then there exists a homomorphism $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \longrightarrow {}_2\mathrm{Br}\, E$ such that $\varepsilon \circ \zeta = \epsilon$ and $\kappa \circ \varepsilon = 1_{{}_2H^1(\Gamma, \overline{E})}$ is the identical map.

Proof. This is entirely analogous to the proof of Proposition 4.1. The only difference is that instead of τ we have to use the homomorphism $H^1(\Gamma, M) \stackrel{\text{res}}{\to} H^1(\Theta, M)$. □

Keeping the above notation we may reformulate Proposition 5.1 in terms of central simple algebras. We distinguish two cases.

Theorem 5.2 *Suppose that the curve E_L is split. Let $f(x) = (x-a)(x-b)(x-c)$, where $a, b, c \in L = K(a)$. Let $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\mathrm{Br}\, E$ be the section for the homomorphism κ described in Proposition 5.1 and $I = \mathrm{Im}\, \varepsilon$. Then*

$$ {}_2\mathrm{Br}\, E \cong {}_2\mathrm{Br}\, K \oplus I $$

and any element in I has the form

$$ \mathrm{cor}_{L/K}\, [(r, x-b) \otimes (s, x-c)] $$

where $r, s \in L^*$. Conversely, any such a class of algebras is unramified over $K(E)$ and it is trivial in I if and only if it coincides with a class

$$ \mathrm{cor}_{L/K}\, [(u-c, x-b) \otimes (u-b, x-c)], $$

where u is the abscissa of a point in $E(K)$.

Proof. Since ϵ is the composition

$$ H^1(\Gamma, M) \xrightarrow{\text{res}} H^1(\Theta, M) \xrightarrow{\epsilon_L} {}_2\mathrm{Br}\, E_L \xrightarrow{\text{cor}} {}_2\mathrm{Br}\, E\ , $$

it follows that ε is the composition

$$ {}_2H^1(\Gamma, \overline{E}) \xrightarrow{\text{res}} {}_2H^1(\Theta, \overline{E}) \xrightarrow{\varepsilon_L} {}_2\mathrm{Br}\, E_L \xrightarrow{\text{cor}} {}_2\mathrm{Br}\, E $$

(an easy diagram chase). Hence

$$ I = \mathrm{Im}\, \varepsilon = \mathrm{cor}(\mathrm{Im}\, \varepsilon_L)\ . $$

According to Theorem 3.6 any element in $\mathrm{Im}\, \varepsilon_L$ is represented by an algebra of type $(r, x-b) \otimes (s, x-c)$ where $r, s \in L^*$. Hence an element in I has the form $\mathrm{cor}_{L/K}\, [(r, x-b) \otimes (s, x-c)]$ for some $r, s \in L^*$.
Let $r, s \in L^*$. Consider the algebra $(r, x-b) \otimes (s, x-c)$ over $L(E)$. It is unramified because its class lies in the image of the homomorphism ϵ_L. Therefore the class

$$ \alpha = \mathrm{cor}_{L/K}\, [(r, x-b) \otimes (s, x-c)] \in \mathrm{Br}\, K(E) $$

is also unramified. Assume that $\alpha \in I$. If

$$ \alpha = \mathrm{cor}_{L/K}\, [(u-c, x-b) \otimes (u-b, x-c)] $$

where u is the abscissa of a point in $E(K)$, then $\alpha = 0$ because

$$ [(u-c, x-b) \otimes (u-b, x-c)] = 0 $$

in $\operatorname{Im}\varepsilon_L$, by Theorem 3.6. Conversely, if $\alpha = 0$ in I then α grows up (via ζ and ε) from the image of the connecting homomorphism δ. By the construction all homomorphisms δ, ζ, ε commute with restriction homomorphisms. It follows that α is equal to a class of algebras coming from $E(K)/2$, that is of type $\operatorname{cor}_{L/K}[(u-c, x-b)\otimes(u-b, x-c)]$ where u is the abscissa of a point in $E(K)$. $\square$

THEOREM 5.3 *Suppose that the curve E_L is semisplit. Let $f(x) = (x-a)g(x)$, where $a \in L$, $g(x)$ is an irreducible quadratic polynomial over L and $g(x) = (x-b)(x-c)$ over $\overline{K}$. Let $\varepsilon : {}_2H^1(\Gamma, \overline{E}) \to {}_2\operatorname{Br} E$ be the section for the homomorphism κ described in Proposition 5.1 and $I = \operatorname{Im}\varepsilon$. Then*

$$ {}_2\operatorname{Br} E \cong {}_2\operatorname{Br} K \oplus I $$

and every element in I is represented either by a class

$$ \operatorname{cor}_{L/K}[r, x-a], $$

where $r \in L^*$, or a class of the form

$$ \operatorname{cor}_{L/K}\left[(t, r^2 - h^2t^2)\otimes(t(x-h)+r, (r^2-t^2h^2)g(x))\right] $$

where $h = (b+c)/2 \in L$, $r, t \in L$ and $t \neq 0$. Conversely, every such a class is unramified over $K(E)$. It is trivial in I if and only if it coincides with a class

$$ \operatorname{cor}_{L/K}[x-h+u, (u+h-a)(x-a)] $$

where u is the abscissa of a point in $E(K)$.

Proof. The proof is similar to that of Theorem 5.2. The difference is just that we use Proposition 4.13 instead of Proposition 3.6. Indeed, we have $I = \operatorname{cor}(\operatorname{Im}\varepsilon_L)$. According to Theorem 4.13 any element in $\operatorname{Im}\varepsilon_L$ is represented by either a quaternion algebra of the form $A = (r, x-a)$, where $r \in K^*$, or a biquaternion algebra of the form

$$ B = (t, r^2 - h^2t^2)\otimes(t(x-h)+r, (r^2-t^2h^2)g(x)), $$

where $h = (b+c)/2 \in K$, $r, t \in K$ and $t \neq 0$. Therefore an element in I is equal to either $\operatorname{cor}_{L/K}[A]$ or $\operatorname{cor}_{L/K}[B]$.
An algebra of the types A or B lies in $\operatorname{Im}\varepsilon_L$ and hence it is unramified. Therefore, classes $\operatorname{cor}_{L/K}[A]$ and $\operatorname{cor}_{L/K}[B]$ are also unramified. They are trivial in I if and only if they come from the image of the connecting homomorphism δ via the homomorphisms ζ and ε. Since δ, ζ and ε commute with the corresponding restriction homomorphisms, it follows (using the second assertion of Proposition 4.13) that the classes $\operatorname{cor}_{L/K}[A]$ and $\operatorname{cor}_{L/K}[B]$ are trivial in I if and only if they coincide with a class

$$ \operatorname{cor}_{L/K}[x-h+u, (u+h-a)(x-a)], $$

where u is the abscissa of a point in $E(K)$. □

The generators of ${}_2\mathrm{Br}\, E$ given in Theorems 5.2 and 5.3 are represented as classes $\mathrm{cor}_{L/K}[A]$, where A is a quaternion or biquaternion algebra over the cubic extension $L(E)/K(E)$. We close this section by showing how one can rewrite these generators as tensor products of quaternion algebras defined over $K(E)$.

Let P/K be a cubic extension and let $P = K(s)$ for some element $s \in P$.

Lemma 5.4 *Every element $a \in P$ can be written in the form*

$$a = \frac{\theta_1 + \theta_2 s}{\theta_3 + \theta_4 s}\ ,$$

where $\theta_1, \theta_2, \theta_3, \theta_4 \in K$.

Proof. Let $V = \{\theta_1 + \theta_2 s \mid \theta_1, \theta_2 \in K\}$ be a two-dimensional vector space over F. Since aV is also a two-dimensional vector space over K, the intersection $V \cap aV$ has dimension at least one. Let $b \in V \cap aV$ be a non-zero element. Then there exists $\theta_1, \theta_2, \theta_3, \theta_4 \in K$ such that

$$b = \theta_1 + \theta_2 s = (\theta_3 + \theta_4 s)a.$$

It follows that

$$a = \frac{\theta_1 + \theta_2 s}{\theta_3 + \theta_4 s}\ ,$$

as required. □

Lemma 5.5 *Let $a, b \in K$ be such that $a + b \neq 0$. Then*

$$\mathrm{cor}_{P/K}[a+s, b-s] = \big[a+b, (a+b)\, N_{P/K}((a+s)(b-s))\big]\ .$$

Proof. Let

$$t = \frac{a+s}{a+b} \quad \text{and} \quad l = \frac{b-s}{a+b}\ .$$

Then $t + l = 1$, whence $[t, l] = [t, 1-t] = 0$ in $\mathrm{Br}\, P$. Substituting t, l, we have

$$0 = [t, l] = \left[\frac{a+s}{a+b}, \frac{b-s}{a+b}\right] =$$

$$[a+s, b-s] + [a+b, b-s] + [a+s, a+b] + [a+b, a+b]\ .$$

Taking $\mathrm{cor}_{P/F}$ and using Lemma 4.9 we obtain that

$$0 = \mathrm{cor}_{P/K}[a+s, b-s] + [a+b, N_{P/K}(b-s)] +$$

$$[N_{P/K}(a+s), a+b] + [a+b, (a+b)^3]\ .$$

Therefore,

$$\mathrm{cor}_{P/F}[a+s, b-s] = [a+b, N_{P/K}(b-s)] + [N_{P/K}(a+s), a+b] + [a+b, a+b]\ ,$$

as required. □

Lemma 5.6 *Let* $u_1, v_1, u_2, v_2 \in K$*,* $v_1 \neq 0$*,* $v_2 \neq 0$ *and* $v_1u_2 \neq u_1v_2$*. Then*

$$\operatorname{cor}_{P/K}[u_1 + v_1 s, u_2 + v_2 s] = \left[v_1(v_1u_2 - u_1v_2), N_{P/K}(u_1 + v_1 s)\right] +$$

$$\left[v_2(u_1v_2 - v_1u_2), v_1(v_1u_2 - u_1v_2)\, N_{P/K}(u_2 + v_2 s)\right] \ .$$

Proof. This is entirely analogous to the proof of Lemma 4.11 and so we omit the details to the reader. □

Using Lemmas 5.4, 4.9, 5.5 and 5.6 one can easily produce explicit formulas to compute all algebras in Theorems 5.2 and 5.3. However we do not present them because of their bulk.

6 *Elliptic curves over local fields*

In the next few sections we demonstrate the efficiency of the above cohomological methods by considering an elliptic curve E defined over a local non-dyadic field K. To get an explicit description of ${}_2\mathrm{Br}\, E$, by Theorems 3.6, 4.13, 5.2 and 5.3, we only need to explicitly describe all relations between the generators indicated in these theorems which is equivalent to the description of the image of the boundary map $\delta : E(K)/2 \to H^1(\Gamma, M)$.
For an elliptic curve over local fields there is a natural p-adic filtration on the group of K-points with finite quotients. Examining each quotient individually one can very quickly find generators for the group $E(K)/2$. This leads in turn to the required description of $\operatorname{Im}\delta$. All necessary facts for our further argument can be easily elicited from standard textbooks, for example from [Hu87] and [Sil85]. For the convenience of the reader we start with recalling them.

For the rest of the paper we use the following specific notation:

K – a local non-dyadic field, i.e. a finite extension of the p-adic field $\mathbb{Q}_p$, $p \neq 2$;
v – the discrete valuation on K;
$\mathcal{O} = \mathcal{O}_K$ – the ring of integers of K;
$\mathcal{O}^* = \mathcal{O}_K^*$ – the unit group of O;
$\alpha = \alpha_K \in \mathcal{O}^*$ – a non-square element;
$\pi = \pi_K$ – a uniformizer for O;
$k = \mathcal{O}/\pi\,\mathcal{O}$ – the residue field of K.

Theorem 6.1 *There is a natural isomorphism*

$$H^1(\Gamma, \overline{E}) \cong \mathrm{Hom}_{cont}(E(K), \mathbb{Q}/\mathbb{Z}) \ .$$

Proof. See [Tate57] or [Mi86]. □

Corollary 6.2 $|\,{}_2\mathrm{Br}\, E| = 2 \cdot \sqrt{|H^1(\Gamma, M)|}$.

Proof. By Theorem 6.1, we have

$$|\,{}_2H^1(\Gamma,\overline{E})| = |\,{}_2\mathrm{Hom}_{cont}(E(K),\mathbb{Q}/\mathbb{Z})| =$$

$$|\mathrm{Hom}_{cont}(E(K)/2,\mathbb{Q}/\mathbb{Z})| = |E(K)/2|\;.$$

On the other hand, sequence (8) shows that

$$|\,{}_2H^1(\Gamma,\overline{E})| = |H^1(\Gamma,M)|/|E(K)/2|\;.$$

Therefore,

$$|E(K)/2|^2 = |H^1(\Gamma,M)|$$

and the result follows. □

Proposition 6.3 *Let n be a natural number. Then*

$$|E(K)/nE(K)| = |{}_nE(K)|\cdot|\mathcal{O}/n\mathcal{O}|.$$

Proof. See, for example, [Mi86], p. 52. □

Corollary 6.4 *Let E be a non-split elliptic curve defined over a local non-dyadic field K. Then ${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K$.*

Proof. Clearly, we have

$$|\,{}_2\mathrm{Br}\,E\,| = |\,{}_2\mathrm{Br}\,K\,|\cdot|\,{}_2H^1(\Gamma,\overline{E})\,| = |\,{}_2\mathrm{Br}\,K\,|\cdot|E(K)/2\,|\;.$$

Since E is non-split, it follows that every nontrivial element from M is not defined over K. Therefore, ${}_2E(K) = 0$ and, by Proposition 6.3, we obtain that $E(K)/2 = 0$. This implies that $|\,{}_2\mathrm{Br}\,E\,| = |\,{}_2\mathrm{Br}\,K\,|$, as required. □

Let E be an elliptic curve over K and let

$$y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

be a Weierstrass equation for the curve E/K with all coefficients $a_i \in \mathcal{O}$. Since its discriminant Δ is also an integer and since v is discrete we can look for an equation with $v(\Delta)$ as small as possible. A Weierstrass equation is called a *minimal* equation for E if $v(\Delta)$ is minimized subject to the condition a_1, a_2, a_3, a_4, $a_6 \in \mathcal{O}$.

It is known (see [Sil85], Proposition 1.3, p. 172) that a minimal (Weierstrass) equation is unique up to a change of coordinates

$$x = u^2x' + r\;,\qquad y = u^3y' + u^2sx' + t$$

with $u \in \mathcal{O}^*$ and r, s, $t \in \mathcal{O}$. Since, by our assumption, $2 \in \mathcal{O}^*$, a coordinate change $y \to y' = y + (a_1x + a_3)/2$ shows that we may always assume that $a_1 = a_3 = 0$, i.e. E is given by a minimal equation of the form

$$y^2 = x^3 + a_2x^2 + a_4x + a_6\;. \tag{11}$$

Later we need to know when (11) is a minimal equation for E. Let

$$b_2 = 4a_2 \ , \qquad b_4 = 2a_4 \ , \qquad b_6 = 4a_6 \ , \qquad b_8 = 4a_2a_6 - a_4^2 \ ,$$

$$c_4 = b_2^2 - 24b_4 \ , \qquad c_6 = b_2^3 + 36b_2b_4 - 216b_6$$

be the usual combinations of the a_i's and let

$$\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6$$

be the discriminant of equation (11) (see [Sil85], p. 46).

Proposition 6.5 *Equation (11) with integer coefficients a_2, a_4, a_6 is minimal if and only if either $v(\Delta) < 12$ or $v(c_4) < 4$.*

Proof. See [Sil85], page 186, Exercises 7.1. □

We assume that our elliptic curve E is given by a minimal equation (11). Reducing its coefficients modulo π we obtain the curve (possibly singular) $\widetilde{E}$ over k:

$$y^2 = x^3 + \tilde{a}_2x^2 + \tilde{a}_4x + \tilde{a}_6 \ .$$

The curve $\widetilde{E}$ is called the *reduction* of E modulo π.
Next let $P \in E(K)$. We can find homogeneous coordinates $P = [x_0, y_0, z_0]$ with integers x_0, y_0, z_0 such that at least one of them is in $\mathcal{O}^*$. Then the reduced point $\widetilde{P} = [\tilde{x}_0, \tilde{y}_0, \tilde{z}_0]$ is in $\widetilde{E}$. This gives a reduction map

$$E(K) \longrightarrow \widetilde{E}(k), \qquad P \longrightarrow \widetilde{P} \ .$$

Since the curve $\widetilde{E}$ can be singular, we denote its set of nonsingular points by $\widetilde{E}_{ns}(k)$ and we put

$$E_0(K) = \{P \in E(K) \mid \widetilde{P} \in \widetilde{E}_{ns}(k)\}$$

$$E_1(K) = \{P \in E(K) \mid \widetilde{P} = \widetilde{O}\} \ .$$

Proposition 6.6 *The following natural sequence of abelian groups*

$$0 \to E_1(K) \longrightarrow E_0(K) \longrightarrow \widetilde{E}_{ns}(k) \to 0$$

is exact.

Proof. See [Sil85], Proposition 2.1, p. 174. □

Proposition 6.7 *The group $E_1(K)$ is uniquely divisible by 2; in particular, we have $E_1(K) = 2E_1(K)$.*

Proof. See [Hu87], Corollary 1.3, p. 264. □

Let E/K be an elliptic curve and let $\widetilde{E}/k$ be the reduced curve for a minimal Weierstrass equation. One says that

(a) E has *good* reduction over K if $\widetilde{E}$ is nonsingular;

(b) E has *multiplicative* reduction over K if $\widetilde{E}$ has a node; in this case the reduction is said to be *split* (respectively *non-split*) if the slopes of the tangent lines at the node are in k (respectively not in k);

(c) E has *additive* reduction over K if $\widetilde{E}$ has a cusp.

Proposition 6.8 *Let E/K be an elliptic curve given by a minimal Weierstrass equation* (11).

(a) *E has good reduction if and only if $v(\Delta) = 0$;*
(b) *E has multiplicative reduction if and only if $v(\Delta) > 0$ and $v(c_4) = 0$;*
(c) *E has additive reduction if and only if $v(\Delta) > 0$ and $v(c_4) > 0$.*

Proof. See [Sil85], Proposition 5.1, p. 180. □

7 *Generators of $E(K)/2$ for a split elliptic curve over a local field*

Let E be a split elliptic curve given by a minimal equation (11). Since M is a trivial Γ-module, it follows that all roots of the cubic polynomial $f(x) = x^3 + a_2x^2 + a_4x + a_6$ are in K. Then these roots, clearly, belong to $\mathcal{O}$, so that we may assume that E is given by a minimal equation of the form

$$y^2 = (x-a)(x-b)(x-c) \tag{12}$$

with all a, b, c in $\mathcal{O}$. In this coordinate system M consists of the points

$$O, \quad P = (a,0), \quad Q = (b,0), \quad T = (c,0)\ .$$

Recall also that, by Proposition 6.3, we have $|E(K)/2| = |M| = 4$.

7.1 *Additive reduction*

Lemma 7.1 *The group $E_0(K)$ is divisible by* 2.

Proof. Since E has additive reduction, we have $E_0(K)/E_1(K) \cong k^+$; in particular the finite group $E_0(K)/E_1(K)$ is divisible by 2. Then the result follows from Proposition 6.7. □

Proposition 7.2 *The elements O, P, Q, T are representatives of $E(K)/2$.*

Proof. In view of Lemma 7.1 we have $E_0(K) \subset 2E(K) \subset E(K)$ and by [Sil85], Theorem 6.1, p. 183, the group $E(K)/E_0(K)$ is finite of order at most 4. Since $|E(K)/2| = 4$, we get $E_0(K) = 2E(K)$ and it remains to note that the points P, Q, T do not belong to $E_0(K)$. □

7.2 *Multiplicative reduction*

By our assumption, among the residues $\tilde{a}$, $\tilde{b}$, $\tilde{c}$ there are exactly two coinciding elements; say $\tilde{a} = \tilde{b}$. Changing coordinates, if necessary, we may assume that E is given by a minimal equation of the form

$$y^2 = x(x + \pi^m \beta)(x + \gamma)$$

with $\beta \in \mathcal{O}^*$, $m \geq 1$ and $\gamma \in \mathcal{O}^*$. Recall that in the case of non-split reduction γ coincides modulo squares with α; otherwise γ is a square in $\mathcal{O}^*$.

Lemma 7.3 *There exists a point $R_1 = (u, v) \in E_0(K)$ such that*

$$u = \alpha t^2\,, \quad u + \pi^m \beta = \alpha q^2\,, \quad u + \gamma = s^2\,, \quad v = \alpha\, t\, q\, s$$

with t, q, s in $\mathcal{O}^*$.

Proof. The proof is easy. Namely, we have to find a solution of the system

$$\begin{cases} \alpha x^2 + \pi^m \beta &= \alpha y^2 \\ \alpha x^2 + \gamma &= z^2 \end{cases}$$

According to standard facts from the theory of quadratic forms over finite and local fields the quadratic form $\tilde{\alpha} x^2 - z^2$ represents $-\gamma \in k^*$, whence, by the Hensel lemma, we can pick units t, $s \in \mathcal{O}^*$ satisfying the second equation. Substitute t in the first equation. Since the residues of the elements $\alpha t^2 + \pi^m \beta$ and α coincide modulo squares, again, applying the Hensel lemma we can find $q \in \mathcal{O}^*$ satisfying the equation $\alpha t^2 + \pi^m \beta = \alpha y^2$. $\square$

Remark 7.4 *Since the abscissa u of R_1 is not a square in K^*, Proposition 3.3 shows that $\delta(R_1) \neq (1, 1)$. Then it follows that $R_1 \notin 2E(K)$.*

Lemma 7.5 *There exists a point $R_2 = (u, v) \in E(K) \setminus E_0(K)$ with $u = \pi d$, $d \in \mathcal{O}$, and such that its image in the group $E(K)/E_0(K)$ is not divisible by* 2.

Proof. The abscissa of every point from $E(K) \setminus E_0(K)$ is of the form πd with $d \in \mathcal{O}$ because its residue is the node. Further, we have $\Delta = 16(\pi^m \beta \gamma (\pi^m \beta - \gamma))^2$ and $\pi^m \beta - \gamma \in \mathcal{O}^*$, so that $v(\Delta)$ is even. Then, by [Hu87], p. 266, the order of the finite group $E(K)/E_0(K)$ is divisible by 2, whence such a point exists. $\square$

Remark 7.6 *If the reduction is non-split, we can take $R_2 = (0, 0)$, because in this case the group $E(K)/E_0(K)$ has order 2 (loc. cit.) and, of course, $R_2 = (0, 0) \notin E_0(K)$.*

Proposition 7.7 *The points R_1, R_2 from the above two lemmas are generators of $E(K)/2E(K)$.*

Proof. Since $|E(K)/2| = 4$, we have $E(K)/2E(K) \cong \mathbb{Z}/2 \times \mathbb{Z}/2$. By our construction and by Remark 7.4, the images of R_1, R_2 in $E(K)/2E(K)$ are not trivial and they do not coincide. $\square$

8 *Generators of $E(K)/2$ for a semisplit elliptic curve over a local field*

We may assume that E is given by a minimal equation of the form

$$y^2 = (x-a)(x^2-d), \tag{13}$$

where $a, d \in \mathcal{O}$ and the polynomial $g(x) = x^2 - d$ is irreducible over K. Let $L = K(\sqrt{d})$ be its splitting field and let $\Lambda = \mathrm{Gal}\,(\overline{K}/L)$. As it was mentioned in Section 4, the module M is isomorphic to the induced module $M_\Gamma^\Lambda(\mathbb{Z}/2)$. This gives the isomorphisms

$$H^1(\Gamma, M) \cong L^*/L^{*2}\ , \qquad H^1(\Lambda, M) \cong L^*/L^{*2} \times L^*/L^{*2}\ .$$

Recall also that under this identification the restriction map $H^1(\Gamma, M) \to H^1(\Lambda, M)$ is given by the formula $l \to (l^\sigma, l)$, where $l \in L^*$ and σ is the nontrivial automorphism L/K; in particular, res is injective (see Section 4). It then follows from the commutative square

$$\begin{tikzcd}
E(L)/2 \arrow[r, hook, "\delta_L"] & L^*/L^{*2} \oplus L^*/L^{*2} \\
E(K)/2 \arrow[u, "\eta"] \arrow[r, hook, "\delta"] & L^*/L^{*2} \arrow[u, hook, "\mathrm{res}"']
\end{tikzcd}$$

that $\eta : E(K)/2 \to E(L)/2$ is also injective. Applying Proposition 6.3 we have $|E(K)/2| = |\,{}_2E(K)| = 2$. Now we want to explicitly describe the image $\eta(E(K)/2)$. The answer depends on the type of reduction.

8.1 *Multiplicative reduction.*

For an elliptic curve given by (13) one has $\Delta = 64d(a^2-d)^2$ and $c_4 = 16(a^2+3d)$. Since, by Proposition 6.8, $v(\Delta) > 0$ and $v(c_4) = 0$, we obtain that $v(d) > 0$ and $a \in \mathcal{O}^*$. Then, according to Proposition 6.5, (13) is a minimal equation for E_L. Hence E_L has multiplicative reduction (again by Proposition 6.8). Note that in view of $v(d) > 0$ and $a \in \mathcal{O}^*$ we have $a^2 - d \in O^*$, whence $v(\Delta) = v(d)$. We say that we are in case:

(M1) if either $v(d)$ is odd or 4 divides $v(d)$ and E has non-split multiplicative reduction;

(M2) if $v(d)$ is even and either E has split multiplicative reduction or 4 does not divide $v(d)$.

Proposition 8.1 *Let R_1, R_2 be the points in $E(L)$ introduced in 7.2. Then in case* (M1) *the nontrivial element of $\eta(E(K)/2)$ coincides with $R_1 + 2E(L)$ and in case* (M2) *it coincides with $R_2 + 2E(L)$.*

Proof. Consider case (M1). If $v(d)$ is odd, then by, [Hu87], p. 266, the group $E(K)/E_0(K)$ has an odd order. So we may choose a representative R of a unique nontrivial element in $E(K)/2$ among elements of $E_0(K)$. Since $E_0(K) \subset E_0(L)$ and η is injective, R coincides with R_1 modulo $2E(L)$.

Next suppose that 4 divides $v(d)$ and E has non-split multiplicative reduction. Since $v(d)$ is even, the extension L/K is unramified, so that $[k_L : k] = 2$, where k_L is the residue field of the local field L. It follows that E_L has split multiplicative reduction and, by [Hu87], p. 266, the group $E(L)/E_0(L)$ is cyclic of order $v(\Delta_L) = v(\Delta_K) = v(d)$; in particular, 4 divides $|E(L)/E_0(L)|$.

Let R be a representative of the nontrivial element of $E(K)/2$. Since E has non-split multiplicative reduction, it follows that $|E(K)/E_0(K)| = 2$ (loc. cit.), hence R can be chosen among elements $E(K)\backslash E_0(K)$. To show that $\eta(R)$ coincides with R_1 modulo $2E(L)$ consider the 2-Sylow subgroup G in $E(L)/E_0(L)$. It is clear that $R + E_0(L) \in G$ and it has order 2. Then $R + E_0(L)$ is divisible by 2 in G and so in $E(L)/E_0(L)$. But, by our construction (see Lemma 7.5), the element R_2 is not divisible by 2 in $E(L)/E_0(L)$, so we obtain $R+2E(L) \neq R_2+2E(L)$ and similarly we have $R+2E(L) \neq R_1+R_2+2E(L)$. It follows that $R + 2E(L) = R_1 + 2E(L)$, as required.

Consider case (M2). We have already mentioned that (13) is a minimal equation for E_L. It follows that $E_0(K) \subset E_0(L)$ and that the natural embedding $E(K) \subset E(L)$ induces the injection $\psi : E(K)/E_0(K) \to E(L)/E_0(L)$.

Suppose that E has split multiplicative reduction and $v(d)$ is even. Then L/K is unramified and again, by [Hu87], p. 266, the groups $E(K)/E_0(K)$ and $E(L)/E_0(L)$ are cyclic of the same order $v(\Delta) = v(\Delta_L) = v(d)$ implying ψ is a bijection. Since $v(d)$ is even, we can choose a representative R of the nontrivial element of $E(K)/2$ such that $R + E_0(K)$ is not divisible by 2 in $E(K)/E_0(K)$. Then it is not divisible by 2 in $E(L)/E_0(L)$; hence $R + 2E(L) = R_2 + 2E(L)$.

Suppose that E has non-split multiplicative reduction. Then according to [Hu87], p. 266, we have $|E(K)/E_0(K)| = 2$ and $|E(L)/E_0(L)| = v(d)$. Since 4 does not divide $v(d)$, the group $\psi\,(E(K)/E_0(K))$ is a 2-Sylow subgroup in $E(L)/E_0(L)$. Hence again picking an element R with the same property as above we easily get $R + 2E(L) = R_2 + 2E(L)$. □

8.2 *Additive reduction*

Proposition 8.2 (1) *If L/K is unramified, then $E(K)/2$ is generated by $P = (a, 0)$.*

(2) *Let L/K be ramified. If $a - \sqrt{d}$ is not a square in L^*, then $E(K)/2$ is again generated by $P = (a, 0)$. If $a - \sqrt{d} = s^2$, $s \in L^*$, then $E(K)/2$ is generated by the point $U = (u, w) \in E(K)$, where $u = N_{L/K}(s) + a$ and $w = N_{L/K}(s)\ Tr_{L/K}(s)$.*

Proof. First let L/K be unramified. Then E_L has additive reduction and by Proposition 7.2, we have $P \notin 2E(L)$. It follows that $P \notin 2E(K)$, as required.

Next let L/K be ramified. Recall that, by Lemma 7.1, we have $E_0(K) \subset 2E(K)$ and that $E(K)/E_0(K)$ is a group of order at most 4 (see [Sil85], p. 183).
If $a-\sqrt{d}$ is not a square in L^*, then, by Proposition 3.3, $\delta_L(P) \neq (1,1)$, hence $P \notin 2E(L)$ and the result follows.
Let $a-\sqrt{d}=s^2$, $s \in L^*$. Then it is easy to check that $2U=P$. This implies that $P \in 2E(K) \setminus E_0(K)$ and so $|2E(K)/E_0(K)| \geq 2$. But $|E(K)/2E(K)|=2$ and $|E(K)/E_0(K)| \leq 4$. It follows that $|2E(K)/E_0(K)|=2$, whence $U \notin 2E(K)$, as required. □

For the description of ${}_2\mathrm{Br}\,E$ we will also need to know whether $(\delta_L \circ \eta)\,(E(K)/2)$ belongs to the unramified part of the subset $\mathrm{res}\,(L^*/L^{*2}) \subset L^*/L^{*2} \times L^*/L^{*2}$. In other words, we will need to know whether $v_L(a+\sqrt{d})$ and $v_L(u+\sqrt{d})$ are odd or even. Here u is the abscissa of the above point U. It turns out that the answer depends on the coefficients of the minimal equation (13) only.
Let $a=\pi^m a'$, $d=\pi^{2k+\lambda}d'$ with $a', d' \in \mathcal{O}^*$ and $\lambda=0,1$. Using Propositions 6.5 and 6.8 one can easily make sure that $m>0$, $2k+\lambda>0$ and that $m=1$ or $2k+\lambda \leq 3$. We will say that we are in case:

($A1$) if one of the following conditions holds:
 (a) $\lambda=0$, i.e. L/K is unramified,
 (b) $\lambda=1$, $m=1$, $k=0$,
 (c) $\lambda=1$, $m>1$;

($A2$) if $\lambda=1$, $m=1$, $k \geq 1$ and $a-\sqrt{d} \notin L^{*2}$.

($A3$) if $\lambda=1$, $m=1$, $k \geq 1$ and $a-\sqrt{d} \in L^{*2}$,

Lemma 8.3 (i) *In case* (A1) *the group $E(K)/2$ is generated by P and $v_L(a+\sqrt{d})$ is odd.*
(ii) *In case* (A2) *the group $E(K)/2$ is generated by P and $v_L(a+\sqrt{d})$ is even.*
(iii) *In case* (A3) *the group $E(K)/2$ is generated by U and $v_L(u+\sqrt{d})$ is odd.*

Proof. First examine case ($A1$).
(a) Here L/K is unramified and at least one of the numbers k and m equals 1. So, obviously, $v_L(a+\sqrt{d})=1$.
(b) Since L/K is ramified, we have $v_L(a)=v_L(\pi)=2$ and $v_L(\sqrt{d})=1$. So $v_L(a+\sqrt{d})=1$.
(c) We have $v_L(a)=2m \geq 4$ and $v_L(\sqrt{d})=2k+1$. Since $2k+\lambda \leq 3$, we obtain that $v_L(a+\sqrt{d})=v_L(d)=2k+1$ is odd.
Case ($A2$). Since L/K is ramified, we have $v_L(a)=v_L(\pi)=2$ and $v_L(\sqrt{d})=2k+1 \geq 3$. It follows that $v_L(a+\sqrt{d})=2$.
Case ($A3$). Keeping the notation of Proposition 8.2 we have $a-\sqrt{d}=s^2$ and $u=N_{L/K}(s)+a$. It easily follows that $v_L(s)=1$. Further, letting σ be the nontrivial automorphism L/K we have

$$u+\sqrt{d}=N_{L/K}(s)+a+\sqrt{d}=ss^\sigma+s^\sigma s^\sigma=(s+s^\sigma)s^\sigma.$$

Therefore, $v_L(u+\sqrt{d})=v_L(s+s^\sigma)+1$ and it remains to note that $v_L(s+s^\sigma)$ is even because $s+s^\sigma \in K$. □

9 *Computing* ${}_2\mathrm{Br}\, E$ *over non-dyadic local fields: split case*

Putting together the results of the previous sections one can easily obtain an explicit and very short description of the 2-torsion subgroup of $\mathrm{Br}\, E$ for split and semisplit elliptic curves (note that for non-split curves it was done in Corollary 6.4). Namely, let $\delta : E(K)/2 \to H^1(\Gamma, M)$ be the boundary map. The description of generators of $E(K)/2$ and their images under the map δ given in Sections 7 and 8 enables us to explicitly construct a subgroup in $H^1(\Gamma, M)$ that complements $\delta(E(K)/2)$. If we then restrict the section $\epsilon : H^1(\Gamma, M) \to {}_2\mathrm{Br}\, E$ constructed in Sections 3 and 4 at this subgroup, we immediately obtain a description of the second summand in the decomposition ${}_2\mathrm{Br}\, E = {}_2\mathrm{Br}\, K \oplus \mathrm{Im}\, \epsilon$ as, by Proposition 3.4, and Lemma 4.8, the equality $\epsilon\, (\mathrm{Im}\, \delta) = 0$ holds.
In this section we consider a split elliptic curve E given by a minimal equation of the form

$$y^2 = x(x-b)(x-c)\ , \tag{14}$$

with b, c in the integer ring $\mathcal{O}$. Its 2-torsion consists of the points O, $P = (0,0)$, $Q = (b, 0)$ and $T = (c, 0)$. As in Section 3, we may identify

$$M = \langle Q\rangle \oplus \langle T\rangle \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2$$

and

$$H^1(\Gamma, M) \cong K^*/K^{*2} \oplus K^*/K^{*2}\ .$$

According to Proposition 3.3 the connecting homomorphism

$$\delta : E(K)/2 \hookrightarrow K^*/K^{*2} \oplus K^*/K^{*2}$$

is given by the formula

$$\delta(u, v) = \begin{cases} (u-c, u-b) & \text{if } u \neq b \text{ and } u \neq c, \\ (b-c, b(b-c)) & \text{if } u = b, \\ (c(c-b), c-b) & \text{if } u = c, \\ (1,1) & \text{if } u = \infty, \end{cases} \tag{15}$$

where $(u, v) \in E(K)$. Let

$$C_\alpha = [\alpha, x-c]\ , \quad C_\pi = [\pi, x-c]\ , \quad B_\alpha = [\alpha, x-b] \quad \text{and} \quad B_\pi = [\pi, x-b] \tag{16}$$

be the classes of quaternion algebras over $K(E)$. We distinguish the following three cases.

9.1 *Good reduction*

We start with the following

Lemma 9.1 *$\delta(E(K)/2)$ is generated by the pairs $(\alpha, 1)$ and $(1, \alpha)$.*

Proof. Let K^{nr}/K be a maximal unramified extension. It suffices to show that the images of our pairs under the natural map $\zeta : H^1(\Gamma, M) \to {}_2H^1(\Gamma, \overline{E})$ are trivial. To do so, first recall that, by [LT58] and [L56], we have

$$H^1(\mathrm{Gal}\,(K^{nr}/K), E(K^{nr})) = H^1(\mathrm{Gal}\,(\overline{k}/k), \widetilde{E}) = 0 \ .$$

This implies that $\mathrm{res} : H^1(\Gamma, \overline{E}) \to H^1(K^{nr}, \overline{E})$ is injective. On the other hand, obviously we have $(\mathrm{res} \circ \zeta)(\alpha, 1) = (\mathrm{res} \circ \zeta)(1, \alpha) = 1$, so the result follows. □

Proposition 9.2 *We have*

$${}_2\mathrm{Br}\, E = {}_2\mathrm{Br}\, K \oplus \{0, B_\pi, C_\pi, B_\pi + C_\pi\} \ .$$

Proof. It suffices to note that the subgroup generated by the pairs $(\pi, 1)$ and $(1, \pi)$ complements the subgroup $\delta(E(K)/2)$ and that ϵ takes these pairs to the classes B_π and C_π. □

9.2 *Additive reduction*

We may assume that $v(b) \geq 1$, $v(c) \geq 1$ and that at least one of these numbers is 1. Let $b = \pi^m d$ and $c = \pi e$, where d and e are units and $m \geq 1$. Proposition 7.2 shows that $E(K)/2$ is generated by the points P, Q, T. Applying (15) we get

Lemma 9.3 $\delta(E(K)/2)$ *is generated by the pairs*

$$\delta(P) = (-\pi e, -\pi^m d) \quad \textit{and} \quad \delta(T) = (\pi e(\pi e - \pi^m d), \pi e - \pi^m d) \ .$$

Proposition 9.4 *We have*

$${}_2\mathrm{Br}\, E = {}_2\mathrm{Br}\, K \oplus \{0, B_\alpha, C_\alpha, B_\alpha + C_\alpha\} \ .$$

Proof. It easily follows from Lemma 9.3 that the subgroup generated by the pairs $(\alpha, 1)$ and $(1, \alpha)$ complements $\delta(E(K)/2)$ in $K^*/K^{*2} \oplus K^*/K^{*2}$ and it remains to note that ϵ takes these pairs to the classes B_α and C_α. □

9.3 *Non-split multiplicative reduction*

We may assume that E is given by a minimal equation of the form

$$y^2 = x(x + \pi^m \beta)(x + \alpha) \ ,$$

with $m \geq 1$ and $\beta \in \mathcal{O}$. Note that in the notation of formulas (15) and (16) we have that

$$b = -\pi^m \beta \qquad \text{and} \qquad c = -\alpha \ .$$

Lemma 9.5 *$\delta(E(K)/2)$ is generated by the pairs $(1,\alpha)$ and $(\alpha,\pi^m\beta)$.*

Proof. Let R_1, R_2 be two points introduced in 7.2. It then follows from Lemma 7.3, Remark 7.6 and formula (15) that $\delta(R_1) = (1,\alpha)$ and $\delta(R_2) = (\alpha,\pi^m\beta)$, as required. □

Proposition 9.6 *We have*

$$ {}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0, B_\pi, C_\pi, B_\pi + C_\pi\} \, . $$

Proof. The subgroup generated by the pairs $(\pi,1)$ and $(1,\pi)$ complements $\delta(E(K)/2)$, so the result follows. □

9.4 *Split multiplicative reduction*

We may assume that E is given by a minimal equation of the form

$$ y^2 = x(x+\pi^m\beta)(x+1). $$

Lemma 9.7 *$\delta(E(K)/2)$ is generated by the pairs $(1,\alpha)$ and $(1,\pi)$.*

Proof. As above, we have $\delta(R_1) = (1,\alpha)$. Further, it follows from the construction that the abscissa of the point $R_2 = (u,v)$ is of the form $u = \pi d$. So applying formula (15), we obtain that $\delta(R_2) = (1,\pi u + \pi^m\beta)$. But $|\delta(E(K)/2)| = 4$, whence $v(\pi u + \pi^m\beta)$ is odd and the result follows. □

Proposition 9.8 *We have*

$$ {}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ B_\alpha,\ B_\pi,\ B_{\alpha\pi}\} \, . $$

Proof. This follows from the fact that the subgroup generated by the pairs $(\alpha,1)$ and $(\pi,1)$ complements $\delta(E(K)/2)$. □

10 *Computing ${}_2\mathrm{Br}\,E$ over non-dyadic local fields: semisplit case*

We keep the notation introduced in Section 8. Assume that E is given by a minimal equation of the form (13). Then $E(K)/2$ and $H^1(\Gamma,M)$ are groups of order 2 and 4 respectively, so that $\delta(E(K)/2)$ can be complemented inside $H^1(\Gamma,M)$ by a single element. We will find such an element among elements $\mathrm{cor}\,(H^1(\Lambda,M))$. Recall that δ_L denotes the homomorphism $E(L)/2 \hookrightarrow H^1(\Lambda,M)$.

Lemma 10.1 *Let $\theta \in H^1(\Lambda,M)$ satisfies the condition $(\mathrm{res}\circ\mathrm{cor})\,(\theta) \notin (\delta_L \circ \mathrm{res})\,(E(K)/2)$. Then $\mathrm{cor}\,(\theta)$ complements $\delta\,(E(K)/2)$.*

Proof. By our assumption,

$$\mathrm{res}\,(\mathrm{cor}\,(\theta)) \notin (\delta_L \circ \mathrm{res})\,(E(K)/2) = (\mathrm{res} \circ \delta)\,(E(K)/2)\ ,$$

so that $\mathrm{cor}\,(\theta)$ does not lie in $\delta(E(K)/2)$. □

Let α_L and π_L be a non-square unit and a uniformizer of the integer ring $\mathcal{O}_L$ of $L = K(\sqrt{d})$ respectively.

10.1 *Good Reduction*

Proposition 10.2 ${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ [\pi, x-a]\}$.

Proof. Clearly, $(\delta_L \circ res)\,(E(K)/2)$ belongs to the unramified part of $H^1(\Lambda, M) \cong L^*/L^{*2} \oplus L^*/L^{*2}$. Since we have good reduction, d is a unit, whence $\pi_L = \pi$. We put $\theta = (1, \pi)$. The equation $(\mathrm{res} \circ \mathrm{cor})(\theta) = (\pi, \pi)$ shows that θ satisfies the condition of Lemma 10.1. It then follows from Theorem 4.12 that ${}_2\mathrm{Br}\,E$ is generated by ${}_2\mathrm{Br}\,K$ and

$$(\mathrm{cor} \circ \epsilon_L)\,[1, \pi] = \mathrm{cor}\,[\pi, x + \sqrt{d}] = [\pi, x^2 - d] = [\pi, x - a].$$

□

10.2 *Additive reduction*

Proposition 10.3 (1) *In cases* (A1) *and* (A3) *we have*

$${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ \mathrm{cor}\,[\alpha_L, x - \sqrt{d}]\}.$$

(2) *In case* (A2) *we have*

$${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ \mathrm{cor}\,[\pi_L, x - \sqrt{d}]\}.$$

Proof. It suffices to note that, by Lemma 8.3, in the first (resp. second) case the pair $\theta = (1, \alpha_L)$ (resp. $\theta = (1, \pi_L)$) satisfies the condition of Lemma 10.1. □

10.3 *Multiplicative Reduction*

Proposition 10.4 *In case* (M1) *we have*

$${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ \mathrm{cor}\,[\pi_L, x - \sqrt{d}]\}.$$

and in case (M2) *we have*

$${}_2\mathrm{Br}\,E = {}_2\mathrm{Br}\,K \oplus \{0,\ \mathrm{cor}\,[\alpha_L, x - \sqrt{d}]\}\ .$$

Proof. Denote a representative of a unique nontrivial element in $E(K)/2E(K)$ by R. Consider first case (M1). Let L^{nr} be a maximal unramified extension of L. According to Proposition 8.1 we have $\eta(R) = R_1 + 2E(L)$. Since, by construction, $R_1 \in E_0(L)$ and $E_0(L^{nr})/2E_0(L^{nr}) = 0$ (see [Sil85], p. 187), it follows that $\delta_L(\eta(R))$ belongs to the unramified part of the group $H^1(\Lambda, M) \cong L^*/L^{*2} \oplus L^*/L^{*2}$. Therefore one can take $\theta = (1, \pi_L)$ and the result follows. In case (M2) we have $\eta(R) = R_2 + 2E(L)$. Since $v(d)$ is even, the extension L/K is unramified and E_L has split multiplicative reduction. We know that the abscissa u of R_2 is of the form $u = \pi u'$, so that $\delta_L(R_2) = (\pi u' + \sqrt{d}, \pi u' - \sqrt{d})$. It is easy to make sure that $v(\pi u' + \sqrt{d})$ is odd. Then $\theta = (1, \alpha_L)$ satisfies the condition of Lemma 10.1 and the result follows. $\square$

References

[AM72] Artin M., Mumford D.: *Some elementary examples of unirational varieties which are not rational.* Proc. London Math. Soc. 3, 75 – 95 (1972)

[CEP71] Cassels J., Ellison W., Pfister A.: *On sums of squares and on elliptic curves over function fields.* J. of Number Theory 3, 125 – 149 (1971)

[CG00] Chernousov V., Guletskiĭ V.: 2*-torsion of the Brauer group of an elliptic curve: generators and relations.* Preprint 00 – 37, Universität Bielefeld (2000)

[Co88] Colliot-Thélène, J.-L. (with the collaboration of Sansuc, J.-J.): *The rationality problem for fields of invariants under linear algebraic groups (with special regard to the Brauer group),* Unpublished Lecture Notes from the 9th ELAM: Santiago de Chile 1988

[CSS98] Colliot-Thélène, J.-L., Skorobogatov A., Swinnerton-Dyer P.: *Hasse principle for pencils of curves of genus one whose Jacobians have rational* 2*-division points.* Invent. math. 134, 579 – 650 (1998)

[Fadd51] Faddeev, D. K.: *Simple algebras over a function field in one variable,* Proc. Steklov Inst. 38, 321 – 344 (1951); English transl. in AMS Transl. 3, 15 – 38 (1956)

[G99] Guletskiĭ, V.: *Algebras of Exponent 2 over an Elliptic Curve.* Preprint 99 – 112, Universität Bielefeld (1999)

[GMY97] Guletskiĭ, V., Margolin, G., Yanchevskiĭ V.: *Presentation of two-torsion part in Brauer groups of curves by quaternion algebras* (in Russian). Dokl. NANB. 41, no. 6, 4–8 (1997)

[Hu87] Husemöller, D.: *Elliptic Curves*, Berlin Heidelberg New York: Springer 1987

[LT58] Lang, S., Tate, J.: *Principal homogeneous spaces over abelian varieties.* Amer. J. of Math. 80, 659 – 684 (1958)

[L56] Lang, S.: *Algebraic groups over finite fields.* Amer. J. of Math. 78, 555 – 563 (1956)

[Lich69] Lichtenbaum, S.: *Duality Theorems for Curves over P-adic Fields.* Invent. Math. 7, 120 – 136 (1969)

[Mi81] Milne, J.S.: *Comparison of the Brauer group with the Tate-Šafarevič group.* J. Fac. Science, Univ. Tokyo, Sec. IA 28, 735–743 (1981)

[Mi86] Milne, J.S.: *Arithmetics Duality Theorems.* Progress in Math. Vol. 1.: Academic Press 1986

[P82] Pierce, R.S.: *Associative algebras.* Graduate Texts in Mathematics 88: Springer-Verlag 1982

[Pu98] Pumplün, S.: *Quaternion algebras over elliptic curves.* Comm. Algebra 26, no. 12, 4357–4373 (1998)

[Sch69] Scharlau, W.: *Über bie Brauer-Gruppe eines algebraischen Funktionen-körpers in einer Variabein.* J. für die reine und angew. Math. 239/240, 1–6 (1969)

[Serre64] Serre, J.-P.: *Cohomologie Galoisienne,* Berlin Heidelberg New York: Springer-Verlag 1964

[Serre79] Serre, J.-P.: *Local Fields,* New York Heidelberg Berlin: Springer-Verlag 1979

[Sil85] Silverman, J.: *The Arithmetic of Elliptic Curves,* Berlin Heidelberg New York: Springer-Verlag 1985

[S99] Skorobogatov A.: *Beyond the Manin obstruction.* Invent. math. 135, 399–424 (1999)

[Tate57] Tate, J.: *WC-groups over $\mathfrak{p}$-adic fields.* Séminaire Bourbaki, Décembre 1957, no. 1556

[YM96] Yanchevskiĭ, V., Margolin G.: *The Brauer groups and the torsion of local elliptic curves.* St. Petersburg Math. J. 7, no. 3, 473–505 (1996)

V. Chernousov
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
33615 Bielefeld
Germany
chernous@mathematik.uni-bielefeld.de

V. Guletskiĭ
Institute of Mathematics
Surganova str. 11
220072 Minsk, Belarus
guletskii@im.bas-net.by

Metric Spaces in Pure and Applied Mathematics

A. Dress, K. T. Huber[1], V. Moulton[2]

Received: September 25, 2001
Revised: November 5, 2001

Communicated by Ulf Rehmann

Abstract. The close relationship between the theory of quadratic forms and distance analysis has been known for centuries, and the theory of metric spaces that formalizes distance analysis and was developed over the last century, has obvious strong relations to quadratic-form theory. In contrast, the first paper that studied metric spaces *as such* – without trying to study their embeddability into any one of the standard metric spaces nor looking at them as mere 'presentations' of the underlying topological space – was, to our knowledge, written in the late sixties by John Isbell. In particular, Isbell showed that in the category whose objects are metric spaces and whose morphisms are *non-expansive* maps, a unique *injective hull* exists for every object, he provided an explicit construction of this hull, and he noted that, at least for finite spaces, it comes endowed with an intrinsic polytopal cell structure.

In this paper, we discuss Isbell's construction, we summarize the history of — and some basic questions studied in — *phylogenetic analysis*, and we explain why and how these two topics are related to each other. Finally, we just mention in passing some intriguing analogies between, on the one hand, a certain stratification of the cone of all metrics defined on a finite set X that is based on the combinatorial properties of the polytopal cell structure of Isbell's injective hulls and, on the other, various stratifications of the cone of positive semi-definite quadratic forms defined on $\mathbb{R}^n$ that were introduced by the Russian school in the context of reduction theory.

2000 Mathematics Subject Classification: 15A63, 05C05, 92-02, 92B99

[1]The author thanks the Swedish National Research Council (VR) for its support (grant# B 5107-20005097/2000).
[2]The author thanks the Swedish Natural Science Research Council (VR) for its support (grant# M12342-300)

Keywords and Phrases: injective hull, tight span, phylogenetic tree, quadratic form

1 Introduction

The close relationship between distance analysis and quadratic-form theory was known already in pre-Pythagorean times: A ceramic slab found in the near east, for instance, presents the triples of integers 3,4,5; 5,12,13; 7,24,25; ... and it is very likely that these integers were of interest to Babylonean builders as they allowed to build walls at right angles without any particular tool except a long string with 12 = 3+4+5 or 30 = 5+12+13 or ... equidistant nodes (see Figure 1).

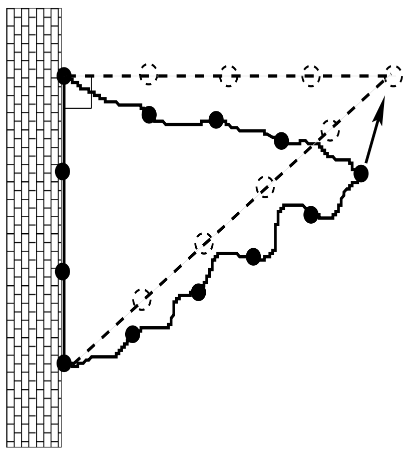


Figure 1: *The figure shows how a wall can be built at a right angle with a string of length 12.*

The Pythagorean Theorem puts this knowledge into more formal terms. And since then, the analysis of distance relationships has always been closely intertwined with that of quadratic forms. The development of differential geometry since Gauss as well as the development of geometric algebra in the 19th century — culminating in the definition of Clifford and Cayley algebras and Hamilton's definition of quarternian fields — clearly testifies to this fact.

In the early 20th century, attempts to develop appropriate conceptual frameworks for dealing with topological phenomena led Frechet to the definition of metric spaces. While this caused most mathematicians to think of metric spaces as just a rather convenient tool to define and to deal with topological spaces, a few began to study metric spaces for their own sake. Menger and Blumenthal in particular began developing *distance geometry* providing and investigating necessary and sufficient conditions for a given metric space to be isometrically embeddable into *standard* metric spaces, e.g. the n-dimensional Euclidean or some hyperbolic, elliptic, or L_p space — rediscovering, by the way, an important result of Cayley's regarding the significance of the now famous *Cayley-Menger determinants* in this context[3]. While, at their time, this effort did not stimulate much of a response among their fellow mathematicians, it turned out to be crucial later on for developing algorithms that would identify the spatial structure of proteins from *two-dimensional* NMR data (cf. [7]).

2 John Isbell's Contribution

Perhaps the first paper that studied metric spaces *as such* – without trying to study their embeddability into standard metric spaces nor looking at them as mere 'presentations' of the underlying topological space – was, to our knowledge, written in the late sixties by John Isbell (cf. [23]). Trying to capture the decisive aspects of distance relationships, he proposed to define the *category of metric spaces* as follows: Its objects — for sure — are the metric spaces. But, noting that

- using continuous maps as morphisms would create too flexible a category, overemphasizing the topological aspects and neglecting the true metric structure (e.g. any bijection between two finite metric spaces would then be an isomorphism)

while

- using isometries only would result in too rigid a category without enough morphisms,

he proposed to use the *non-expansive* maps from a metric space A into a metric space B as the set of morphisms from A to B, that is, those maps $f : A \to B$ for which the distance in B between the image $f(a)$ and $f(a')$ of two points a and a' from A never exceeds their distance in A (or, in other words, the continuous maps from A to B for which the Weierstrass δ can always be chosen to be equal to the Weierstrass ε).

Isbell then went on to show that a unique *injective hull* exists in this category for every one of its objects, providing an explicit construction of this hull for all spaces and noting that it comes endowed, at least for finite spaces, with an

[3] Actually, Cayley's original paper dealing with these determinants was the first to introduce the present notation for determinants.

intrinsic polytopal structure.

More precisely, Isbell presented the following intriguing observations:

(i) There exist *injective* metric spaces, that is, metric spaces $X = (X, d) = (X, d : X \times X \to \mathbb{R})$ such that, for every isometric embedding $\alpha : X \hookrightarrow X'$ of (X, d) into another metric space (X', d'), there exists a non-expansive *retract* $\alpha' : X' \to X$, that is, a non-expansive map α' from X' into X with $\alpha' \circ \alpha = Id_X$.

(ii) Every metric space (X, d) can be embedded isometrically into an injective metric space $(\hat{X}, \hat{d})$.

(iii) Given any such isometric embedding $\alpha : X \hookrightarrow \hat{X}$ of a metric space (X, d) into an injective metric space $(\hat{X}, \hat{d})$, there exists a unique smallest injective subspace $(\bar{X}, \bar{d})$ of $(\hat{X}, \hat{d})$ containing $\alpha(X)$. This subspace depends – up to isometry – only on (X, d):

- The map

$$\bar{X} \to \mathbb{R}^X : \bar{x} \mapsto (h_{\bar{x}} : X \to \mathbb{R} : x \mapsto \bar{d}(\alpha(x), \bar{x}))$$

is easily seen to define an isometric embedding of $\bar{X}$ into the set $\mathbb{R}^X$ of all maps from X into $\mathbb{R}$ endowed with the supremum norm (or l_∞*metric*)

$$||f, g||_\infty := \sup(|f(x) - g(x)| : x \in X) \ \ (f, g \in \mathbb{R}^X).$$

- And its image consists exactly of all those maps $f \in \mathbb{R}^X$ that satisfy the condition

$$f(x) = \sup(d(x, y) - f(y) : y \in X)$$

for all $x \in X\}$.

In [9], this subset of $\mathbb{R}^X$ has also been called the *tight span* $T(X, d)$ of (X, d) — a tradition that we will follow is in this paper, too.

(iv) In addition, the above embedding identifies X with the set

$$\{h_x : X \to \mathbb{R} : y \mapsto d(y, x) : x \in X\}$$

and, hence, with the subset

$$T^0(X, d) := \{f \in T(X, d) : 0 \in f(X)\}$$

of $T(X, d)$.

(v) The above definition/construction of $T(X, d)$ identifies it with a subset of the convex set

$$P(X, d) := \{f \in \mathbb{R}^X : f(x) + f(y) \geq d(x, y) \text{ for all } x, y \in X\},$$

more precisely, it identifies it with the set of all *minimal* maps in $P(X, d)$ (relative to the partial order $P(X, d)$ inherits from the partial order of $\mathbb{R}^X$ defined, as usual, by $f \leq g \iff f(x) \leq g(x)$ for all $x \in X$). Thus, it consists of a locally finite collection of (low-dimensional) faces of $P(X, d)$ whenever this convex set is actually a convex polytope (i.e. determined by a 'locally finite' collection of half spaces) which is surely the case if X itself is finite.

(v) $T(X, d)$ is always contractible. More precisely, there exists always a continuous family f_t $(t \in [0, 1])$ of non-expansive maps

$$f_t : T(X, d) \to T(X, d)$$

with $f_0 = Id_{T(X,d)}$ and $\#f_1(T(X, d)) = 1$.

Although these notions may appear to be somewhat strange at first, the tight span of small metric spaces (X, d) can be described in simple geometric terms as follows: In case X consists of just two points of distance c, its tight span is exactly the interval of length c, its end points being just the two points from X (thus the name "tight span"). In case X consists of just three points of distance c_1, c_2, c_3, its tight span is the union of three intervals of length $(c_1+c_2-c_3)/2$, $(c_1+c_3-c_2)/2$, and $(c_2+c_3-c_1)/2$, respectively, all identified at one end point while the other three end points are the three points from X. In Figure 2, we picture the tight span of a generic 4-point metric space:
In general, the tight span of a *finite* metric space (X, d) coincides exactly with the union of all compact faces of the polytope $P(X, d)$. Using this fact, it is possible to determine the polytopal structure of the tight span for a generic metric space of cardinality up to 5, cf. [9]. For finite metric spaces of larger cardinality, it is also possible in principle to determine their tight span, though it can be a tricky combinatorial problem to do this explicitly for any particular given metric space (see e.g.[11, 20]).
It is worthwhile to note that Isbell's construction does not really need a *metric* d to perform its task. It also works just as well for *every* map D from the set $\mathcal{P}_{\text{fin}}(X)$ of all finite subsets of a set X into $\underline{\mathbb{R}} := \mathbb{R} \cup \{-\infty\}$ (rather than only the map $D = D_d : \mathcal{P}_{\text{fin}}(X) \to \underline{\mathbb{R}}$ defined by $D(Y) := d(x, y)$ in case $Y = \{x, y\}$ for some x, y in X, and $D(Y) := -\infty$ else): Indeed, if such a map D is given, we may define

$$P(X, D) := \{f \in \mathbb{R}^X : \sum_{x \in Y} f(x) \geq D(Y) \text{ for all } Y \in \mathcal{P}_{\text{fin}}(X)\}$$

and

$$T(X, D) :=$$

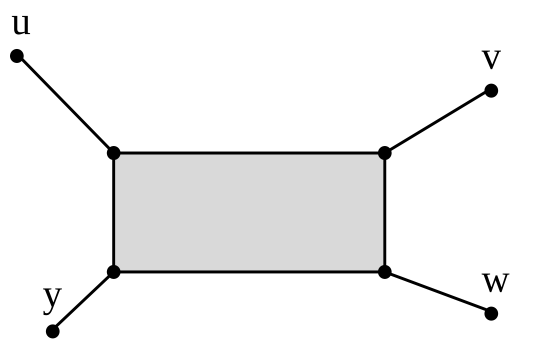


Figure 2: *The tight span of a generic metric d on the set $\{u, v, w, y\}$ for which $d(u, w) + d(v, y)$ is the largest of the three sums $d(u, w) + d(v, y)$, $d(u, v) + d(w, y)$, and $d(u, y) + d(v, w)$; it consists of eight 0-cells, eight 1-cells, and one 2-cell.*

$$\{f \in \mathbb{R}^X : f(x) = \sup(D(Y \cup \{x\}) - \sum_{y \in Y} f(y)) \text{ for all } Y \in \mathcal{P}_{\text{fin}}(X - \{x\})\}$$

just as before (so that $P(X, D_d) = P(X, d)$ and $T(X, D_d) = T(X, d)$ holds for every metric d and the map D_d associated with it according to the definition above). It is then not too difficult to establish, in this much more general setting, most of the results collected above in the special case considered originally by John Isbell.

Perhaps a bit surprisingly, this generalization can be used to construct affine buildings of GL-type. Assume that K is field with a valuation

$$\text{val} : K \to \underline{\mathbb{R}}$$

that satisfies the usual conditions

(i) $\text{val}(x) = -\infty \iff x = 0$,

(ii) $\text{val}(xy) = \text{val}(x) + \text{val}(y)$,

(iii) $\text{val}(x + y) \leq \max(\text{val}(x), \text{val}(y))$

for all $x, y \in K$ and consider, for some natural number n, the set $X := K^n$ and the map $D : \mathcal{P}_{fin}(X) \to \underline{\mathbb{R}}$ defined by

$$D(Y) := \text{val}(\det(x_1, ..., x_n))$$

if $Y = \{x_1, ..., x_n\}$ and $n = \#Y$, and

$$D(Y) = -\infty$$

else ($Y \in \mathcal{P}_{\text{fin}}(X)$). Then, it is easily seen that $T(X, D)$ coincides — together with its induced polytopal structure – with the affine building associated with $GL(n, K)$ provided the valuation in question is discrete while, in general, it provides at least a useful generalization that should also coincide with generalizations proposed so far for non-discrete valuations [15].

We expect that, in addition, the following example is of relevance in the context of symplectic and orthogonal groups: Let X be any vector space over K on which a sesqui-linear form $\langle \cdot \,|\, \cdot \rangle$ from $X \times X$ into K is defined and assume that $\langle \cdot \,|\, \cdot \rangle$ is also "almost symmetric" (i.e. that $\langle\, x \,|\, y \,\rangle = 0 \iff \langle\, y \,|\, x \,\rangle = 0$ holds for all x, y in X). It is then easy to see that the map D defined by

$$D(Y) := \text{val}(\det(\langle x_i | x_j \rangle)_{i,j=1,\dots,n})$$

if $Y = \{x_1, \dots, x_n\}$ and $n = \#Y$ holds, defines indeed a well-defined map from $\mathcal{P}_{\text{fin}}(X)$ into $\underline{\mathbb{R}}$ to which Isbell's construction can be applied. We have not yet checked, but expect $T(X, D)$ to coincide with the corresponding affine building of the symplectic group $Sp(2n, K)$ if X is of dimension $2n$ and the form $\langle \cdot \,|\, \cdot \rangle$ is non degenerate and skew-symmetric. We are not so sure about what happens in case $\langle \cdot \,|\, \cdot \rangle$ is non degenerate and symmetric. But we know, of course, that Isbell's construction at least provides in any case a nice contractible space on which the symmetry group of $(X, \langle \cdot \,|\, \cdot \rangle)$ acts in a canonical fashion (cf. [9]).

3 Phylogenetic Analysis

Isbell's construction was rediscovered in 1982 (see [9]) when the process of (re)constructing phylogenetic trees from distance data was scrutinized to develop methods for checking the suitability of data for and to improve the reliability of phylogenetic analysis (and, curiously enough, it was rediscovered again in 1994 in a completely different context, cf. [6]).

The goal of phylogenetic analysis is to derive a complete, consistent and, hopefully, true picture of the evolutionary branching process that produced a class of present — and, sometimes also some extinct — species from their last common ancestor, e.g. the evolution of all the various forms of tetrapodes from the first *amphibia*-like beings crawling out of the sea around 400 million years ago.

The first such phylogenetic tree encompassing all plant and animal kingdoms then known was constructed in 1866 (see Figure 3) just seven years after the publication, in 1859, of Charles Darwin's (1809-1882) *The Origin of Species*[4] by the German biologist Ernst Haeckel (1834-1919), the most ardent supporter of Darwin in that time in Germany. While Darwin never made much effort

[4] or, more correctly, *On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life*

to construct phylogenetic trees explicitly (even though he was, of course, fully aware that his theory implies the existence of such a tree and remarked "*As we have no record of the lines of descent, the pedigree can be discovered only by observing the degrees of resemblance between the beings which are to be classed*"), it was not too difficult for Ernst Haeckel to design his tree. All he had to do was to give a *Darwinian* dynamic interpretation of the static systems previously put forward (in form of tableaux) by Carolus Linnaeus (1707-1778), Georges Cuvier (1769-1832) and others.

Linnaeus had become famous very early in his life for his analysis of gender in plants, thus recognizing an amazing universality of certain basic laws of life in the then known living world. In his *Systema Naturae, Sive Regna Tria Naturae Systematice Proposita*[5], published in 1735 in Leiden, Linnaeus followed the most rigorous scientific traditions of his time. These had been established by John Ray (1628-1705) in his writings since 1660, culminating in his *Methodus Plantorum Nova* from 1682 and his posthumously published *Synopsis Avium et Piscium* from 1713. Ray was probably the first scientist to recognize and to conceptualize the *invariance* of species as the fundamental basis of life science. Linnaeus followed Ray's insights and constructed a whole binary *hierarchy* of *phyla, kingdoms, genera, families, subfamilies* etc. to classify biological species according to their intrinsic similarities.

These ideas were then taken up by scientists like August Quirinus Rivinus (1652-1723) in Germany and Joseph Pitton de Tournefort (1656-1708) in France as well as, a little later, by Linnaeus in Sweden. Like Ray, Linnaeus insisted that the living world (except for a few species doomed by the great deluge and documented in the fossil record) had been created in that very order in which it presents itself to us today and that the task of taxonomy was to search for a "natural system" that would reflect the Divine Order of creation. Darwin's ideas allowed to reinterpret Linnaeus' classes as *clades*, i.e. as collections of *all* those species derived from *one* common ancestor. Thus, the static Linnaean system could immediately be transformed into Haeckel's dynamic tree.

However, there are always many details in such trees that are hotly debated, and the evidence that can be used for tree (re)construction is often scarce, inconsistent and contradictory. For instance, it is not yet fully known whether the *monotremata* — the Australian *duck-billed platypus* and the *spiny anteaters* (*echidna aculeata* and *echidna Bruynii*) — are more closely related to the *marsupalia* (opossums, kangaroos, etc.) than to us (the *placental mammals* or *eutheria*) or whether, the third alternative, the placental mammals and the marsupalia are more closely related to each other than both are to the platypus and the echidnas (even though the most recent molecular data appears to sup-

[5]The System of Nature, or the Three Kingdoms of Nature Presented Systematically

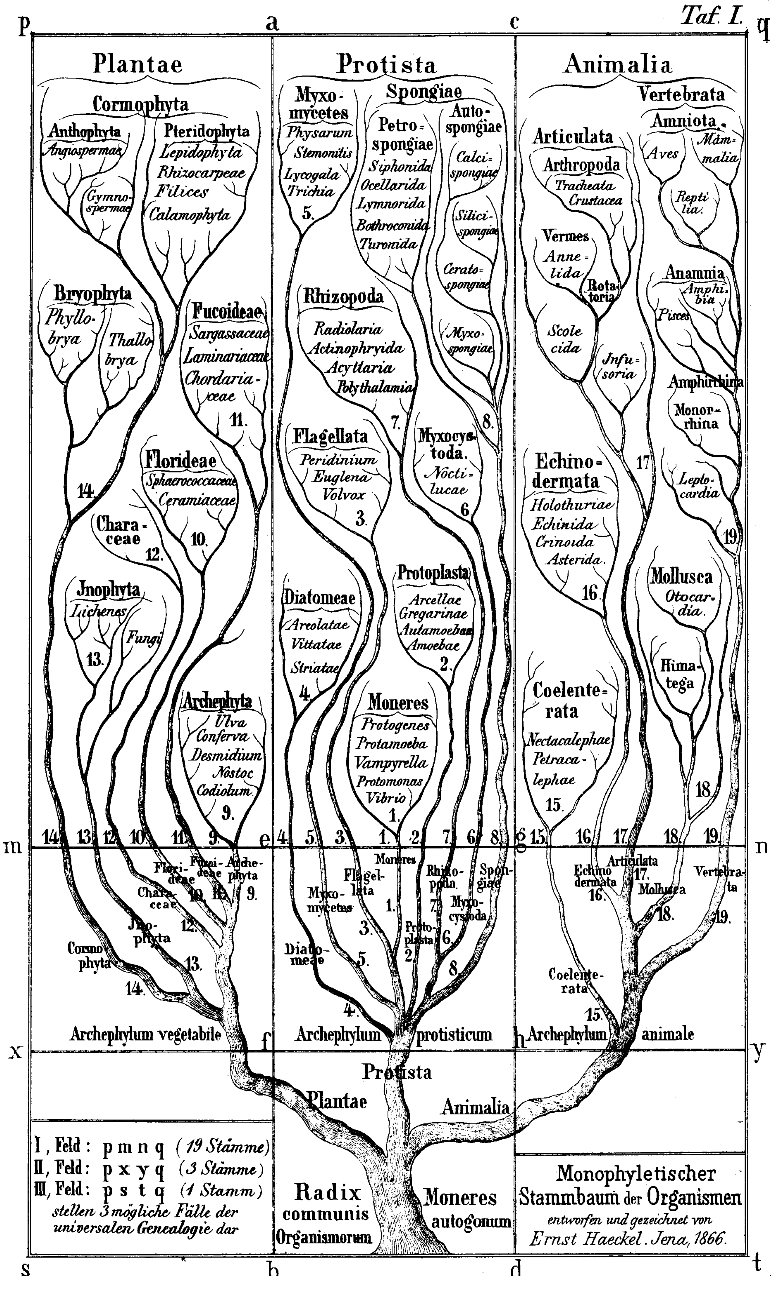


Figure 3: Haeckel's tree of life (1866).

port the first alternative). And even less clear are at present the phylogenetic relationships among the various groups of placental mammals (cf. [28] and also http://phylogeny.arizona.edu/tree for fascinating up to date information regarding the present view of Haeckel's *Tree of Life*[6]).

Consequently, biologists have always been looking for further evidence – in addition to morphological evidence, from all parts of the organism in all stages of its development, and metabolic peculiarities – on which phylogenetic conclusions could be based. So, when the amino acid sequence of closely related proteins from distinct species (and encoded by related though not identical genes all supposedly derived from *one* common ancestral gene by accumulating successive mutations) became known in sufficient abundance in the late 1960's, some biologists realized quickly that such documents of *molecular* evolution might provide the most convincing evidence on which to build phylogenetic trees.

The first paper exploiting this idea that appeared in Science was written by Walter Fitch and Emanuel Margoliash almost thirty five years ago. It was entitled simply *Construction of Phylogenetic Trees* (cf. [19]) and it caused a revolution in taxonomy. It used the amino acid sequences of cytochrom C, a protein of decisive importance in oxygen metabolism in all eucariots, derived from more than 20 species from all eucariot kingdoms. Fitch and Margoliash estimated the *genetic distance* $d(S_1, S_2)$ between any two of these sequences S_1 and S_2 in terms of the easily computed number of mismatches between S_1 and S_2 relative to a *multiple alignment* of all of the sequences in question that had been constructed simply by hand — in this specific case a comparatively simple task in view of the large overall similarity of the sequences.
They then constructed their tree automatically by employing the following very simple standard algorithm from cluster-analysis textbooks:
Given a finite set X together with a symmetric map d from $X \times X$ into $\mathbb{R}$, one defines the set $V(X, d)$ of nodes of the tree $T_{F\&M}(X, d)$ to be constructed to consist of those subsets Y of X that constitute, for some real number c, a connected component of the graph $\Gamma_c := (X, E_c)$ whose vertex set is the given set X and whose edges consist of all pairs of elements x, y from X with $d(x, y) \leq c$. And two such nodes Y_1, Y_2 are connected by an edge if and only if $Y_1 \subset Y_2$ holds and there is no Y in $V(X, d)$ with $Y_1 \subset Y \subset Y_2$ — or, equivalently, if $\#\{Y \in V(X, d) : Y_1 \subseteq Y \subseteq Y_2\} = 2$ holds.
At that time, most taxonomists were appalled by this approach. The definitive result of a scholar's whole life of research could apparently now be produced in less than a minute by an insightless machine. Others, impressed by the obvious potential of this new approach (which had almost simultaneously also been conceived independently by at least one further research group) took

[6]Or just visit the American Museum of Natural History in New York where the fourth floor has been devoted to actually spreading out all along the floor our present version (or vision?) of that tree!

immediately to the road to visit the authors of that paper.

Today, essentially every paper dealing with phylogenetics offers trees produced automatically from sequence data by appropriate computer programs. It also became obvious in the mean time that such trees are not the end of scientific investigation in taxonomy. Rather to the contrary, it needs the full knowledge and expertise of experienced scientists to discuss the computer-generated trees and to point out their weak as well as their strong points.

Clearly, the obvious idea any tree-reconstruction algorithm must use is that, given any three sequences that have been derived by the process of replication, mutation, and selection from one common ancestral sequence, the last common ancestral sequence of the two more similar among those three sequences should have existed later than the last common ancestral sequence of all three sequences. This suggests the following tree-construction algorithm: First, identify each sequence S from the set X of sequences in question with the corresponding one-element clade $\{S\}$ consisting of S, only. Then, using any appropriately defined dissimilarity measure $d : X \times X \to \mathbb{R}$ (e.g. the mismatch or *Hamming* distance employed by Fitch and Margoliash), search for those two sequences S_1, S_2 that have minimal dissimilarity and, supposing that no other sequence in X can be an offspring of the last common ancestral sequence of S_1 and S_2, fuse S_1 and S_2 into one larger *d-clade* $\{S_1\} \cup \{S_2\}$. Then replace the set X by a smaller set X' representing all maximal, presently identified (d-)clades (that is, the one d-clade of cardinality 2 and the additional, not yet processed single-element clades at that stage) and define a new dissimilarity measure on those clades by defining the distance $d(Y_1, Y_2)$ of any two such clades Y_1, Y_2 to be some function of the dissimilarities $d(y_1, y_2)$ with $y_1 \in Y_1$ and $y_2 \in Y_2$. And then, repeat the above process to identify the next two clades that are to be fused into one new, larger d-clade, and so on. Obviously, if $d(Y_1, Y_2)$ is defined by $d(Y_1, Y_2) := \min\{d(y_1, y_2) | y_1 \in Y_1, y_2 \in Y_2\}$ for any two d-clades Y_1, Y_2, this will lead exactly to the tree $T_{F\&M}(X, d)$ described above.

However, this procedure is obviously bound to make mistakes: Assume, we have four sequences S_1, S_2, S_3, S_4 and that, during the evolution of those four sequences from their common ancestor sequence S, there were first two distinct offsprings sequences S', S'' of S so that S_1 and S_2 were later derived from S' and S_3 and S_4 from S''. Assume furthermore that S_1 remained very similar to S' and S_3 remained very similar to S'' and S_2 as well as S_4 diverged very far from their respective ancestor sequences. Then, the above algorithm will inevitably form a wrong clade $\{S_1, S_3\}$ (see Figure 4).
Many algorithms have therefore been designed to deal with this particular problem. And quite a few of them accept the dissimilarities computed from the input sequences as a starting point, yet they search for a tree that provides the best global approximation of the given dissimilarity pattern, i.e. a tree

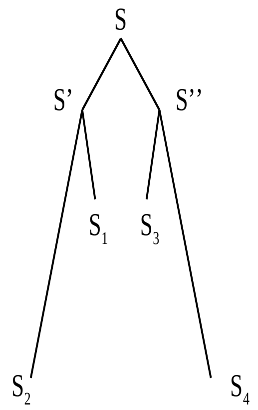


Figure 4: *As explained in the text, the incorrect clade* $\{S_1, S_3\}$ *is formed by the agglomeration algorithm and the 'true topology' of the tree is not found.*

whose leaves are labeled by the elements from X, and to whose branches appropriate edge lengths are attached so that the resulting induced *tree metric* (that associates to any pair of elements x, y from X the total length of the unique path from the two leaves labeled with x and y) matches the given dissimilarities *in toto* as closely as possible.

To imagine the task one has to perform using the approach it is worthwhile to observe that the *space* of all possible dissimilarities that can be defined on an n-set X has dimension $\binom{n}{2}$ while the subspace of *tree-like* dissimilarities that can be defined on X has dimension $2n-3$ (the maximal number of branches in a tree with n leaves) and forms a rather complex low-dimensional network of large codimension $\binom{n}{2} - 2n + 3$ within this cone. Consequently, while trying to identify the best global 'tree-like' approximation of the given dissimilarity pattern, there may be many rather distinct, yet essentially equally good tree-like approximations to a given arbitrary dissimilarity d and to find the best one will naturally be very hard (e.g. the tree-like dissimilarities form a space of dimension 17 and, hence, of codimension 28 in the 45-dimensional space of all dissimilarities that can be defined on a set of 10 points — so its much worse than looking for a needle in the hay stack, a codimension 2 (or, at most, 3) problem — or than trying to find the closest river mouth to a given point on earth).

4 Tree Reconstruction and the Tight Span

Nevertheless, this approach suggests a number of interesting, purely mathematical questions which to pursue might still be helpful in this context: E.g., it leads to the question which dissimilarities are *tree like* dissimilarities, i.e. which dissimilarities would fit exactly into a tree, and whether that tree would be completely determined by those dissimilarities. Fortunately, these two questions have simple answers that have been discovered in the sixties and seventies of the last century independently by various mathematicians (cf. [5, 29, 30]):

(i) A dissimilarity d is tree like if and only if

$$d(x,y) + d(u,v) \leq \max\{d(x,u) + d(y,v), d(x,v) + d(y,u)\}$$

holds for all x, y, u, v from X.

(ii) If this condition is fulfilled, there is only one tree that fits the given dissimilarity (up to isomorphism, and except for additional branches not involved with the given data).

Remarkably, once we define a metric on *all* points of that tree (whether a branching point, an end point, or just a point somewhere on some branch) by associating again to any two such points x, y the total length of the unique path from x to y, the resulting metric space, necessarily an $\mathbb{R}-$tree (by the very definition of $\mathbb{R}-$trees) actually coincides with the injective hull of the metric defined on its leaves. This establishes not only the uniqueness of the tree in question; it can also be used to study the structure of that tree in terms of the metric defined on its leaves. More importantly, it suggests to use the injective hull in any case, whether or not the input dissimilarities satisfy the above four-point condition, as a good substitute for the tree in question — at least, it is always simply connected (though not always of dimension one).

In particular, if there exists some subset K of small diameter within this injective hull T not containing any leaf, yet such that its complement $T - K$ has several connected components, the (labels of the) leaves in at least all but one of these components have a good chance to form one of those clades within X that phylogenetic analysis is designed to find.

It was exactly this observation which lead to the rediscovery of Isbell's construction in 1982 mentioned above. And it also motivated and initiated many further investigations regarding the structure of injective metric spaces and their relevance in phylogenetic analysis (cf. [10, 11, 13, 14]). In particular, the analysis of injective hulls of finite metric spaces made it obvious that the injective hull of a sum $d = d_1 + d_2 + \ldots + d_k$ of k metrics $d_1, d_2, \ldots, d_k$ defined on a finite set X is closely related to that of the summands $d_1, d_2, \ldots, d_k$ provided these metrics form a *coherent decomposition* of the metric d, i.e. provided there exist, for every map $f : X \to \mathbb{R}$ with

$f(x) + f(y) \geq d_1(x, y) + d_2(x, y) + \ldots + d_k(x, y)$ for all $x, y \in X$, some maps $f_1, f_2, \ldots, f_k : X \to \mathbb{R}$ such that $f_i(x) + f_i(y) \geq d_i(x, y)$ holds for all $x, y \in X$ and for all $i = 1, 2, \ldots, k$ (cf. [2, 24, 25, 26]).

Moreover, defining a metric d to be

- a *split* — or a *cut* — metric if there are exactly two subsets of X in the set X/d of equivalence classes of elements of X relative to the equivalence relation $\simeq$ defined on X by $x \simeq y \Leftrightarrow d(x, y) = 0$, and
- a *split-prime* metric if it cannot be decomposed into a coherent sum of a split metric and another metric,

it could be shown that

- every metric d defined on a finite set X has a unique coherent decomposition — also called the *canonical split decomposition* of d — into a sum $d = d_1 + d_2 + \ldots + d_k + d_0$ of pairwise linearly independent split metrics $d_1, d_2, \ldots, d_k$ and a split-prime metric d_0 (possibly the 0-metric),
- the metrics $d_1, d_2, \ldots, d_k$ occurring in this decomposition are always linearly independent (as elements in the vector space of all maps from $X \times X$ into $\mathbb{R}$) — and so are $d_1, d_2, \ldots, d_k, d_0$ if $d_0 \neq 0$ holds,
- the metrics $d_1, d_2, \ldots, d_k$ occurring in this decomposition are – up to scaling – exactly those split metrics d' defined on X for which $d - d'$ is also a metric and the two metrics $d', d - d'$ form a coherent decomposition of d,
- if d is a tree-like metric, then the split-prime metric d_0 in the corresponding canonical coherent decomposition $d = d_1 + d_2 + \ldots + d_k + d_0$ of d into a sum of pairwise linearly independent split metrics $d_1, d_2, \ldots, d_k$ and a split-prime metric d_0 vanishes while the split metrics $d_1, d_2, \ldots, d_k$ correspond in a one-to-one fashion to the branches of the associated tree (cf. Figure 5).

This was of considerable interest within the context of phylogenetic analysis: If a split metric d' occurs as a summand in a coherent component of a metric d derived from a family of phylogenetically related sequences, there is a good chance that at least one of the two equivalence classes in X/d' is one of those clades within X that we want to find.

In particular, given any metric d defined on a set X of cardinality n, the linear independence of the split metrics occurring in the canonical decomposition of d implies that there exist, up to scaling, at most $\binom{n}{2}$ split metrics d' such that (i) $d - d'$ is also a metric and (ii) the two metrics $d', d - d'$ are coherent, – clearly too many to fit into a tree (because a tree with n leaves has at most $2n - 3$ edges), but surely much less than $2^{n-1} - 1$, the number of all split

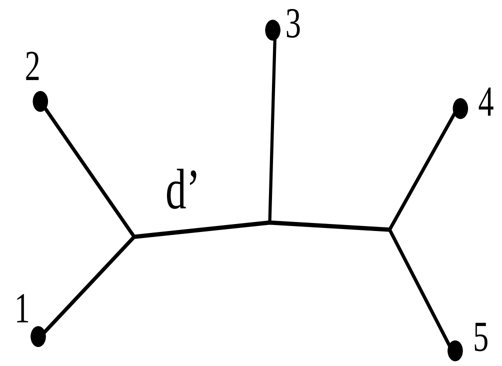


Figure 5: *A tree with leaves labeled by the finite set* $\{1, 2, 3, 4, 5\}$. *The branch separating the vertices* $1, 2$ *from the vertices* $3, 4, 5$ *corresponds to a split metric* d' *with* $X/d' = \{\{1, 2\}, \{3, 4, 5\}\}$.

metrics that, up to scaling, can be defined on an n-set.

In addition, it might even be helpful when analyzing a given data set to realize that several competing evolutionary interpretations of the data are possible (as indicated by the existence of two split metrics d', d'' in the canonical decomposition of d for which $\#(X/(d' + d'')) = 4$ holds) or that, at least, some additional feature (e.g. some sort of *convergence*) might be present in the data.

Consequently, algorithms were developed to compute, given any metric D, all split metrics d for which the above conditions are fulfilled as well as to visualize the resulting *split network* (cf. [3, 12, 22]). The resulting SplitsTree program has proven useful in diverse phylogenetic applications. Moreover, as Figure 6 shows, it can as well be applied to all sorts of distance data: The split networks in Figure 6(left) was computed for the distances between the towns of Wellington on the North Island, and Christchurch, Greymouth etc. on the South Island of New Zealand that were taken from a mileage chart. If one compares this graph with a map of New Zealand a good correlation between the distribution of vertices and the geographical locations of the towns is observed. It has also been applied to analyze the perceived similarity of colors and — in *stemmatology* — the "kinship" relations between the various handwritten versions of Chaucer's *Canterbury tales* written by Geoffrey Chaucer about 100 years before book printing was invented (in central Europe) (cf. [4]).

These examples illustrate that split networks can give meaningful representations of data even if they are not necessarily tree-like in character. Within

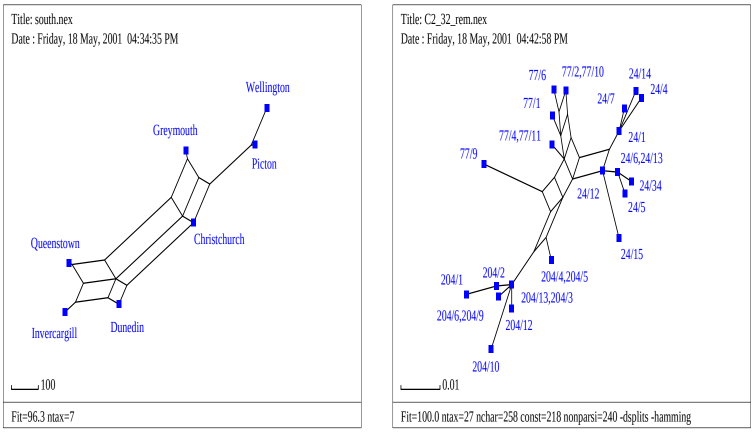


Figure 6: *Split networks for a mileage chart of New Zealand (left) and a hepatitis C virus data set (right).*

biology, non tree-like distances often arise when analyzing viral data sets, a phenomenon that is probably caused by more complex evolutionary processes such as recombination. In Figure 6 (right), we present a split network that was computed for a hepatitis C data set which was presented in [1]. In this graph, a complex relationship between various viral sequences (represented by the labeled vertices) is observed. However, there is a clear separation between the three sets of vertices labeled with prefixes 204, 77, and 24, and indeed this reflects the fact that the viruses corresponding to vertices prefixed by 204 and 77 were taken from recipients of blood transfusions from a donor who was infected with the viruses corresponding to the vertices prefixed by 24.

For more applications of the SplitsTree program to biological data see e.g. [8, 12, 16, 21, 27]. The latest version of SplitsTree, written by Daniel Huson, can be obtained from:

http://www.mathematik.uni-bielefeld.de/∼huson

There is also a www version of the program running at:

http://bibiserv.techfak.uni-bielefeld.de/splits

Some further references and discussions of related topics can be found on the following www pages:

http://www.fmi.mh.se/∼vince/publications/publications.html
http://www.mathematik.uni-bielefeld.de/∼terhalle

and further phylogenies by Haeckel can be found on the following web pages:

http://www.boga.ruhr-uni-bochum.de/spezbot/Folien/Abb1_Stammbaum_Haeckel.html
http://genome.imb-jena.de/stammbaum.html

5 Back to Mathematics and Quadratic forms

In addition to these applications, there are also striking analogies between split-decomposition theory and the theory of positive semi-definite quadratic forms as developed by the Russian school: In both fields, one considers a large convex cone (either consisting of all metrics defined on a finite set or consisting of all positive semi-definite quadratic forms defined on some finite-dimensional vector space), one has good reasons to decompose this cone — in one way or the other — into a family of finitely generated convex subcones, and one wants to understand the combinatorics of the resulting stratification of the large cone. In split-decomposition theory, it is the concept of *coherence* that gives rise to the stratification in question: given any two metrics d and d', defined on a fixed finite set X, one may define the metric d' to be a *coherent specialization* of the metric d if there exists some positive real number ρ such that $d'' := \rho d - d'$ is also a metric and the two metrics d', d'' form a coherent decomposition of d. One can show that, given any metric d defined on X, the collection of metrics d' that are coherent specializations of d forms a finitely generated convex subcone $C(d)$ of the cone of all metrics defined on X. Moreover, some (not at all obvious) conditions on d are known from split-decomposition theory which imply that $C(d)$ is a simplicial cone while this does not seem to hold in general for every metric d.

Very similar problems have been (and still are being) studied in the theory of positive semi-definite quadratic forms while trying to understand the process of reduction of quadratic forms (cf. [17, 18]). And in both areas, the extremals of the convex cones in question — the positive semi-definite quadratic forms of rank one on the one hand and the split metrics as well as some further, not yet well understood metrics on the other — appear to be of special significance.

Thus, it might prove rather useful trying not only to develop both theories in parallel, but also to understand the deeper reason for the striking analogy between them.

Acknowledgment The authors would like to thank Olaf Breidbach for his very helpful comments on the historical part of this note.

References

[1] J.-P. Allain, Y. Dong, A.-M. Vandamme, V. Moulton, M. Salmei, Evolutionary rate and genetic drift of hepatitis C virus are not correlated with the host immune response: studies of infected donor-recipient clusters, Journal of Virology 74 (2000) 2541–2549.

[2] H.-J. Bandelt, A. Dress, A canonical decomposition theory for metrics on a finite set, Adv. in Math. 92 (1992) 47–105.

[3] H.-J. Bandelt, A. Dress, Split decomposition: a new and useful approach to phylogenetic analysis of distance data, Molecular Phylogenetics and Evolution 1(3) (1992) 242–252.

[4] A. C. Barbrook, C. J. Howe, N. Blake, P. Robinson, The phylogeny of The Canterbury Tales, Nature 394 (1998) 839.

[5] P. Buneman, The recovery of trees from measures of dissimilarity. In F. Hodson et al., Mathematics in the Archaeological and Historical Sciences, (pp.387–395), Edinburgh University Press, 1971.

[6] M. Chrobak, L. Lamore, Generosity helps or an 11-competitive algorithm for three servers, Journal of Algorithms 16 (1994) 234–263.

[7] G. M. Crippen, T. F. Havel, Distance Geometry and Molecular Confirmation, Wiley, Chinchester, 1981.

[8] J. Dopazo, A. Dress, A. von Haeseler, Split decomposition: A technique to analyze viral evolution, PNAS 90 (1993) 10320–10324.

[9] A. Dress, Trees, tight extensions of metric spaces, and the cohomological dimension of certain groups: A note on combinatorial properties of metric spaces, Adv. in Math. 53 (1984) 321–402.

[10] A. Dress, K. T. Huber, V. Moulton, A comparison between the median and the tight-span completion of finite split systems, Annals of Combinatorics, 2, 1998, 299–311.

[11] A. Dress, K. T. Huber, V. Moulton, An explicit computation of the injective hull of certain finite metric spaces in terms of their associated Buneman complex, Advances in Mathematics, to appear.

[12] A. Dress, D. Huson, V. Moulton, Analyzing and visualizing distance data using SplitsTree, Discrete Applied Mathematics 71 (1996) 95–110.

[13] A. Dress, V. Moulton, M. Steel, Trees, taxonomy and strongly compatible multi-state characters, Advances in Applied Mathematics 19 (1997) 1–30.

[14] A. Dress, V. Moulton, W. Terhalle, T-theory: An overview, European Journal of Combinatorics 17 (1996) 161–175.

[15] A. Dress, W. Terhalle, The tree of life and other affine buildings. Proceedings of the International Congress of Mathematicians, Vol. III (Berlin, 1998). Doc. Math. 1998, Extra Vol. III, 565–574.

[16] A. Dress, R. Wetzel, The human organism - A place to thrive for the immuno-deficiency virus, in Proceedings of IFCS, Paris.

[17] P. Engel, V. Grishukhin, An example of a non-simplicial L-type domain, European J. Combin. 22 (2001) 491–496.

[18] R. M. Erdahl, Zonotopes, dicings, and Voronoi's conjecture on parallelohedra, European J. Combin. 20 (1999) 527–549.

[19] W. M. Fitch, E. Margoliash, Construction of phylogenetic trees, Science 155 (1967) 279–284.

[20] O. Goodmann, V. Moulton, On the tight span of an antipodal graph, Discrete Mathematics 218 (2000) 73–96.

[21] E. Holmes, M. Worobey, A. Rambaut, Phylogenetic evidence for recombination in dengue virus, Mol. Biol. Evol. 16(3) (1999) 405–409.

[22] D. Huson, SplitsTree: a program for analyzing and visualizing evolutionary data, Bioinformatics 14(1) (1998) 68–73.

[23] J. Isbell, Six theorems about metric spaces, Comment. Math. Helv. 39 (1964) 65–74.

[24] J. Koolen, V. Moulton, A note on the uniqueness of coherent decompositions, Advances in Applied Mathematics 19 (1997) 444–449.

[25] J. Koolen, V. Moulton, U. Toenges, The coherency index, Discrete Mathematics 192 (1998) 205–222.

[26] J. Koolen, V. Moulton, U. Toenges, A classification of the six-point prime metrics, European Journal of Combinatorics 21 (2000) 815–829.

[27] P. Plikat, K. Nieselt-Struwe, A. Meyerhans, Genetic drift can dominate short-term HIV-1 nef quasispecies evolution in vitro, Journal of Virology 71 (1997) 4233–4240.

[28] D. Penny, M. Hasegawa, The platypus put in its place, Nature 387 (1997) 549–550.

[29] J.M.S. Simões-Pereira, A note on tree realizability of a distance matrix, J. Comb. Theory (B) 6 (1969) 303–310.

[30] K.A.Zaretsky, Reconstruction of a tree from the distances between its pendent vertices, Uspekhi Math. Nauk, Russian Mathematical Surveys 20 (1965) 90–92 (In Russian).

A. Dress
FSPM-Strukturbildungsprozesse
University of Bielefeld
D-33501 Bielefeld, Germany
dress@mathematik.uni-bielefeld.de

K. T. Huber
FMI, Mid Sweden University
S 851-70 Sundsvall
Sweden
kathi@dirac.fmi.mh.se

V. Moulton
FMI, Mid Sweden University
S 851-70 Sundsvall
Sweden
vince@dirac.fmi.mh.se

Isotropy and Factorization in Reduced Witt Rings

Robert W. Fitzgerald

Received: April 5, 2001

Communicated by Ulf Rehmann

Abstract. We consider reduced Witt rings of finite chain length. We show there is a bound, in terms of the chain length and maximal signature, on the dimension of anisotropic, totally indefinite forms. From this we get the ascending chain condition on principal ideals and hence factorization of forms into products of irreducible forms.

2000 Mathematics Subject Classification: 11E81, 12D15
Keywords and Phrases: isotropy, quadratic forms, Witt ring

R will denote a (real) reduced Witt ring. A form $q \in R$ is totally indefinite if $|\mathrm{sgn}_\alpha q| < \dim q$ for all orderings α of R. It is well-known that such a form need not be isotropic. However, when R has finite chain length, $\mathrm{cl}(R)$, we show there are restrictions on the possible dimensions of anisotropic, totally indefinite forms. To be specific,

$$\dim q \leq \tfrac{1}{2}\mathrm{cl}(R) \max_\alpha\{|\,\mathrm{sgn}_\alpha q|^2\},$$

unless $R = \mathbb{Z}$ and q is one-dimensional. The proof depends on Marshall's classification of reduced Witt rings of finite chain length.

This bound allows us to show that R, of finite chain length, satisfies the ascending chain condition on principal ideals. One consequence of this result is that chains of basic clopen sets $H(a_1, \dots, a_n)$, for fixed n, stabilize. Another consequence is that non-zero, non-units of R factor into a finite product of irreducible elements (in the sense of Anderson and Valdes-Leon). This had been previously known only for odd dimensional forms in rings with only finitely many orderings.

Conversely, we show, for a wide class of reduced Witt rings R, that the ascending chain condition on principal ideals implies R has finite chain length. The proof relies on Marshall's notion of a sheaf product. We close with examples of factorization into irreducible elements. These illustrate how the factorization of even dimensional forms is less well behaved than the factorization of odd dimensional forms studied in [8].

We set some of the notation. R will be an abstract Witt ring, in the sense of Marshall [11], and reduced. The main case of interest is the Witt ring of a Pythagorean field. X_R, or just X if the ring is understood, denotes the set of orderings (equivalently, signatures) on R. We always assume X is non-empty. For a form $q \in R$ and ordering $\alpha \in X$, the signature of q at α will be denoted by either $\mathrm{sgn}_\alpha q$ or $\hat{q}(\alpha)$.

We let G_R, or just G when R is understood, denote the group of one-dimensional forms of R. When R is the Witt ring of a field, $G = F^*/F^{*2}$. Forms in R are written as $\langle a_1, \dots, a_n\rangle$, with each $a_i \in G$. An n-fold Pfister form is a product $\langle 1, a_1\rangle\langle 1, a_2\rangle \cdots \langle 1, a_n\rangle$, denoted by $\langle\langle a_1, a_2, \dots, a_n\rangle\rangle$. The set of orderings X has a topology with basic clopen sets

$$H(a_1, \dots, a_n) = \{\alpha \in X : a_i >_\alpha 0 \quad \text{for all } i\},$$

where each $a_i \in G$. The *chain length* of R, denoted by $\mathrm{cl}(R)$, is the supremum of the set of integers k for which there is a chain

$$H(a_0) \subsetneq H(a_1) \subsetneq \cdots \subsetneq H(a_k)$$

of length k (each $a_i \in G$).

A subgroup $F \subset G$ is a *fan* if it satisfies : any subgroup $P \supset F$ such that $-1 \notin P$ and P has index 2 in G is an ordering. The index of the fan is $[G : F]$. The set of orderings P that contain F is denoted X/F. Note that $|X/F| = 2^{n-1}$ if F has index 2^n. The *stability index* of R, denoted by $\mathrm{st}(R)$, is the supremum of $\log_2 |X/F|$ over all fans in G.

If R_1 and R_2 are reduced Witt rings then so is the product

$$R_1 \sqcap R_2 = \{(r_1, r_2) : r_1 \in R_1, r_2 \in R_2 \quad \text{and} \quad \dim r_1 \equiv \dim r_2 \pmod 2\}.$$

E will always denote a group of exponent 2. If R is a reduced Witt ring then so is the group ring generated by E, denoted by $R[E]$. E_k will denote the group of exponent 2 and order 2^k. We will always take $t_1, \dots, t_k$ as generators of E_k (except when $k = 1$ when we use just t). For an arbitrary E we use $t_1, t_2, \dots$ as generators. When E is uncountable we are assuming the use of infinite ordinals as indices. Lastly, if $S \subset G$ we write $sp(S)$ for the subgroup generated by S.

1. Isotropy.

Over $\mathbb{R}$ a form q is hyperbolic iff $\mathrm{sgn}\, q = 0$ and isotropic iff $|\mathrm{sgn}\, q| < \dim q$. The first statement holds for any reduced Witt ring but not the second. Our goal is to find a limit on the difference between $|\mathrm{sgn}\, q|$ and $\dim q$ for anisotropic forms. We restrict ourselves to reduced Witt rings with a finite chain length. Recall [12, 4.4.2] ([5] in the field case) that such rings are built up from copies of $\mathbb{Z}$ by finite products and arbitrary group ring extensions. The decomposition is unique except that $\mathbb{Z} \sqcap \mathbb{Z} = \mathbb{Z}[E_1]$.

We introduce some notation. Recall that E_k is generated by $t_1, \dots, t_k$. We fix a listing $x_1, \dots, x_{2^k}$ of the elements of E_k as follows. The list for E_1 is

$1, t_1$. The list for E_{k+1} is the list of E_k followed by t_{k+1} times the list for E_k. We also fix a listing $\alpha_1, \dots, \alpha_{2^n}$ of the orderings on $\mathbb{Z}[E_k]$. For the $k = 1$ we take α_1 to be the ordering with t_1 positive and α_2 to be the ordering with t_1 negative. The list for $\mathbb{Z}[E_{k+1}]$ consists of the orderings on $\mathbb{Z}[E_k]$ extended by taking t_{k+1} positive, followed by the extensions with t_{k+1} negative. Lastly, we define P_k to be the $2^k \times 2^k$-matrix whose (i, j) entry is the sign of x_j at the α_i ordering. Thus $P_1 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Lemma 1.1. *For each $k \geq 1$*

(1) *P_k is symmetric.*
(2) *$P_k^2 = 2^k I$.*
(3) *For $q = \sum n_i x_i \in \mathbb{Z}[E_k]$ let $s_i = \hat{q}(\alpha_i)$. Set $\bar{n} = (n_1, \dots n_{2^k})^T$, where T denotes the transpose, and $\bar{s} = (s_1, \dots, s_{2^k})^T$. Then $P_k \bar{n} = \bar{s}$.*

Proof. We use induction on k to prove (1) and (2). Both are clear for $k = 1$. By our construction,

$$P_{k+1} = \begin{pmatrix} P_k & P_k \\ P_k & -P_k \end{pmatrix}.$$

Thus P_k symmetric implies P_{k+1} is also. And

$$P_{k+1}^2 = \begin{pmatrix} 2P_k^2 & 0 \\ 0 & 2P_k^2 \end{pmatrix} = 2^{k+1} I.$$

Statement (3) is simple to check. $\square$

The reader may notice that each P_k is a Hadamard matrix, indeed the simplest examples of Hadamard matrices, namely Kronecker products of copies of $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$.

Notation. Let $M(q) = \max\{|\hat{q}(\alpha)| : \alpha \in X\}$.

Proposition 1.2. *Let $R = \mathbb{Z}[E]$, where E is an arbitrary group of exponent two. Suppose $q \in R$ is anisotropic. Then $\dim q \leq M(q)^2$.*

Proof. We may assume $q \in \mathbb{Z}[E_k]$ for some k. Write $q = \sum n_i x_i$ where $n_i \in \mathbb{Z}$ and the x_i form the list of the elements of E_k described above. Let $\bar{n}$ and $\bar{s}$ be as in (1.1). Then:

$$\begin{aligned} P_k \bar{n} &= \bar{s} \\ 2^k \bar{n} = P_k^2 \bar{n} &= P_k \bar{s} \\ \sum n_i^2 = \bar{n}^T \bar{n} &= \frac{1}{2^{2k}} \bar{s}^T P_k^T P_k \bar{s} \\ &= \frac{1}{2^k} \bar{s}^T \bar{s} = \frac{1}{2^k} \sum s_i^2. \end{aligned}$$

Now for each i we have $s_i^2 \leq M(q)^2$. So $\sum n_i^2 \leq M(q)^2$. Further, $|n_i| \leq n_i^2$ so $\dim q = \sum |n_i| \leq M(q)^2$. $\square$

Remarks. (1) The bound in (1.2) is sharp infinitely often. Let $\epsilon = (\epsilon_1, \dots, \epsilon_k)$ be a choice of signs, that is, each $\epsilon_i = \pm 1$. Pick a one-to-one correspondence between the 2^k many sign choices and the elements of $sp\{t_{k+1}, \dots, t_{2k}\}$, say $\epsilon \mapsto x_\epsilon$. Then consider

$$q = \sum_\epsilon x_\epsilon \langle\langle \epsilon_1 t_1, \dots, \epsilon_k t_k \rangle\rangle \in \mathbb{Z}[E_{2k}],$$

where the sum is over all possible sign choices. At each ordering of $\mathbb{Z}[E_k]$ exactly one of the Pfister forms has signature 2^k, the others having signature zero. In any extension of this ordering to $\mathbb{Z}[E_{2k}]$ we get $\operatorname{sgn} q = \pm 2^k$. Thus q is anisotropic, $\dim q = 2^{2k}$ and $M(q) = 2^k$. Hence $\dim q = M(q)^2$.
(2) The bound of (1.2) is not sharp for M's that are not 2-powers. For instance, suppose q is anisotropic and $M(q) = 3$. We may assume (see (2.6)) that q has signature 3 or -1 at each ordering. Let $q_0 = (q-1)_{an}$, the anisotropic part. Then $M(q_0) = 2$ and so $\dim q_0 \le 4$. Thus $\dim q \le 5 < M(q)^2$.
The bound of (1.2) can also be improved if k is fixed. For instance, one can show for anisotropic $q \in \mathbb{Z}[E_3]$ that $\dim q \le \frac{5}{2} M(q)$.

Theorem 1.3. *Suppose R is a reduced Witt ring of finite chain length. Let $q \in R$ be anisotropic. Then $\dim q \le \frac{1}{2} cl(R) M(q)^2$, unless $R = \mathbb{Z}$ and q is one-dimensional.*

Proof. The result is clear if $\dim q = 1$ so assume $\dim q \ge 2$. We may thus ignore the exceptional case. We will prove the result for $R = S[E]$, any E, by induction on the chain length of S. Say $\mathrm{cl}(S) = 1$ so that $S = \mathbb{Z}$. If $E = 1$ then $\dim q = M(q) \le \frac{1}{2} M(q)^2$ as $\dim q \ge 2$. If $E \ne 1$ then we are done by (1.2) as $\mathrm{cl}(\mathbb{Z}[E]) = 2$.
In the general case we may assume $S = S_1 \sqcap S_2$, with at least one of S_1 or S_2 not $\mathbb{Z}$. Then both S_1 and S_2 have smaller chain length than S and so we are assuming the result holds for $S_i[E]$, $i = 1, 2$ and any E.
First suppose $E = 1$. Write $q = (a, b)$ with $a \in S_1$ and $b \in S_2$. We may assume that $\dim a \ge \dim b$. Then $\dim q = \dim a$. We have by induction

$$\begin{aligned}\dim q = \dim a &\le \tfrac{1}{2}\mathrm{cl}(S_1) M(a)^2 \\ &\le \tfrac{1}{2}\mathrm{cl}(R) M(a)^2, \quad \text{since } \mathrm{cl}(R) = \mathrm{cl}(S_1) + \mathrm{cl}(S_2) \\ &\le \tfrac{1}{2}\mathrm{cl}(R) M(q)^2,\end{aligned}$$

as $\hat{q}(\alpha) = \hat{a}(\alpha)$ or $\hat{b}(\alpha)$ for every $\alpha \in X$ so that $M(a) \le M(q)$.
Next suppose $E \ne 1$. Since q has only finitely many entries we may assume that $q \in (S_1 \sqcap S_2)[E_k]$, for some k. Write $q = \sum (a_i, b_i) x_i$, where each $a_i \in S_1$ and $b_i \in S_2$ and the x_i's are our listing of the elements of E_k. Set

$$\varphi = (\sum a_i x_i, 0) + r(0, \sum b_i x_i) \in (S_1[E_k] \sqcap S_2[E_k])[E_1].$$

Now

$$\dim q = \sum_i \max\{\dim a_i, \dim b_i\}$$
$$\dim \varphi = \sum \dim a_i + \sum \dim b_i \geq \dim q.$$

We check the signatures. If $\alpha \in X_{S_1}$ and α^ϵ is an extension of α to $R = (S_1 \sqcap S_2)[E_k]$ then $\hat{q}(\alpha^\epsilon) = \sum \hat{a}_i(\alpha)\epsilon_i$ (here $\epsilon_i = \pm 1$ depending on the sign of x_i in the extension). Similarly, if $\beta \in X_{S_2}$ and β^ϵ is an extension to R then $\hat{q}(\beta^\epsilon) = \sum \hat{b}_i(\beta)\epsilon_i$.
We may also view α^ϵ as an extension of α to $S_1[E_k]$ and hence to $S_1[E_k] \sqcap S_2[E_k]$. Let $\alpha^{\epsilon+}$ denote the further extension to $(S_1[E_k] \sqcap S_2[E_k])[E_1]$ with r positive. We also have the other extensions $\alpha^{\epsilon-}, \beta^{\epsilon+}$ and $\beta^{\epsilon-}$. Then:

$$\hat{\varphi}(\alpha^{\epsilon+}) = \sum \hat{a}_i(\alpha)\epsilon_i$$
$$\hat{\varphi}(\alpha^{\epsilon-}) = \sum \hat{a}_i(\alpha)\epsilon_i$$
$$\hat{\varphi}(\beta^{\epsilon+}) = \sum \hat{b}_i(\beta)\epsilon_i$$
$$\hat{\varphi}(\beta^{\epsilon-}) = -\sum \hat{b}_i(\beta)\epsilon_i.$$

Thus $M(\varphi) = M(q)$.
Set $\varphi_1 = \sum a_i x_i \in S_1[E_k]$ and $\varphi_2 = \sum b_i x_i \in S_2[E_k]$. Then by induction we have:

$$\dim \varphi_1 \leq \tfrac{1}{2}\mathrm{cl}(S_1) M(\varphi_1)^2$$
$$\dim \varphi_2 \leq \tfrac{1}{2}\mathrm{cl}(S_2) M(\varphi_2)^2.$$

The previous computation shows that for any ordering γ of $(S_1[E_k] \sqcap S_2[E_k])[E_1]$ that $\hat{\varphi}(\gamma)$ equals $\hat{\varphi}_1(\alpha)$ or $\pm\hat{\varphi}_2(\beta)$ where γ restricts to either α on $S_1[E_k]$ or β on $S_2[E_k]$. Thus $M(\varphi_i) \leq M(\varphi)$ for $i = 1, 2$. We obtain

$$\begin{aligned}\dim \varphi = \dim \varphi_1 + \dim \varphi_2 &\leq \tfrac{1}{2}(\mathrm{cl}(S_1) + \mathrm{cl}(S_2)) M(\varphi)^2 \\ &= \tfrac{1}{2}\mathrm{cl}(R) M(\varphi)^2,\end{aligned}$$

using [12, 4.2.1]. Lastly, we have already checked that $\dim q \leq \dim \varphi$ and $M(q) = M(\varphi)$, giving the desired bound. $\square$

Remarks. (1) The bound of (1.3) is sometimes achieved. For example, in

$$R = (\mathbb{Z}[E_2] \sqcap \mathbb{Z}[E_2] \sqcap \mathbb{Z}[E_2])[E_2],$$

where the last E_2 is generated by s_1, s_2, let $\varphi = \langle 1, t_1, t_2, -t_1 t_2 \rangle$ and set $q = (\varphi, 0, 0) + s_1(0, \varphi, 0) + s_2(0, 0, \varphi)$. Then q is anisotropic, $\dim q = 12$, $M(q) = 2$ and $\mathrm{cl}(R) = 6$. Thus $\dim q = \frac{1}{2}\mathrm{cl}(R) M(q)^2$.

(2) Bröcker [3] has a result that looks similar to (1.3) but is apparently unrelated. There, in the version of [12, 7.7.3], if q is anisotropic, $\hat{q}(\alpha) = \pm 2^k$ for all α and $Y = \{\alpha : \hat{q}(\alpha) = 2^k\}$ is the union of basic open sets each of stability index at most $k+1$, then $\dim q \leq 2^{2k} = M(q)^2$.
(3) Bonnard [2] also has a result that looks like (1.3), which in fact uses Bröcker's result in the proof. In our notation, her result is: if R has finite stability index s and $q \in R$ is anisotropic then $\dim q \leq 2^{s-1}M(q)$. Her bound is slightly better than this. Chain length and stability index are independent invariants so again there is no apparent connection between (1.3) and Bonnard's result.
Recall that a form q is *weakly isotropic* if mq is isotropic for some $m \in \mathbb{N}$.

Corollary 1.4. *Let R be a real Witt ring (not necessarily reduced) of finite chain length. Let $q \in R$ be a form of dimension at least 2. If $\dim q > \frac{1}{2} cl(R)M(q)^2$ then q is weakly isotropic.*

Proof. Let $q_r = q + R_t \in R_{red}$, the reduced Witt ring. Then q_r is isotropic by (1.3). Hence $q_r \simeq \langle 1, -1\rangle + \varphi_r$, for some form $\varphi_r = \varphi + R_t \in R_{red}$. Then $2^k q \simeq 2^k\langle 1, -1\rangle + 2^k\varphi$, for some k, and so q is weakly isotropic. □

2. Chains of principal ideals.

We use the standard abbreviation ACC for *ascending chain condition.*

Proposition 2.1. *If ACC holds for the principal ideals of R then R has finite chain length.*

Proof. Suppose we have a tower

$$H(a_1) \supseteq H(a_2) \supseteq \cdots \supseteq H(a_n) \supseteq \cdots .$$

Set $q_n = \langle 1, 1, a_n\rangle$. Then $\hat{q}_n(\alpha)$ is 1 or 3, with $\hat{q}_n(\alpha) = 3$ iff $\alpha \in H(a_n)$. In particular, for every n we have $\hat{q}_{n+1}(\alpha)$ divides $\hat{q}_n(\alpha)$, for every $\alpha \in X$. Then q_{n+1} divides q_n by [7, 1.7]. Thus we have a tower of principal ideals :

$$(q_1) \subseteq (q_2) \subseteq \cdots \subseteq (q_n) \subseteq \cdots .$$

The ACC implies there exists a N such that $(q_N) = (q_m)$ for all $m > N$. Then $\hat{q}_N(\alpha)$ divides $\hat{q}_m(\alpha)$ for all $\alpha \in X$ and so $H(a_N) = H(a_m)$, for all $m > N$. □

We need some technical terms for the next result.

Definitions. A *fan tower* is a strictly decreasing tower of fans $F_1 > F_2 > \cdots > F_n > \cdots$, each of finite index plus a fixed choice of complements C_n where $G = C_n \times F_n$. We set $F_\infty = \cap F_n$. A *separating set of fan towers* is a finite set of fan towers $s_1, \ldots, s_\ell$, with $s_i = \{F_{in}\}$ such that

(1) Given any $q \in R$ there exists m, possibly depending on q, such that all entries of q are in $C_{im}F_{i\infty}$, for each i between 1 and ℓ.

(2) Given $K \subset \mathbb{Z}$ and forms $q_1, q_2 \in R$, there exists N, depending on q_1 and q_2 but not K, such that if for some $n > N$

$$\hat{q}_1^{-1}(K) \cap (X/F_{in}) = \hat{q}_2^{-1}(K) \cap (X/F_{in})$$

for all i then $\hat{q}_1^{-1}(K) = \hat{q}_2^{-1}(K)$.

Example. For a simple example, let $R = \mathbb{Z}[E]$ with E countably infinite. Let $F_i = sp\{t_{i+1}, t_{i+2}, \dots\}$ and $C_i = sp\{-1, t_1, \dots, t_i\}$. Then each F_i is a fan of finite index, each C_i is a complement and the F_i are strictly decreasing. Hence $\{F_i\}$ is a fan tower. Note that here $F_\infty = 1$. This fan tower is a separating (singleton) set of fan towers. A given form q has entries involving only a finite number of t_i's and so its entries lie in some C_m; this is the first condition. If we are given two forms q_1 and q_2 then again all of their entries lie in some C_N. So the signatures of the q_i depend only on the signs of $t_1, \dots t_N$ in that ordering. Hence if $\hat{q}_1$ and $\hat{q}_2$ agree on X/F_N then they agree at every ordering. This is the second condition.
Roughly, our fan towers will look like this example. When there is a product we will need one tower in each coordinate, hence a separating set.

Lemma 2.2. *If R has finite chain length then R has a separating set of fan towers.*

Proof. We prove this by induction on the chain length. When $\mathrm{cl}(R) = 1$ then $R = \mathbb{Z}$ and the result is clear. We first consider the case $R = S_1 \sqcap S_2$. Write G_1 and X_1 for G_{S_1} and X_{S_1} and similarly for G_2 and X_2. Let $\{s_1^1, \dots, s_{\ell_1}^1\}$ be a separating set of fan towers for S_1. Here $s_k^1 = \{F_{ki}^1\}$ with complements C_{ki}^1. Set $F_{ki} = F_{ki}^1 \times G_2$, which is a fan in $G = G_1 \times G_2$ with complement $C_{ki} = C_{ki}^1 \times 1$. Then for $1 \le k \le \ell_1$, $r_k = \{F_{ki}\}$ is a fan tower. Note that $F_{k\infty} = F_{k\infty}^1 \times G_2$.
Similarly, let $\{s_1^2, \dots, s_{\ell_2}^2\}$ be a separating set of fan towers for S_2, with $s_k^2 = \{F_{ki}^2\}$ and complements C_{ki}^2. Set $F_{\ell_1+k\,i} = G_1 \times F_{ki}^2$ and $C_{\ell_1+k\,i} = 1 \times C_{ki}^2$. Then for $1 \le k \le \ell_2$, $r_{\ell_1+k} = \{F_{\ell_1+k\,i}\}$ is a fan tower. We check that $r_1, \dots, r_{\ell_1}, r_{\ell_1+1}, \dots, r_{\ell_1+\ell_2}$ is a separating set of fan towers for R.
We check the first condition. We are given a form $q = \langle (a_1, b_1), \dots, (a_n, b_n) \rangle \in R$. By induction, there exists a m_1 such that $a_1, \dots, a_n \in C_{km_1}^1 F_{k\infty}^1$ for all k. So

$$(a_1, b_1), \dots, (a_n, b_n) \in C_{km_1} F_{k\infty} = C_{km_1}^1 F_{k\infty}^1 G_2,$$

for all k with $1 \le k \le \ell_1$. Similarly, there exists a m_2 such that $b_1, \dots, b_n \in C_{km_2}^2 F_{k\infty}^2$, for all $1 \le k \le \ell_2$. Hence $(a_1, b_1), \dots, (a_n, b_n) \in C_{km_2} F_{k\infty}$ for all k with $\ell_1 < k \le \ell_1 + \ell_2$. So take m to be the maximum of m_1 and m_2.
We next check the second condition. We are given $K \subset \mathbb{Z}$ and forms $q_1 = (u_1, v_1)$ and $q_2 = (u_2, v_2)$. Note that $\hat{q}_1^{-1}(K) = \hat{u}_1^{-1}(K) \cup \hat{v}_1^{-1}(K) \subset X_1 \cup X_2$, a disjoint union. By induction there exists a N_1 satisfying the second condition for K, u_1 and u_2 and a N_2 satisfying the second condition for K, v_1 and v_2. Let N be the maximum of N_1 and N_2. Suppose for some $n > N$ we have

$$\hat{q}_1^{-1}(K) \cap (X/F_{kn}) = \hat{q}_2^{-1}(K) \cap (X/F_{kn}),$$

for all $1 \leq k \leq \ell_1 + \ell_2$. For $1 \leq k \leq \ell_1$ we have:

$$\begin{aligned}\hat{q}_1^{-1}(K) \cap (X/F_{kn}) &= \hat{q}_1^{-1}(K) \cap (X_1/F_{kn}^1) \\ &= \hat{u}_1^{-1}(K) \cap (X_1/F_{kn}^1).\end{aligned}$$

We thus obtain

$$\hat{u}_1^{-1}(K) \cap (X_1/F_{kn}^1) = \hat{u}_2^{-1}(K) \cap (X_1/F_{kn}^1),$$

for all $1 \leq k \leq \ell_1$. By the second condition on S_1 we have $\hat{u}_1^{-1}(K) = \hat{u}_2^{-1}(K)$. Similarly, $\hat{v}_1^{-1}(K) = \hat{v}_2^{-1}(K)$ and so $\hat{q}_1^{-1}(K) = \hat{q}_2^{-1}(K)$.

Now suppose $R = S[E]$. Set $T_i = sp\{t_{i+1}, t_{i+2}, \dots\}$. Let $\{s_1, \dots, s_\ell\}$ be a separating set of fan towers for S where $s_k = \{F'_{ki}\}$ and the complements are C'_{ki}. Then $F_{ki} = F'_{ki}T_i$ is a fan of finite index in R with complement $C_{ki} = C'_{ki}sp\{t_1, \dots, t_i\}$. Then $r_k = \{F_{ki}\}$ is a fan tower. Note that $F_{k\infty} = F'_{k\infty}$. We show that $\{r_1, \dots, r_\ell\}$ is a separating set of fan towers for R.

For the first condition we are given a form $q \in R = S[E]$. There exists a p such that $q \in S[E_p]$. Write $q = \sum a_i x_i$ where each $a_i \in S$ and the x_i's are some list of the elements of E_p. By induction, for each i there exists a $m(i)$ such that every entry of a_i is in $C'_{km(i)}F'_{k\infty}$ for all k, $1 \leq k \leq \ell$. Let m be the maximum of the $m(i)$ and p. Then every entry of every a_i lies in $C'_{km}F'_{k\infty} \subset C_{km}F_{k\infty}$ and each x_i lies in $sp\{t_1, \dots, t_p\} \subset C_{km}$. So every entry of q lies in $C_{km}F_{k\infty}$, for all k.

For the second condition we are given $K \subset \mathbb{Z}$ and two forms $q_1, q_2 \in R$. Again there exists a p such that $q_1, q_2 \in S[E_p]$. Write $q_1 = \sum a_i x_i$ and $q_2 = \sum b_i x_i$ with $a_i, b_i \in S$ and the x_i as before. Let $\epsilon \in \{\pm 1\}^p$ be a choice of sign for $t_1, \dots, t_p$. Let $\epsilon(x_i)$ be the resulting sign of x_i. Set:

$$q_1^\epsilon = \sum a_i \epsilon(x_i) \qquad q_2^\epsilon = \sum b_i \epsilon(x_i),$$

both forms in S. For each ϵ there exists a N_ϵ so that condition 2 holds for q_1^ϵ and q_2^ϵ. Let N be the maximum of the N_ϵ and p.

If $\alpha \in X_S$ we let α^ϵ be the extension of α to $S[E_p]$ with $t_i > 0$ iff $\epsilon(t_i) = 1$. Then we claim that:

$$\hat{q}_1^{-1}(K) \cap X_{S[E_p]} = \bigcup_\epsilon [(\hat{q}_1^\epsilon)^{-1}(K)]^\epsilon.$$

Namely if $\alpha^\epsilon \in X_{S[E_p]}$ and $\hat{q}_1(\alpha^\epsilon) \in K$ then

$$\hat{q}_1(\alpha^\epsilon) = \sum \hat{a}_i(\alpha)\epsilon(x_i) = \hat{q}_1^\epsilon(\alpha).$$

Hence $\alpha^\epsilon \in (\hat{q}_1^\epsilon)^{-1}(K)^\epsilon$. The reverse inclusion is similar.

Now let $\alpha^{\epsilon e}$ denote any extension of α^{ϵ} to $R = S[E]$. Then by the claim we have:

$$\hat{q}_1^{-1}(K) = \bigcup_e \left(\bigcup_{\epsilon} [(\hat{q}_1^{\epsilon})^{-1}(K)]^{\epsilon} \right)^e \tag{2.3}$$

So $\hat{q}_1^{-1}(K) \cap (X/F_{kn}) = \hat{q}_2^{-1}(K) \cap (X/F_{kn})$ implies that

$$(\hat{q}_1^{\epsilon})^{-1}(K) \cap (X_s/F'_{kn}) = (\hat{q}_2^{\epsilon})^{-1}(K) \cap (X_s/F'_{kn}),$$

for all sign choices ϵ. Hence by condition 2 applied to S we obtain $(\hat{q}_1^{\epsilon})^{-1}(K) = (\hat{q}_2^{\epsilon})^{-1}(K)$ for all ϵ. Then (2.3) gives $\hat{q}_1^{-1}(K) = \hat{q}_2^{-1}(K)$. $\square$

Lemma 2.4. *Suppose R has a separating set of fan towers $\{s_1, \dots, s_\ell\}$. Let $q \in R$ and $K \subset \mathbb{Z}$. Let m be the index such that every entry of q lies in $C_{km}F_{k\infty}$, for all $1 \le k \le \ell$. Let $n > m$. Then for each k we have:*

$$|\hat{q}^{-1}(K) \cap (X/F_{kn})| = \frac{|X/F_{kn}|}{|X/F_{km}|} |\hat{q}^{-1}(K) \cap (X/F_{km})|.$$

Proof. Pick a k with $1 \le k \le \ell$. $F_{kn} \subset F_{km}$ are both fans of finite index so we can write $F_{km} = H \times F_{kn}$ with H spanned by $h_1, \dots, h_p$, where $2^p = |X/F_{kn}|/|X/F_{km}|$. Every $\alpha \in X/F_{km}$ has 2^p extensions to X/F_{kn}, one for each choice of signs (± 1) for the h_i. Specifically, if ϵ is a sign choice for the h_i and $h \in H$, let $\epsilon(h)$ be the resulting sign of h. Since $G = C_{km}HF_{kn}$, the extension of $\alpha \in X/F_{km}$ to X/F_{kn} via ϵ is: $\alpha^{\epsilon}(chf) = \alpha(c)\epsilon(h)$, where $c \in C_{km}$, $h \in H$ and $f \in F_{kn}$. We thus have

$$X/F_{kn} = \bigcup_{\epsilon} (X/F_{km})^{\epsilon}.$$

Write $q = \langle a_1, a_2, \dots \rangle$. By assumption, each a_i is in $C_{km}F_{k\infty} \subset C_{km}F_{kn}$. Hence $\alpha^{\epsilon}(a_i) = \alpha(a_i)$. Thus :

$$\hat{q}^{-1}(K) \cap (X/F_{kn}) = \bigcup_{\epsilon} \left(\hat{q}^{-1}(K) \cap (X/F_{km}) \right)^{\epsilon}.$$

So $|\hat{q}^{-1}(K) \cap (X/F_{kn}) = 2^p |\hat{q}^{-1}(K) \cap (X/F_{km})|$, and the result follows. $\square$

Lemma 2.5. *Let $q \in R$ be a form of dimension n. Let F be a fan of finite index and let $K \subset \mathbb{Z}$. Then :*

$$|\hat{q}^{-1}(K) \cap (X/F)| = \frac{k}{2^n} |X/F|,$$

for some integer k, $0 \leq k \leq 2^n$.

Proof. Write $q = \langle a_1, \dots, a_n\rangle$. Then $\hat{q}^{-1}(K)$ is a disjoint union of $H(\epsilon_1 a_1, \dots, \epsilon_n a_n)$ for various choices of $\epsilon = (\epsilon_1, \dots, \epsilon_n) \in \{\pm 1\}^n$. Set $\rho_\epsilon = \langle\langle \epsilon_1 a_1, \dots, \epsilon_n a_n\rangle\rangle$. Then by the easy half of the representation theorem

$$\sum_{\alpha \in X/F} \hat{\rho}_\epsilon(\alpha) \equiv 0 \pmod{|X/F|}$$
$$2^n |H(\epsilon_1 a_1, \dots, \epsilon_n a_n) \cap (X/F)| = k_\epsilon |X/F|,$$

for some non-negative integer k_ϵ. Then :

$$|\hat{q}^{-1}(K) \cap (X/F)| = \sum_\epsilon \frac{k_\epsilon}{2^n}|X/F| = \frac{k}{2^n}|X/F|,$$

for some non-negative integer k. $\square$

The following is essentially from [9]. For a form $q = \langle a_1, \dots, a_n\rangle$ the *discriminant* is $\operatorname{dis} q = (-1)^{n(n-1)/2} a_1 \cdots a_n$. This is sometimes called the signed discriminant.

Lemma 2.6. *Let q be an odd dimensional form.*

(1) *$dis\, q >_\alpha 0$ iff $\hat{q}(\alpha) \equiv 1 \pmod 4$.*
(2) *$sgn_\alpha dis(q) q \equiv 1 \pmod 4$ for all $\alpha \in X$.*
(3) *If $0 \neq a = bc$ and $\hat{a}(\alpha) = \pm \hat{b}(\alpha)$ for all $\alpha \in X$ with $\hat{a}(\alpha) \neq 0$ then there exists $d \in G$ such that $\langle d\rangle a = b$.*

Proof. (1) Suppose $n = \dim q$. Let $s = \hat{q}(\alpha)$. If r is the number of α-negative entries in q then

$$\mathrm{sgn}_\alpha \mathrm{dis}\, q = (-1)^{n(n-1)/2}(-1)^r = (-1)^{\frac{n(n-1)}{2} + \frac{n-s}{2}} = (-1)^{(n^2-s)/2}.$$

This is positive iff $n^2 - s \equiv 0 \pmod 4$. As n is odd we get that the discriminant is positive iff $\hat{q}(\alpha) = s \equiv n^2 \equiv 1 \pmod 4$.
(2) is easy to check. For (3), let $A = \{\alpha \in X : \hat{a}(\alpha) \neq 0\}$. Then $\hat{c}(\alpha) = \pm 1$ for all $\alpha \in A$. In particular c is odd dimensional and $\hat{a}(\alpha) = 0$ iff $\hat{b}(\alpha) = 0$.Let $d = \operatorname{dis} c$. Then $\langle d\rangle c$ has signature 1 for all $\alpha \in A$ by (2). Hence $\langle d\rangle bc$ and b have the same signature at each $\alpha \in B$, and also at each $\alpha \notin A$ (as both have signature 0 there). Thus $\langle d\rangle a = \langle d\rangle bc = b$. $\square$

Theorem 2.7. *Let R be a reduced Witt ring. Then ACC holds for principal ideals iff the chain length of R is finite.*

Proof. (2.1) gives $(\longrightarrow)$. For the converse, let $(q) \subset (q_1) \subset (q_2) \subset \cdots$ be an ascending chain of principal ideals in R. Note that as each q_i divides q we have $M(q_i) \leq M(q)$. Let $M = M(q)$. Then (1.3) gives $\dim q_i \leq \frac{1}{2}\mathrm{cl}(R)M^2$ for all i (note q is not one-dimensional else all $(q_i) = R$).

We begin with some simple reductions. If all q_i are 0 then the result is clear. If some q_i is not zero then all the later q_i's are not zero. We may start our tower there, that is, we may assume $q \neq 0$. For a non-zero form φ define $\deg \varphi$ to be the largest d such that 2^d divides $\hat{\varphi}(\alpha)$ for all $\alpha \in X$. Since q_{i+1} divides q_i we have $\deg q_{i+1} \leq \deg q_i$. Let d_0 be the minimum of the degrees of the q_i. We may start our tower at a q_j of minimal degree, that is, we may assume that $\deg q = \deg q_i$ for all i. Now we may write $q = q_i \varphi_i$ for some form φ_i. We check that φ_i is odd dimensional. If instead φ_i is even dimensional then 2 divides $\hat{\varphi}_i(\alpha)$ for all α and so 2^{d+1} divides $\hat{q}(\alpha)$ for all α, contradicting our reduction to a tower of uniform degree. Hence φ_i is odd dimensional. In particular, $\hat{q}(\alpha) = 0$ iff $\hat{q}_i(\alpha) = 0$.

Let D be the set of integers $d > 1$ that divide some non-zero $\hat{q}(\alpha)$, $\alpha \in X$. Write $D = \{d_1, \dots, d_z\}$ with $d_1 < d_2 < \cdots < d_z$. Set $A(i, d_j) = \hat{q}_i^{-1}(\pm d_j)$. Let d_k be the largest element of D (if any) for which $\{A(i, d_k) : i \geq 1\}$ is not finite. Our goal is to show that there is in fact no such d_k. Our assumption on d_k means that for each $j > k$ we have a t_j such that $A(t, d_j) = A(t_j, d_j)$ for all $t \geq t_j$. Let T be the maximum of the t_j, $j > k$. Then by starting our tower of ideals with q_T, we may assume $A(i, d_j) = A(1, d_j)$ for all $j > k$ and all $i \geq 1$.

We first check that $A(i+1, d_k) \subset A(i, d_k)$ for any i. Namely, $q_i = q_{i+1}\varphi$ for some form φ. So if $\alpha \in A(i+1, d_k)$ then $\pm d_k$ divides $\hat{q}(\alpha)$. Also $|\hat{q}_i(\alpha)|$ is not of the d_j with $j > k$ else $\alpha \in A(i, d_j) = A(i+1, d_j)$, which is impossible as $\alpha \in A(i+1, d_k)$. Thus $|\hat{q}_i(\alpha)| \leq d_k$ and is divisible by d_k. Hence $\hat{q}_i(\alpha) = \pm d_k$ and $\alpha \in A(i, d_k)$ as desired.

Let $s = \{F_m\}$ be one fan tower in a separating set of fan towers for R (which exists by (2.2)). The first condition for a separating set, plus a simple induction argument, shows that for each i there exists a least $m(i)$ with every entry of $q_1, \dots, q_i$ in $C_{m(i)}F_\infty$. Note that $m(i+1) \geq m(i)$. Let $p(i)$ be the number of distinct values of

$$\frac{|A(j, d_k) \cap (X/F_{m(i)})|}{|X/F_{m(i)}|} \equiv \gamma(i, j),$$

over j with $1 \leq j \leq i$. Now, by (2.4)

$$\begin{aligned} \gamma(i+1, j) &= \frac{|A(j, d_k) \cap (X/F_{m(i+1)})|}{|X/F_{m(i+1)}|} \\ &= \frac{|X/F_{m(i+1)}|}{|X/F_{m(i)}|} \frac{|A(j, d_k) \cap (X/F_{m(i)})|}{|X/F_{m(i+1)}|} \\ &= \gamma(i, j). \end{aligned}$$

Hence $p(i+1) \geq p(i)$, with only $\gamma(i+1, i+1)$ possibly being a new value.

Since every $\dim q_i \leq \frac{1}{2}\mathrm{cl}(R)M^2$, (2.5) implies each $p(i) \leq 2^{\mathrm{cl}(R)M^2/2} + 1$. Hence there is a t_0 such that $p(t) = p(t_0)$ for all $t \geq t_0$. Let $p = p(t_0)$ and $m = m(t_0)$. Say $\gamma(t_0, j_1), \dots, \gamma(t_0, j_p)$ are the distinct γ-values over $1 \leq j \leq t_0$. Let $t > t_0$

and set $n = m(t)$. Then $\gamma(t,t) = \gamma(t_0, j_s)$ for some j_s. That is,

$$\frac{|A(t,d_k) \cap (X/F_n)|}{|X/F_n|} = \frac{|A(j_s,d_k) \cap (X/F_m)|}{|X/F_m|} = \frac{|A(j_s,d_k) \cap (X/F_n)|}{|X/F_n|},$$

using (2.4) again. Further, $A(t,d_k) \subset A(t_0,d_k) \subset A(j_s,d_k)$ so that we have

$$|A(t,d_k) \cap (X/F_n)| = |A(t_0,d_k) \cap (X/F_n)|,$$

and this holds for all $t \geq t_0$.
We can repeat this argument for each fan tower in the separating set. Let $\{s_1, \dots, s_\ell\}$ be the separating set and let $s_i = \{F_{in}\}$. Hence there exist an N and a T such that $|A(t,d_k) \cap (X/F_{in})| = |A(T,d_k) \cap (X/F_{in})|$ for all $1 \leq i \leq \ell$ and all $t \geq T$. By the second property of a separating set we have $A(t,d_k) = A(T,d_k)$ for all $t \geq T$. This contradicts our choice of d_k.
Hence we have a T such that $A(t,d_j) = A(T,d_j)$ for all $t \geq T$ and all $d_j \in D$. Thus $\hat{q}_t(\alpha) = \pm\hat{q}_T(\alpha)$ for all α in the union of the $A(T,d_j)$, that is, for all α with $\hat{q}(\alpha) \neq 0$. By our early reduction, $\hat{q}(\alpha) \neq 0$ iff $\hat{q}_T(\alpha) \neq 0$. Thus $\hat{q}_t(\alpha) = \pm\hat{q}_T(\alpha)$ for all α with $\hat{q}_T(\alpha) \neq 0$ and also q_t divides q_T. By (2.6) we obtain $(q_t) = (q_T)$, for all $t \geq T$. □

Corollary 2.8. *Let R be a real (but necessarily reduced) Witt ring. If R has finite chain length then ACC holds for principal ideals generated by odd dimensional forms.*

Proof. Every ideal containing an odd dimensional form contains the torsion ideal R_t by [7, 1.5]. Hence passing to the reduced Witt ring maintains a tower of principal ideals generated by odd dimensional forms. This reduced tower stabilizes by (2.7). Hence the original tower stabilizes. □

Corollary 2.9. *Let (G,X) be a space of orderings. Let $\mathcal{S}$ denote the collection of subsets of G of order n. If X has finite chain length then any tower*

$$H(S_1) \subset H(S_2) \subset \cdots H(S_k) \subset \cdots .$$

with each $S \in \mathcal{S}$, stabilizes.

Proof. Suppose $S_i = \{a_{i1}, \dots, a_{in}\}$. Set $q_i = \langle\langle a_{i1}, \dots, a_{in}\rangle\rangle + 1$. Then $\hat{q}_i(X) = \{1, 2^n + 1\}$ and $\hat{q}_i^{-1}(2^n + 1) = H(S_i)$. Thus $\hat{q}_{i+1}(\alpha)$ divides $\hat{q}_i(\alpha)$ for all $\alpha \in X$. So q_{i+1} divides q_i by [7, 1.7]. We thus have a tower of principal ideals $(q_1) \subset (q_2) \subset \cdots$. This stabilizes by (2.7) and so the tower of $H(S_i)$'s also stabilizes. □

3. Factorization.

Anderson and Valdes-Leon [1] have several notions of an associate in a commutative ring R. We need three of these. Two elements a and b are *associates* if their principal ideals are equal, $(a) = (b)$. They are *strong associates* if $a = bu$, for some unit $u \in R$. Lastly, a and b are *very strong associates* if $(a) = (b)$ and either $a = b = 0$ or $a \neq 0$ and $a = br$ implies r is a unit.

An non-unit a is *irreducible* if $a = bc$ implies either b or c is an associate of a. Similarly, a is *strongly irreducible* (*very strongly irreducible*) if $a = bc$ implies either b or c is a strong associate (respectively, very strong associate) of a. Lastly, R is *atomic* if every non-zero non-unit of R can be written as a finite product of irreducible elements. Define *strongly atomic* and *very strongly atomic* similarly.

Proposition 3.1. *Let R be a reduced Witt ring and let $a, b \in R$. Then a, b are associates iff a, b are strong associates. In particular, R is atomic iff R is strongly atomic.*

Proof. Strong associates are always associates so we check the converse. Suppose $(a) = (b)$. Write $a = bx$ and $b = ay$. Then $a = axy$ and $a(1 - xy) = 0$. Let $Z = \{\alpha \in X : \hat{a}(\alpha) = 0\}$. Then for all $\alpha \notin Z$ we have $\hat{x}(\alpha) = \pm 1$. From $a = bx$ and (2.6) we get $\langle d \rangle a = b$ for some $d \in G$. Clearly $\langle d \rangle$ is a unit. □

Strong associates need not be very strong associates in a reduced Witt ring. If $\pm 1 \neq g \in G$ then $\langle 1, g \rangle$ is not even a very strong associate of itself. Namely, $\langle 1, g \rangle = \langle 1, g \rangle \langle 1, 1, -g \rangle$ and $\langle 1, g \rangle \neq 0$ and $\langle 1, 1, -g \rangle$ is not a unit. So, except for $R = \mathbb{Z}$, R will not be very strongly atomic.

Corollary 3.2. *Let R be a real Witt ring (not necessarily reduced) and suppose R has finite chain length.*

(1) *Every odd dimensional form can be written as a finite product of irreducible forms.*
(2) *If R is reduced then R is atomic.*

Proof. These are standard consequences of (2.8) and (2.7), see [1, 3.2]. □

We are unable to prove the converse to (3.2)(2) for all reduced Witt rings R. However, we can prove the converse for a wide class of rings. For this we need Marshall's notion of a sheaf product [11]. Start with a non-empty Boolean space I, a collection of reduced Witt rings R_C, one for each clopen $C \subset I$ and a collection of ring homomorphisms $\mathrm{res}_{C.D} : R_C \to R_D$, defined whenever $D \subset C$ are clopen in I. We assume the usual sheaf properties, namely,

(1) $R_\emptyset = \mathbb{Z}/2\mathbb{Z}$ and $R_C \neq \mathbb{Z}/2\mathbb{Z}$ if $C \neq \emptyset$.
(2) $\mathrm{res}_{C,C}$ is the identity map on C.
(3) If $E \subset D \subset C$ then $\mathrm{res}_{C,E} = \mathrm{res}_{D,E}\mathrm{res}_{C,D}$.
(4) If $C = \cup_j C_j$ and if $r_j \in R_j$ are given such that

$$\mathrm{res}_{C_j, C_j \cap C_k}(r_j) = \mathrm{res}_{C_k, C_j \cap C_k}(r_k),$$

for all j, k, then there exists a unique $r \in R_C$ such that $\mathrm{res}_{C,C_j}(r) = r_j$, for all j.

For fixed $i \in I$ we form the *stalk*

$$R_i = \varinjlim_{i \in C} R_C.$$

Each R_i is a reduced Witt ring. We call the reduced Witt ring R_I the *sheaf product* of the R_i's and write $R_I = \prod_{i \in I} R_i$. When I is finite and discrete this is the usual product of Witt rings.

We next define a sequence of classes of reduced Witt rings (which is slightly different from the sequence of Marshall [11, p. 219]). Let $\mathcal{C}_1$ denote the class of finitely generated reduced Witt rings. Inductively define $\mathcal{C}_n$ to be sheaf products of $R_i[E^i]$, where E^i is a group of exponent two (not necessarily finite) and $R_i \in \mathcal{C}_m$ for some $m < n$. Lastly, let $\mathcal{C}_\omega$ be the union of all $\mathcal{C}_n$. This is a large class. Already $\mathcal{C}_2$ contains all SAP reduced Witt rings and $\mathcal{C}_\omega$ contains all reduced Witt rings where X has only a finite number of accumulation points [11, 8.17].

We will prove that $R \in \mathcal{C}_\omega$ atomic implies R has finite chain length. We begin with a lemma.

Lemma 3.3. *Let $S = R[E]$ and let $T \subset G_S$ be a fan of finite index. Set $T_0 = T \cap G_R$.*

(1) *T_0 is a fan in G_R.*
(2) *Suppose $X_R/T_0 = \{P, Q\}$. Then X_S/T consists of extensions of P, Q to S. If $x \in G_S \setminus G_R$ then either none, exactly half or all of the extensions of P that lie in X_S/T make x positive.*

Proof. (1) Write $T = T_0 H$ for some subgroup H of G_S with $H \cap G_R = 1$. Extend H to subgroup L of G_S such that $G_S = G_R \times L$. Suppose $P \subset G_R$ is a subgroup of index 2, containing T_0 but not -1. Then PL is a subgroup of index at most 2 containing T. If $-1 \in PL$ then for some $p \in P$ and $y \in L$ we have $-p = y \in P \cap L = 1$. But then $-1 = p \in P$, a contradiction. Thus PL is an ordering in G_S. It is easy to check that P is then an ordering in G_R. This shows T_0 is a fan.

(2) The first statement is clear. Suppose $P_1, \ldots P_m, Q_1, \ldots, Q_m$ are the extensions of P, Q that lie in X_S/T. Pick $a \in G_R$ with $\hat{a}(P) = 1$ and $\hat{a}(Q) = -1$. Let k be the number of P_i for which x is positive. From the easy half of the Representation Theorem [11, 7.13]

$$\sum_{\alpha \in X_S/T} \mathrm{sgn}_\alpha \langle\langle a, x \rangle\rangle \equiv 0 \pmod{2m}$$

$$4k \equiv 0 \pmod{2m}.$$

So m divides $2k$ and clearly $k \leq m$. Hence $k = 0, \frac{1}{2}m$ or m. $\square$

Our proof that $R \in \mathcal{C}_\omega$ atomic implies finite chain length is not the usual induction argument since we are unable to show $R[E]$ atomic implies R atomic. Instead we explicitly construct a form which does not factor into a finite product of irreducibles. Unfortunately, the construction requires considerable notation. We introduce this notation by first looking at a special case. Let $*$ denote a group ring extension. A ring in $\mathcal{C}_n$ looks like

$$\begin{aligned} R &= \prod_{\alpha\in A_1} W(\alpha)^* \\ &= \prod_{\alpha\in A_1}\Big(\prod_{\beta\in A_2(\alpha)} W(\alpha,\beta)^*\Big)^* \\ &= \prod_{\alpha\in A_1}\Big(\prod_{\beta\in A_2(\alpha)}\Big(\prod_{\gamma\in A_3(\alpha,\beta)} W(\alpha,\beta,\gamma)^*\Big)^*\Big)^*, \end{aligned}$$

where each $A_1, A_2(\alpha)$ and $A_3(\alpha,\beta)$ is a Boolean space and each $W(\alpha,\beta,\gamma)$ is in $\mathcal{C}_m$, for some $m \le n-3$.
Suppose we want to single out the product over $A_3(\alpha_0,\beta_0)$, for some particular α_0 and β_0. We set :

$$\begin{aligned} R_1 &= \prod_{\gamma\in A_3(\alpha_0,\beta_0)} W(\alpha_0,\beta_0,\gamma)^* \\ R_2 &= \prod_{\substack{\beta\in A_2(\alpha_0)\\ \beta\ne\beta_0}}\Big(\prod_{\gamma\in A_3(\alpha_0,\beta)} W(\alpha_0,\beta,\gamma)^*\Big)^* \\ R_3 &= \prod_{\substack{\alpha\in A_1\\ \alpha\ne\alpha_0}}\Big(\prod_{\beta\in A_2(\alpha)}\Big(\prod_{\gamma\in A_3(\alpha,\beta)} W(\alpha,\beta,\gamma)^*\Big)^*\Big)^*. \end{aligned}$$

Then $R = ((R_1^* \sqcap R_2^*)^* \sqcap R_3^*)^*$.
We will want to single out the first infinite sheaf product. We have:

$$R = ((\dots((R_1^* \sqcap R_2^*)^* \sqcap R_3^*)^* \sqcap \dots)^* \sqcap R_s^*)^*,$$

with R_1 an infinite sheaf product, say

$$R_1 = \prod_{\delta\in A} W(\delta)^*,$$

and each $W(\delta)$ in some $\mathcal{C}_m$, $m \le n-s$. We will need explicit extension groups. We use the notation

$$R = (\dots((R_1[E^1] \sqcap R_2[F^1])[E^2] \sqcap R_3[F^2])[E^3] \sqcap \dots \sqcap R_s[F^{s-1}])[E^s].$$

We further take $\{t^i_j\}$ as generators of E^i.

Lastly, we need notation to express the orderings on R. Let X_i denote X_{R_i}. Let $X_1(\epsilon_1)$ denote the extensions of X_1 to $R_1[E^1]$. Here ϵ_1 is an arbitrary choice of signs. The extension is determined by the values $\epsilon_1(t_j^1) \in \{\pm 1\}$. To save on indices we will write $\epsilon_1(j)$ for $\epsilon_1(t_j^1)$. Next, $X_2(\eta_1)$ denotes the extensions from R_2 to $R_2[F^1]$. $X_1(\epsilon_1, \epsilon_2)$ denotes the extensions from R_1 to $(R_1[E^1] \sqcap R_1[F^1])[E^2]$, with ϵ_2 a sign choice for E^2. Continue with this pattern. We obtain for X_R

$$\bigcup_{\epsilon,\eta}[X_1(\epsilon_1, \dots, \epsilon_s) \cup X_2(\eta_1, \epsilon_2, \dots, \epsilon_s) \cup X_3(\eta_2, \epsilon_3, \dots, \epsilon_s) \cup \dots \cup X_s(\eta_{s-1}, \epsilon_s)].$$

THEOREM 3.4. *Suppose $R \in \mathcal{C}_\omega$. The following are equivalent:*

(1) *R has finite chain length.*
(2) *R has ACC on principal ideals.*
(3) *R is atomic.*

Proof. We need only show R atomic implies R has finite chain length, by (2.7) and (3.2). Suppose $R \in \mathcal{C}_n$ and let s be the first level (if any) with an infinite sheaf product. We follow the above notation. Fix some $\delta_0 \in A$ and define $a \in G_{R_1}$ with -1 in the δ_0 coordinate and 1 in the other coordinates. Set

$$b = ((\dots(a, -1), -1), \dots), -1) \in G_R,$$

and set $q = \langle b, t_1^1, bt_1^1 \rangle$.
Let X_δ be the orderings on $W(\delta)^*$ so that $X_1 = \cup X_\delta$. Set $C = \hat{q}^{-1}(3)$. Then:

$$C = \bigcup_{\substack{\epsilon,\eta \\ \epsilon_1(1)=1}} \left[\left(\bigcup_{\delta \neq \delta_0} X_\delta \right)(\epsilon_1, \dots, \epsilon_s) \cup X_2(\eta_1, \epsilon_2, \dots, \epsilon_s) \cup \dots \cup X_s(\eta_{s-1}, \epsilon_s)) \right].$$

We are assuming R is atomic, so let $q = \varphi_1 \cdots \varphi_r$ with each φ_i irreducible. We may assume $\hat{\varphi}_i(X) = \{3, -1\}$ by (2.6). Set $D_i = \hat{\varphi}_i^{-1}(3)$. Note $D_i \subset C$. We will show that in fact one of the φ_i factors and hence that no sheaf product in R is infinite.
Our first goal is to show that each D_i consists of all extensions, with t_1^1 positive, of some subset of X_1. Pick $P \in X_{\delta_0}$ and $Q \in X_\delta$ with $\delta \neq \delta_0$. Fix some k and j. Let

$$\begin{aligned} e^k &= sp\{t_1^k, \dots, t_{j-1}^k, t_{j+1}^k, \dots\} \\ e^1 &= sp\{t_2^1, t_3^1, \dots\}. \end{aligned}$$

Let T be the fan

$$(\dots(((P \cap Q)[e^1] \sqcap G_{R_2}[F^1])[E^2] \dots)[e^k] \sqcap \dots \sqcap G_{R_s}[F^{s-1}])[E^s].$$

Then X/T has 8 orderings, namely the extensions of P and Q with all t^i_ℓ positive except possibly t^1_1 and t^j_k. Write these orderings as $P(\pm 1, \pm 1)$ and $Q(\pm 1, \pm 1)$, where the first coordinate gives the sign of t^1_1 and the second gives the sign of t^k_j.
$C \cap (X/T) = \{Q(1, \pm 1)\}$ so that $|C \cap (X/T)| = 2$. To ease notation slightly, write D for one of the D_i. Let $w = |D \cap (X/T)|$. Then by the easy part of the Representation Theorem we have:

$$\begin{aligned} \sum_{\gamma \in X/T} \hat{\varphi}(\gamma) &\equiv 0 \pmod{|X/T|} \\ 3w - (8 - w) &\equiv 0 \pmod 8 \\ w &\equiv 0 \pmod 2. \end{aligned}$$

As $D \cap (X/T) \subset C \cap (X/T)$ we have $D \cap (X/T)$ is either empty or all of $C \cap (X/T)$.
Suppose for some k and j we are in the second case, $D \cap (X/T) = C \cap (X/T)$. Choose another pair g, h. Pick the fan T' generated over $P \cap Q$ by E^i for $i \neq 1, k, g$, the same e^1 as before and

$$\begin{aligned} e^{k'} &= sp\{t^k_1, \dots, t^k_{j-1}, -t^k_j, t^k_{j+1} \dots\} \\ e^{g'} &= sp\{t^g_1, \dots, t^g_{h-1}, t^g_{h+1}, \dots\}. \end{aligned}$$

Then X/T' has 8 orderings, namely the extensions of P and Q with all t^i_ℓ positive except t^k_j negative and t^1_1, t^g_h arbitrary. Write these as $P(\pm 1, -1, \pm 1)$ and $Q(\pm 1, -1, \pm 1)$ with the first coordinate the sign of t^1_1, the second coordinate indicating that t^k_j is negative and the third coordinate the sign of t^g_h.
Again $C \cap (X/T')$ consists of two orderings, $Q(1, -1, \pm 1)$. And as before we get that $D \cap (X/T')$ is either empty or all of $C \cap (X/T')$. But $Q(1, -1, 1)$ is the same ordering that was denoted by $Q(1, -1)$ before (that is, with t^1_1 positive, t^k_j negative and all other t's positive). Hence we have $D \cap (X/T') = C \cap (X/T')$.
We continue to assume $D \cap (X/T) = C \cap (X/T)$. If we repeat this argument (first with a fan having t^k_j and t^g_h negative) we get that any extension Q with t^1_1 positive and only a finite number of t^i_ℓ negative is in D. Now $D = \hat{\varphi}^{-1}(3)$ and the entries of φ involve only a finite number of t^i_ℓ. Hence we have that any extension of Q with t^1_1 positive is in D.
The assumption that $D \cap (X/T) \neq \emptyset$ means we are assuming some extension of Q with t^1_1 positive is in D. From this we conclude that all such extensions are in D.
Let X^*_1 denote the orderings on $R_1[E^1]$, namely the extensions ϵ_1 of X_1. Write $D|X^*_1$ for the orderings in D restricted to $R_1[E^1]$. We have shown that $D|X^*_1$ consists of all extensions, with t^1_1 positive, of some subset (call it $D|X_1$) of X_1.
Each factor φ_i of q has its set D_i. We have $C = \cup D_i$ and

$$\cup(D_i|X_1) = C|X_1 = \bigcup_{\substack{\delta \in A \\ \delta \neq \delta_0}} X_\delta.$$

A is infinite so some $D_i|X_1$ meets at least two X_δ's. For simplicity, call this D_i simply D and the corresponding form φ. Suppose $D|X_1$ meets X_{δ_1} and X_{δ_2}, $\delta_1 \neq \delta_2$. Set

$$D_0 = \bigcup_{\epsilon(1)=1} [(D|X_1) \cap X_{\delta_1}](\epsilon_1) \subset X_1^*.$$

In words, D_0 consists of the extensions for X_{δ_1} that lie in $D|X_1^*$. We will use D_0 to construct a factor of φ.

Let $f : X_1^* \to \mathbb{Z}$ by $f(P) = 3$ if $P \in D_0$ and $f(P) = -1$ if $P \notin D_0$. We want to use the Representation Theorem [11,7.13] to show f is represented by a form in $R_1[E^1]$. Let $T \subset G_{R_1}E^1$ be a fan of finite index. Then $T_1 = T \cap G_{R_1}$ is a fan in G_{R_1} by (3.3)

Case 1 : $(X_1/T_1) \subset X_\delta$ for some $\delta \in A$.

Here $X_1^*/T = (X_\delta/T_1)(\epsilon)$, over some set of extensions ϵ to E^1. If $\delta \neq \delta_1$ then $f(P) = -1$ for all $P \in (X_1^*/T)$ since D_0 only has extensions from X_{δ_1}. Thus

$$\sum_{P \in X_1^*/T} f(P) = -|X_1^*/T| \equiv 0 \pmod{|X_1^*/T|}.$$

If $\delta = \delta_1$ then $P \in D_0$ iff $P \in D|X_1^*$ iff some (equivalently, every) extension, with t_1^1 positive, of P to X_R lies in D iff $\hat{\varphi}(P) = 3$. So $f(P) = \hat{\varphi}(P)$ for all $P \in X_1^*/T$. We obtain

$$\sum_{P \in X_1^*/T} f(P) = \sum_{P \in X_1^*/T} \hat{\varphi}(P) \equiv 0 \pmod{|X_1^*/T|}.$$

Case 2 : $(X_1/T_1) \not\subset X_\delta$ for some $\delta \in A$.

Here we must have $|X_1/T_1| = 2$ by [11, 8.12] Write $X_1/T_1 = \{P_\alpha, P_\beta\}$ where α, β are distinct elements of A and $P_\alpha \in X_\alpha$ and $P_\beta \in X_\beta$. Then X_1^*/T consists of some set of extensions, to E^1, applied to P_α and P_β.

Again, if neither α nor β are δ_1 then all $f(P) = -1$ and we are done. So say $\alpha = \delta_1$ (and so $\beta \neq \delta_1$). If $P_\alpha \notin D|X_1$ then no extension is in D_0 and all $f(P) = -1$ again. So suppose $P_\alpha \in (D|X_1) \cap X_{\delta_1}$. Since $P_\beta \notin X_{\delta_1}$ no extension of P_β in X_1^*/T is in D_0. This is half of X_1^*/T. The other half consists of extensions of P_α and by (3.3) either none, exactly half or all of these extensions make t_1^1 positive, and hence lie in D_0. Thus $|D_0 \cap (X_1^*)| = d|X_1^*/T|$, where d is either (i) 0, or (ii) $\frac{1}{4}$ or (iii) $\frac{1}{2}$. In case (i) we have

$$\sum_{P \in X_1^*} f(P) = -|X_1^*/T| \equiv 0 \pmod{|X_1^*/T|}.$$

In case (ii) we have

$$\sum_{P \in X_1^*} f(P) = \tfrac{1}{4}|X_1^*/T| \cdot 3 + \tfrac{3}{4}|X_1^*/T| \cdot (-1) \equiv 0 \pmod{|X_1^*/T|}.$$

In case (iii) we have

$$\sum_{P\in X_1^*} f(P) = \tfrac{1}{2}|X_1^*/T|\cdot 3 + \tfrac{1}{2}|X_1^*/T|\cdot(-1) \equiv 0 \pmod{|X_1^*/T|}.$$

Thus in all cases we have $\sum f(P) \equiv 0 \pmod{|X_1^*/T|}$. By the non-trivial half of the Representation Theorem we have $f = \hat{\psi}$ for some form $\psi \in R_1[E^1]$. By construction $\hat{\psi}(X_1^*) = \{3,-1\}$ and $\hat{\psi}^{-1}(3) = D_0 < D$. Hence by [7, 1.7] ψ is a proper divisor of φ. Hence φ is not irreducible, a contradiction.
We thus have if $R \in \mathcal{C}_n$ is atomic then all sheaf products are finite. Hence $\mathrm{cl}(R) < \infty$, using [12, 4.2.1]. $\square$

Corollary 3.5. *Let $R \in \mathcal{C}_\omega$. If $R[E]$ is atomic then so is R.*

Proof. $R[E]$ atomic implies $R[E]$ has finite chain length by (3.4). Then , as $\mathrm{cl}(R[E]) = \mathrm{cl}(R)$, R has finite chain length and so is atomic by (3.2). $\square$

It is unknown if the reduced Witt rings of finite stability index lie in $\mathcal{C}_\omega$ so the following may improve (3.4), although (3.4) includes many atomic Witt rings with X infinite.

Proposition 3.6. *Suppose R has finite stability index. The following are equivalent:*

(1) *R has finite chain length.*
(2) *R has ACC on principal ideals.*
(3) *R is atomic.*
(4) *X is finite.*

Proof. (1) and (4) are equivalent by [10] (first shown, in the field case in [4]). As in the proof of (3.4) we need only show (3) implies (1). Suppose the stability index of R is n. We can find a prime p congruent to 1 mod 2^n by Dirichlet's Theorem. R is atomic so $p = \varphi_1\cdots\varphi_t$ for some irreducible elements φ_i. Note that for each i we have $|\hat{\varphi}_i(X)| = \{p,1\}$. Let $A_i = \hat{\varphi}_i^{-1}(\pm p)$. The A_i's form a clopen cover of X.
We wish to show R has finite chain length. So suppose we have a tower

$$H(a_1) > H(a_2) > H(a_3) > \cdots .$$

First suppose there is an s, $1 \le s \le t$ and a k such that $A_s \cap H(a_k)$ is a non-empty, proper subset of A_s. Define $f : X \to \mathbb{Z}$ by

$$f(\alpha) = \begin{cases} p, & \text{if } \alpha \in A_s \cap H(a_k) \\ 1, & \text{if } \alpha \notin A_s \cap H(a_k). \end{cases}$$

Let T be a fan, $|X/T| = 2^m$, where $m \le n$ by definition of the stability index. Set $w = |A_s \cap H(a_k) \cap (X/T)|$. Then

$$\sum_{\alpha\in X/T} f(\alpha) = wp + (2^m - w) = w(p-1) \equiv 0 \pmod{2^m},$$

since $p-1$ is a multiple of 2^n. By the Representation Theorem, $f = \hat{\psi}$ for some form ψ. Then $\hat{\psi}(\alpha)$ divides $\hat{\varphi}_s(\alpha)$ for all α and for $\alpha \in A_s \setminus H(a_k)$, $\hat{\psi}(\alpha) \neq \pm\hat{\varphi}_s(\alpha)$. So, using [7, 1.7], we have ψ is proper divisor of φ_s, which is impossible.
Thus there does not exists a pair s, k such that $H(a_k) \cap A_s$ is a non-empty, proper subset of A_s. That is, for all i, j we have $H(a_i) \cap A_j \neq \emptyset$ implies $A_j \subset H(a_i)$. The A_j's cover X so each $H(a_i)$ is a union of A_j's. Let $n(i)$ be the number of A_j's required to cover $H(a_i)$. Then $1 \leq n(i+1) < n(i) \leq t$ for all i. Thus the tower is finite and we are done. □

4. Irreducible elements.

We look at some examples to illustrate factorization in reduced Witt rings.

Proposition 4.1. *If $1 \neq a \in G$ then $\langle 1, -a\rangle$ is irreducible in R.*

Proof. Suppose $\langle 1, -a\rangle = q\varphi$ in R. We may assume q is even dimensional and φ is odd dimensional. If $a <_\alpha 0$ then $2 = \hat{q}(\alpha)\hat{\varphi}(\alpha)$. Thus $\hat{q}(\alpha) = \pm 2 = \pm\mathrm{sgn}_\alpha\langle 1, -a\rangle$, for all α with $\mathrm{sgn}_\alpha\langle 1, -a\rangle \neq 0$. By (2.6) there exists a $d \in G$ such that $\langle d\rangle\langle 1, -a\rangle = q$ and so q is an associate of $\langle 1, -a\rangle$. □

Example. If $R \neq \mathbb{Z}$ then factorization into irreducible elements is not unique. Namely, if $a \neq \pm 1$ then $\langle 1, -a\rangle\langle 1, -a\rangle = \langle 1, 1\rangle\langle 1, -a\rangle$ gives two different factorizations of the Pfister form. This is quite different from the case of factoring odd dimensional forms. When X is finite there is unique factorization of odd dimensional forms if the ideal class group of R is trivial or, equivalently, the stability index is at most 2, by [6, 2.7] and [7, 1.17].

We next find the irreducible elements in $\mathbb{Z}[E_1]$. Note that any form q in this ring is associate to some $n + mt$ with $n \geq |m|$.

Proposition 4.2. *Let $q = n + mt \in \mathbb{Z}[E_1]$ with $n \geq |m|$. Then q is irreducible iff (n, m) or $(n, -m)$ equals one of the following:*

(1) $(1, 1)$
(2) $(2^k + 1, 2^k - 1)$*, for some* $k \geq 0$
(3) $(\frac{1}{2}(p+1), \frac{1}{2}(p-1))$*, for some odd prime* p*.*

Proof. Let q be irreducible. First suppose q is even dimensional. If both n and m are even then 2 is a factor of q. So we have n and m odd. If $n = \pm m$ then n is a factor of q and we must have $n = 1$. Thus $(n, m) = (1, \pm 1)$. We may thus suppose $n + m$ and $n - m$ are non-zero. Write $n + m = 2^g h$ and $n - m = 2^k \ell$ with h and ℓ odd and $g, k \geq 1$. Set

$$\varphi_1 = \tfrac{1}{2}(2^g + 2^k) + \tfrac{1}{2}(2^g - 2^k)t$$
$$\varphi_2 = \tfrac{1}{2}(h + \ell) + \tfrac{1}{2}(h - \ell)t.$$

Then $q = \varphi_1\varphi_2$ and φ_2 is odd dimensional and so not an associate of q. Thus φ_1 is an associate of q. If α is the ordering with t positive then $n + m = \hat{q}(\alpha) =$

$\pm\hat{\varphi}_1(\alpha) = \pm 2^g$. Since $n \geq -m$ we obtain $n+m = 2^g$ and $h = 1$. Similarly, taking signatures at the ordering β with t negative gives $\ell = 1$. If both g and k are at least 2 then n and m are even which is not possible. Suppose $n+m = 2^g$ and $n-m = 2$. Then we get case (2). The reverse , $n+m = 2$ and $n-m = 2^k$ gives case (2) for the pair $(n, -m)$.
Now suppose q is odd dimensional. If $n+m$ is composite, say $n+m = ab$ with $a, b > 1$, then set

$$\varphi_1 = \tfrac{1}{2}(a+1) + \tfrac{1}{2}(a-1)t$$
$$\varphi_2 = \tfrac{1}{2}(b+n-m) + \tfrac{1}{2}(b-n+m)t.$$

Then $q = \varphi_1\varphi_2$. Neither φ_1 nor φ_2 is an associate of q as $\hat{q}(\alpha) = ab$ while $\hat{\varphi}_1(\alpha) = a$ and $\hat{\varphi}_2(\alpha) = b$. Hence $n+m$ is not composite. Similarly, $n-m$ is not composite. If both $n+m$ and $n-m$ are prime then set

$$\varphi_1 = \tfrac{1}{2}(n+m+1) + \tfrac{1}{2}(n+m-1)t$$
$$\varphi_2 = \tfrac{1}{2}(n-m+1) + \tfrac{1}{2}(1-n+m)t.$$

We have $q = \varphi_1\varphi_2$. Neither φ_1 nor φ_2 is an associate of q as $\hat{q}(\alpha) = n+m$ while $\hat{\varphi}_2(\alpha) = 1$ and $\hat{q}(\beta) = n-m$ while $\hat{\varphi}_1(\beta) = 1$. Thus we must have $n+m = p$, p an odd prime, and $n-m = 1$ (or the reverse). This gives case (3).
It is straightforward to check the forms in cases (1) - (3) are irreducible. $\square$

Example. Already for $\mathbb{Z}[E_1]$, and in fact for any $R \neq \mathbb{Z}$, the number of irreducible factors in factorization of a given element can be arbitrarily large. For instance, $\langle 1, 1, t\rangle$ is irreducible (take $p = 3$ in (4.2)(3)) and $\langle 1, -t\rangle\langle 1, 1, t\rangle = \langle 1, -t\rangle$. Hence

$$\langle\langle 1, -t\rangle\rangle = \langle 1, 1\rangle\langle 1, 1, t\rangle^n\langle 1, -t\rangle$$

is a factorization into irreducible elements for any n. Again the situation is quite different if we consider only factorizations of odd dimensional forms. When X is finite, the number of irreducible factors in a factorization is uniquely determined iff the stability index is at most 3 and R has no factor of the type $(\mathbb{Z}^s)[E_2]$, with $s \geq 3$, see [7].

Notice that the even prime of $\mathbb{Z}$ remains irreducible in $\mathbb{Z}[E_1]$ while the odd primes of $\mathbb{Z}$ all factor in $\mathbb{Z}[E_1]$. This holds more generally.

Proposition 4.3. *Let $q \in R$ be irreducible.*

(1) *If q is even dimensional then q remains irreducible in $R[E_1]$.*
(2) *If q is odd dimensional then q remains irreducible in $R[E_1]$ iff q is not associate to $1 + 2q_0$, for some $q_0 \in R$.*

Proof. First say $q = 1 + 2q_0$, for some $q_0 \in R$. Since q is not a unit, there exists an $\alpha \in X_R$ with $\hat{q}(\alpha) \neq \pm 1$. Let α^+ and α^- denote the extensions of α to $R[E_1]$ with, respectively, t positive and t negative. Now

$$q = (1 + q_0\langle 1, t\rangle)(1 + q_0\langle 1, -t\rangle).$$

Neither factor is an associate of q as the first has signature 1 at α^- and the second has signature 1 at α^+. Thus q is not irreducible in $R[E_1]$.
Now suppose we have an irreducible q that factors in $R[E_1]$. We want to show q is odd dimensional and associate to some $1 + 2q_0$. Write $q = (a + b\langle 1, t\rangle)(c + d\langle 1, -t\rangle)$, with $a, b, c, d \in R$ and neither factor an associate of q. The coefficient of t, namely $bc - ad$, must be zero and so $q = ac + ad + bc$. Then

$$q = ac + 2bc = c(a + 2b) \tag{4.4}$$
$$= ac + 2ad = a(c + 2d). \tag{4.5}$$

As q is irreducible in R, (4.4) shows that either c or $a + 2b$ is an associate of q. We may assume c is the associate of q. Namely, if $a + 2b$ is the associate then rewrite q as

$$\begin{aligned} q &= ((c + 2d) + (-d)\langle 1, t\rangle)((a + 2b) + (-b)\langle 1, -t\rangle) \\ &\equiv (a' + b'\langle 1, t\rangle)(c' + d'\langle 1, -t\rangle). \end{aligned}$$

Then $c' = a + 2b$ is associate to q.
Write $uq = c$ for some unit $c \in R$. Equation (4.5) shows that either a or $c + 2d$ is an associate of q. Assume by way of contradiction that $vq = c + 2d$ for some unit $v \in R$. Note $(v - u)q = 2d$; set $\chi = v - u$. Let $Z = \{\alpha \in X_R : \hat{q}(\alpha) \neq 0\}$. From (4.4), $q = qu(a + 2b)$ so that $\hat{u} = \hat{a} + 2\hat{b}$ on Z. Similarly, from (4.5) $q = qva$ so that $\hat{v} = \hat{a}$ on Z. Thus, on Z, $\hat{\chi} = \hat{v} - \hat{u} = -2\hat{b}$. Now u and v are units and so have signatures ± 1 at all orderings. Thus $\hat{\chi}(X_R) \subset \{2, 0, -2\}$. If b is even dimensional then we must have $\hat{b} = 0$ on Z. Then $\hat{\chi} = 0$ on Z and $0 = q\chi = 2d$. But then $d = 0$ and the second factor of q, $c + d\langle 1, -t\rangle = c = uq$ is an associate of q, a contradiction. Hence b is odd dimensional. In particular, $\hat{b}$ is never zero. So $\hat{v} - \hat{u}$ is not zero on Z. We must have $\hat{v} = -\hat{u}$ (as $\hat{u}$ and $\hat{v}$ are always ± 1). So $\hat{\chi} = 2\hat{v}$ on Z. Then $2vq = q\chi = 2d$ and $vq = d$. But then the second factor of q is $c + d\langle 1, -t\rangle = uq + vq\langle 1, -t\rangle = q(u + v - vt) = -vtq$, an associate of q. This is impossible.
Hence we must have that q is an associate of a as well as c. Write $uq = c$ and $vq = a$ for units $u, v \in R$. Equation (4.4) gives $q = uq(a + 2b)$. If q is even dimensional then $a + 2b$ is odd dimensional and so a is odd dimensional. But a is an associate of the even dimensional q so a must be even dimensional, a contradiction.
We have then that q is odd dimensional. Then $q = uq(a+2b)$ implies $u(a+2b) = 1$. So $uvq = ua = 1 - 2ub$, as desired. $\square$

It can be shown that $a + bt \in R[E_1]$ is irreducible if $a + b$ is irreducible in R and $a - b$ is a unit. Thus in the factorization of (4.3) $1 + 2q_0 = (1 + q_0 + q_0 t)(1 + q_0 - q_0 t)$, both factors are irreducible. However, not every irreducible $a + bt \in R[E_1]$ satisfies $a + b$ irreducible and $a - b$ a unit. For instance, one may easily check that $q = \langle 1\rangle + \langle\langle t_1, t_2, t_3\rangle\rangle \in \mathbb{Z}[E_3]$ is irreducible. As a form in $R[E_1]$, where $R = \mathbb{Z}[E_2]$, we have $q = a + bt_3$ with $a = \langle 1\rangle + \langle\langle t_1, t_2\rangle\rangle$ and $b = \langle\langle t_1, t_2\rangle\rangle$. Then $a - b$ is a unit but $a + b = 1 + 2\langle\langle t_1, t_2\rangle\rangle = (1 - \langle\langle t_1, t_2\rangle\rangle)^2$. In fact, we have been unable to determine the irreducible elements of $R[E_1]$ in terms of the irreducibles of R. For products, we can determine only the irreducible odd dimensional forms.

Proposition 4.6. *If $R = R_1 \sqcap R_2$ and $(a, b) \in R$ is odd dimensional then (a, b) is irreducible iff a is irreducible in R and b is a unit or the reverse, a is a unit and b is irreducible.*

Proof. We have $(a, b) = (a, 1)(1, b)$. So (a, b) irreducible implies either a or b is a unit. Say b is a unit. If $a = xy$ then $(a, b) = (x, b)(y, 1)$, so a must be irreducible in R. □

References

1. D. D. Anderson, S. Valdes-Leon, *Factorization in commutative rings with zero divisors*, Rocky Mtn. J. Math. **26** (1996), 439–480.
2. I. Bonnard, *Un critère pour reconnaître les fonctions algébriquement constructibles*, J. reine angew. Math. **526** (2000), 61–88.
3. L. Bröcker, *On basic semi-algebraic sets*, Expositiones Math. **9** (1991), 289–334.
4. T. Craven, *Stability in Witt rings*, Trans. Amer. Math. Soc. **225** (1977), 227–242.
5. ———, *Characterizing reduced Witt rings of fields*, J. Algebra **53** (1978), 68–77.
6. R. Fitzgerald, *Ideal class groups of Witt rings*, J. Algebra **124** (1989), 506–520.
7. ———, *Picard groups of Witt rings*, Math. Z. **206** (1991), 303–319.
8. ———, *Half factorial Witt rings*, J. Algebra **155** (1993), 127–136.
9. M. Knebusch, A. Rosenberg, R. Ware, *Structure of Witt rings, quotients of abelian group rings and orderings of fields*, Bull. Amer. Math. Soc. **77** (1971), 205–210.
10. M. Marshall, *The Witt ring of a space of orderings*, Trans. Amer. Math. Soc. **258** (1980), 505–521.
11. ———, *Abstract Witt Rings*, Queen's Papers in Pure and Applied Math., vol. 57, 1980.
12. ———, *Spaces of Orderings and Abstract Real Spectra*, Lecture Notes in Math., vol. 1636, Springer-Verlag, Berlin/Heidelberg/New York, 1996.

Robert W. Fitzgerald
Southern Illinois University
Carbondale, IL 62901-4408
USA
rfitzg@math.siu.edu

The Zassenhaus Decomposition for the Orthogonal Group: Properties and Applications

Alexander Hahn[1]

Received: May 29, 2001

Communicated by Ulf Rehmann

Abstract. Zassenhaus [17] constructed a decomposition for any element in the orthogonal group of a non-degenerate quadratic space over a field of characteristic not 2 and used it to provide an alternative description of the spinor norm. This decomposition played a central role in the study of question of the length of an element in the commutator subgroup of the orthogonal group with respect to the generating set of all elementary commutators of hyperplane reflections. See Hahn [6]. The current article develops the fundamental properties of the Zassenhaus decomposition, e.g., those of uniqueness and conjugacy, and applies them to sharpen and expand the analysis of [6].

2000 Mathematics Subject Classification: 20G15, 20G25, 20F05.

1. Introduction. We begin with a discussion of the length question just mentioned. For the moment, consider any group G along with a set of generators A (not containing the identity element of G) that satisfies $A^{-1} = A$. Of all the factorizations of an element $\sigma \in G$ as a product of elements from A choose one that involves the smallest number of factors. This smallest number is defined to be the length $\ell(\sigma)$ of σ. One very basic question - necessarily in the context of specific examples - is this: are there parameters attached to σ from which $\ell(\sigma)$ can be read off?

A number of theorems have responded to this question. For G a Weyl group - or more generally a Coxeter group - and A an appropriate set of hyperplane reflections, see Humphreys [7]. Refer to Dyer [3] for a recent result in this

[1]The author wishes to thank the algebraists of Louisiana State University for their splendid organization of *Quadratic Forms 2001* and their warm hospitality throughout the conference.

context. For G a classical group and A a set of canonical elements coming from the underlying geometry, see Hahn-O'Meara [5] for a comprehensive treatment of the theorems of Dieudonné, Wall, and others. For G a classical group and A a set of generators coming from a single conjugacy class of elements, see Ellers-Malzan [2] and Knüppel [10]. In a related direction, interesting codes have been constructed starting with $G = SL_2(\mathbb{Z})$ and carefully selected A. See Margulis [12, 13] and Rosenthal-Vontobel [16] for details. The connection with the length problem is provided by the associated Cayley graph and its diameter.

Example 1. Let G be the symmetric group on $\{1, \dots, n\}$ and let A be the set of transpositions. Let $k(\sigma)$ be the the number of cycles of σ including the trivial cycles. Then $\ell(\sigma) = n - k(\sigma)$.

The fact that $\ell(\sigma) \le n - k(\sigma)$ follows from the decomposition of σ into its disjoint cycles. The other inequality is a consequence of the fact that $k(\sigma\tau) = k(\sigma) \pm 1$ for any transposition τ. A similar (but more complicated) argument provides

Example 2. Let G be the alternating group on $\{1, \dots, n\}$ and let A be the set of three cycles, or equivalently, the set of elementary commutators of transpositions. This time, let $k(\sigma)$ be the number of cycles of odd cardinality again including the trivial cycles. Then $n - k(\sigma)$ is even and $\ell(\sigma) = \frac{1}{2}(n - k(\sigma))$.

We now turn to the orthogonal group and begin by recalling some of the basics. For the details, see [5], especially Sections 5.2A, 5.2B, Chapter 6 (all specialized to the orthogonal case $\Lambda = 0$) and Section 8.2A.

Let V be a non-zero, non-degenerate, $n-$dimensional quadratic space with symmetric bilinear form B over a field F with $\mathrm{char}(F) \ne 2$. Denote $B(x, x)$ by $Q(x)$ and $\frac{1}{2}Q(x)$ by $q(x)$. Check that $B(x, y) = q(x + y) - q(x) - q(y)$. Two vectors x and y are *orthogonal* if $B(x, y) = 0$. A non-zero vector x in V that is orthogonal to itself is *isotropic* and it is *anisotropic* otherwise. A non-degenerate plane that contains isotropic vectors is a *hyperbolic plane* and an orthogonal sum of hyperbolic planes is a *hyperbolic space*. If U and W are orthogonal subspaces that intersect trivially, then $U \oplus W$ is denoted $U \perp W$. The orthogonal complement of a subspace U of V is denoted by $U^\perp$, and the radical of U is defined by $\mathrm{Rad}\ U = U \cap U^\perp$. If W is a complement of $\mathrm{Rad}\ U$ in U, then W is non-degenerate and $U = \mathrm{Rad}\ U \perp W$ is a *radical splitting* of U. Any two such complements of $\mathrm{Rad}\ U$ are isometric.

Let $O_n(V)$ be the orthogonal group of V. For $\sigma \in O_n(V)$, let S be the subspace $S = (\sigma - 1_V)V$ of V. This S is the *space* of σ. Intuitively, this is where the "action" of σ is. In particular, there is no action on the orthogonal complement $S^\perp$ of S; the fact is that $S^\perp = \{x \in V \mid \sigma(x) = x\}$. Clearly, $\sigma = 1_V$ if and only if $S = 0$. It turns out that dim S is even if and only if $\sigma \in O_n^+(V)$, the subgroup of $O_n(V)$ consisting of the elements of determinant 1. If $\eta \in O_n(V)$ commutes with σ, then $\eta S = S$. We will "transfer" properties of S to σ. For example, σ is *non-degenerate*, *degenerate*, or *totally degenerate*, if S is non-degenerate, degenerate, or totally degenerate, that is, if the radical $\mathrm{Rad}\, S$ of S is, respectively, zero, non-zero, or S. In the same way, σ is *anisotropic* if S is anisotropic.

An element $\sigma \in O_n(V)$ is an involution if $\sigma^2 = 1$. It is easy to see that σ is an involution if and only if $\sigma_{|_S} = -1_S$. In particular, involutions have the form $\sigma = -1_S \perp 1_{S^\perp}$ and are non-degenerate. Let v be an anisotropic vector and define τ_v in $O_n(V)$ by

$$\tau_v(x) = x - B(x,v)q(v)^{-1}v \quad \text{for all} \quad x \in V \;.$$

Check that the space of τ_v is Fv and that $\tau_{v}{}_{|_{Fv}} = -1_{Fv}$. So τ_v is an involution. These involutions are the *hyperplane reflections* or *symmetries*.

Theorem 1. (Cartan-Scherk-Dieudonné) Let G be the group $O_n(V)$ and let A be the set of hyperplane reflections. If σ is not totally degenerate, then $\ell(\sigma) = \dim\, S$. If σ is totally degenerate, then $\ell(\sigma) = \dim\, S + 2$.

Theorem 1 in combination with Examples 1 and 2 calls attention to the length problem in the situation where G is the commutator subgroup $\Omega_n(V)$ of $O_n(V)$ and A the set of elementary commutators of symmetries. It seems surprising that this question did not receive scrutiny until recently. John Hsia first called attention to it in the case of a non-dyadic local field and it was solved in this context in Hahn [6]. The answer is not simply a modification of the conclusion of Theorem 1, as a comparison of Examples 1 and 2 might suggest. We will see that, unlike Theorem 1, it depends critically on the arithmetic of the field F.

2. The Zassenhaus Decomposition.

An element σ in $O_n(V)$ is *unipotent* if its minimal polynomial has the form $(X-1)^m$ for some positive integer m. A non-trivial unipotent element is degenerate and can, therefore, exist only if V is isotropic. The elements with minimal polynomial $(X-1)^2$ are precisely the non-trivial totally degenerate elements. A degenerate element σ with dim $S = 2$ is an *Eichler* transformation. Let S be a degenerate plane and put $S = Fu \perp Fv$ with $u \in \mathrm{Rad}\ S$ and $v \in S$. Define $\Sigma_{u,v} \in O_n(V)$ by

$$\Sigma_{u,v}(x) = x + B(x,v)u - B(x,u)v - q(v)B(x,u)u \quad \text{for all} \quad x \in V\,.$$

Then $\Sigma_{u,v}$ is an Eichler transformation and all Eichler transformations have this form. A totally degenerate Eichler transformation has minimal polynomial $(X-1)^2$ and one that is not totally degenerate has minimal polynomial $(X-1)^3$. In particular, all Eichler transformations are unipotent.

Let σ be any element in $O_n(V)$. Consider the subspace

$$X = \{x \in V \mid (\sigma - 1_V)^j x = 0 \ \text{ some } \ j\}$$

of V. This unique largest space on which σ acts as a unipotent transformation turns out to be non-degenerate. Let $R = X^\perp$. Then R is non-degenerate and $X = R^\perp$. Notice that $\sigma R^\perp = R^\perp$. So $\sigma R = R$ and hence $\sigma = \sigma_{|_{R^\perp}} \perp \sigma_{|_R}$. Put $\mu = \sigma_{|_{R^\perp}} \perp 1_R$ and $\rho = 1_{R^\perp} \perp \sigma_{|_R}$. Then

$$\sigma = \mu \cdot \rho$$

with μ unipotent and ρ non-degenerate with space R. This is the *Zassenhaus decomposition or splitting* of σ. Note that μ and ρ commute.

To develop the essential properties of the Zassenhaus splitting, we need the *Wall* form. Let $\sigma \in O_n(V)$. Define

$$(\ ,\ )_\sigma : S \times S \longrightarrow F$$

by the equation $(\sigma x - x, \sigma y - y)_\sigma = B(\sigma x - x, y)$ for all $\sigma x - x$ and $\sigma y - y$ in S. This is the *Wall form* on S. It is non-degenerate and bilinear, but it is almost never symmetric. In fact, $(\ ,\ )_\sigma$ is symmetric if and only if σ is an involution, and in this case, $(s, s')_\sigma = -\frac{1}{2}B(s, s')$ for all s, s' in S. Also, $(\ ,\ )_\sigma$ is alternating if and only if σ is totally degenerate.

The space S is now equipped with both the Wall form $(\ ,\ )_\sigma$ and the restriction of B. When the focus is on $(\ ,\ )_\sigma$, then S is denoted by S_σ. Similarly, the space S_1 of σ_1 in $O_n(V)$ is written S_{σ_1} when $(\ ,\ )_{\sigma_1}$ is under consideration, and analogously for σ_2. The spaces of orthogonal transformations μ, ρ, μ', ρ' and so on, will be denoted by U, R, U', R' and so on, with appropriate subscripts when the focus is on the Wall form.

The key facts are these. Let S_1 be a non-degenerate subspace of S_σ. Then there is a unique $\sigma_1 \in O_n(V)$ - the transformation belonging to S_1 - such that $S_{\sigma_1} = S_1$. Let S_2 be the right complement of S_1 in S_σ. Then S_2 is non-degenerate. If σ_2 is the transformation belonging to S_2, then $\sigma = \sigma_1\sigma_2$. Conversely, if $\sigma = \sigma_1\sigma_2$ with $S_1 \cap S_2 = 0$, then $S_\sigma = S_{\sigma_1} \perp S_{\sigma_2}$. This means that the Wall forms of both S_{σ_1} and S_{σ_2} are obtained by restricting the Wall form $(\ ,\ )_\sigma$ and that $(s_1, s_2)_\sigma = 0$ for all $s_1 \in S_1$ and $s_2 \in S_2$ (but it is not required that $(s_2, s_1)_\sigma = 0$). For example, if $\sigma = \mu\rho$ is the Zassenhaus splitting of σ, then because μ and ρ commute,

$$S_\sigma = U_\mu \perp R_\rho = R_\rho \perp U_\mu \,.$$

Another important fact asserts that elements σ and σ_1 in $O_n(V)$ are conjugate in $O_n(V)$ if and only if the spaces S_σ and S_{σ_1} are isometric.

To conclude this discussion of the Wall form, we note that the map

$$\Theta : O_n^+(V) \ \longrightarrow\ \overset{*}{F}/\overset{*}{F}{}^2$$

defined by $\Theta(\sigma) = (\text{disc } S_\sigma)\overset{*}{F}{}^2$, where disc S_σ is the discriminant of the space S_σ, provides one of the (equivalent) definitions of the spinor norm. Its kernel is denoted by $O'_n(V)$. All unipotent elements are in $O'_n(V)$. It is clear that $O'_n(V) \supseteq \Omega_n(V)$ and it is a standard fact that if V is isotropic, then $O'_n(V) = \Omega_n(V)$. A formula useful for computations is

$$\Theta(\sigma) = \Theta(\rho) = \det\, (\rho - 1_V)_{|R}\ \text{disc}\ R\,,$$

where ρ is the non-degenerate component of the Zassenhaus decomposition of σ and disc $R \in \overset{*}{F}/\overset{*}{F}{}^2$ is the discriminant of the space R relative to the form B.

Proposition 1. Let $\sigma = \mu\rho$ be the Zassenhaus splitting of an element $\sigma \in O_n(V)$. Then $S = U \perp R$, and

i) σ is in $\Omega_n(V)$ if and only if both μ and ρ are in $\Omega_n(V)$.

ii) An element in $O_n(V)$ commutes with σ if and only if it commutes with both μ and ρ.

Proof: Recall that $O_n(V)$ has non-trivial unipotent elements only if V is isotropic. So any non-trivial unipotent element of $O_n(V)$ is in $O'_n(V) = \Omega_n(V)$. This implies (i). As to (ii), observe that if $\eta \in O_n(V)$ commutes with σ, then η stabilizes $X = R^\perp$. So $\eta = \eta_{|R^\perp} \perp \eta_{|R}$ and it follows that η commutes with both μ and ρ. QED.

We next consider the question of the uniqueness of the Zassenhaus splitting. It is not difficult to construct situations of the following sort: a non-degenerate element σ and a non-trivial unipotent element μ_0 with $U_0 \subseteq S$ such that μ_0 commutes with σ and the space of $\rho_0 = \mu_0^{-1}\sigma$ is S. In such a situation, $\sigma = 1_V\sigma = \mu_0\rho_0$ are two different ways of writing σ as a commuting product of a unipotent element and a non-degenerate element. We will see that such situations are in essence the only obstruction to the uniqueness of the Zassenhaus splitting.

Let $\sigma = \mu\rho$ be the Zassenhaus splitting of $\sigma \in O_n(V)$. Suppose that $\sigma = \mu'\rho'$ is any factorization of σ with μ' unipotent, ρ' non-degenerate, and such that μ' and ρ' commute.

Denote by W the orthogonal complement $W = R'^\perp$ of the space R' of ρ'. By an application of Proposition 1 (ii), the elements μ, μ', ρ, and ρ' all commute with each other. In particular, σ commutes with ρ'. So $\sigma R' = R'$ and hence $\sigma W = W$. Therefore, $\sigma = \sigma_{|W} \perp \sigma_{|R'}$. The fact that $\rho'_{|W} = 1_W$, tells us that $\sigma_{|W} = \mu'_{|W}$. So σ is unipotent on W and hence $W \subseteq R^\perp$. Therefore, $R' = W^\perp \supseteq R$. Let T be the orthogonal complement of R in R'. Because R'

and R are both non-degenerate, $R' = T \perp R$ with T non-degenerate. Because $R^\perp$ is the largest space on which σ is unipotent, T is the largest space on which $\sigma_{|_{R'}}$ is unipotent. So

$$\sigma_{|_{R'}} = (\mu_{|_T} \perp 1_R)(1_T \perp \rho_{|_R})$$

is the Zassenhaus splitting of $\sigma_{|_{R'}}$. Notice that $1_T \perp \rho_{|_R} = \rho_{|_{R'}}$ and hence that $\mu_{|_T} \perp 1_R = \mu_{|_{R'}}$. Since μ' and ρ' commute with with both ρ' and ρ, it follows that μ' and ρ' stabilize the spaces R', R, and therefore T. So

$$\mu_{|_T} = \sigma_{|_T} = (\mu'_{|_T})(\rho'_{|_T}) \ .$$

Therefore, $\rho'_{|_T}$ is a product of two commuting unipotent transformations. So $\rho'_{|_T}$ is unipotent. If T where to be non-zero, then ρ' would fix a non-zero vector in T. But this is impossible, because ρ' is non-degenerate with space R'. So $T = 0$. Hence $R' = R$ and $W = R^\perp$. This means that $\mu'_{|_{R^\perp}} = \sigma_{|_{R^\perp}} = \mu_{|_{R^\perp}}$ and hence that $\mu' = \mu_{|_{R^\perp}} \perp \mu'_{|_R}$. Because $\mu'_{|_R} \cdot \rho'_{|_R} = \sigma_{|_R} = \rho_{|_R}$, it follows that $\rho' = 1_{R^\perp} \perp (\mu'_{|_R})^{-1}(\rho_{|_R})$. Therefore the obstruction to the uniqueness of the Zassenhaus splitting is as described earlier.

Notice that $U' \cap R' = U' \cap R = 0$ if and only if $\mu'_{|_R} = 1_R$. In this case, $\mu' = \mu$ and $\rho' = \rho$. If R is anisotropic, then $O_r(R)$ has no non-trivial unipotent elements, and this condition is met. The following uniqueness criterion is a special case of our discussion.

Proposition 2. (Uniqueness) Let $\sigma = \mu\rho$ be the Zassenhaus splitting of $\sigma \in O_n(V)$. Suppose that $\sigma = \mu'\rho'$ where μ' is unipotent, ρ' non-degenerate, and $S = U' \perp R'$. Then

$$\mu' = \mu \ \text{ and } \ \rho' = \rho \ .$$

Proposition 3. (Conjugacy) Let σ and σ_1 be elements in $O_n(V)$ and let $\sigma = \mu\rho$ and $\sigma_1 = \mu_1\rho_1$ be their Zassenhaus splittings. Then σ_1 is conjugate to σ if and only if μ_1 is conjugate to μ and ρ_1 is conjugate to ρ.

Proof: If σ_1 is conjugate to σ then by an application of Proposition 2, μ_1 is conjugate to μ and ρ_1 is conjugate to ρ. As to the converse, observe first that $S_\sigma = U_\mu \perp R_\rho = R_\rho \perp U_\mu$ and similarly for S_{σ_1}. If μ_1 is conjugate to μ and ρ_1 is conjugate to ρ, then U_{μ_1} is isometric to U_μ and R_{ρ_1} is isometric to R_ρ. Therefore S_{σ_1} is isometric to S_σ, and hence σ_1 is conjugate to σ. QED.

3. Application to the Length Problem. Our study of the length problem for the group $\Omega_n(V)$ and its set of generators

$$A = \{\tau_v\tau_w\tau_v\tau_w | \tau_v \text{ and } \tau_w \text{ non-commuting hyperplane reflections in } O_n(V)\}$$

will expand on the results of Hahn [6].

Let $\sigma \in \Omega_n(V)$ be arbitrary. Note that the typical element $\tau_v\tau_w\tau_v\tau_w$ in A is equal to $\tau_v\tau_{\tau_w(v)} = \tau_v\tau_{v'}$ where $Fv \neq Fv'$ and $Q(v) = Q(v')$. Conversely, any such product is an element in A. It is a direct consequence of this fact and Theorem 1, that

$$\ell(\sigma) \geq \frac{1}{2} \dim S.$$

We will therefore define $\sigma \in \Omega_n(V)$ to be *short* if $\ell(\sigma) = \frac{1}{2}\dim S$ and *long* if $\ell(\sigma) > \frac{1}{2}\dim S$.

Our goal is the same as that of Theorem 1, namely the complete description of the long elements of $\Omega_n(V)$ and the determination of their lengths.

Let σ in $\Omega_n(V)$ be an involution. By an application of the Wall form, σ is short if and only if $S = W_1 \perp \cdots \perp W_k$ with $\dim W_i = 2$ and disc $W_i = 1$. Totally degenerate elements are in $\Omega_n(V)$. It follows from Theorem 1 that they are long. We now focus on the elements in $\Omega_n(V)$ that are neither involutions nor totally degenerate.

Theorem 2. Let $\sigma \in \Omega_n(V)$ be long with σ neither totally degenerate nor an involution. Let $\sigma = \mu\rho$ be the Zassenhaus splitting of σ. Then

i) The space of μ satisfies $U = \operatorname{Rad} U \perp T$ with T anisotropic. The element μ is a product of $\frac{1}{2}(\dim U)$ commuting Eichler transformations, exactly $\dim T$ of which are not totally degenerate. In particular, $(\mu - 1_V)^3 = 0$.

ii) The element ρ is long and its space R is anisotropic.

iii) (Splicing Condition) The space $T \perp R$ is anisotropic.

Finally, if V is isotropic, then $\ell(\sigma) = \frac{1}{2}\dim S + 1$.

Proof: In view of Hahn [6] and in particular Proposition 15, only the existence of the factorization in (i) requires proof. By the same proposition, we know that $(\mu - 1_V)U \subseteq \operatorname{Rad} U$. If $T = 0$, then μ is totally degenerate. By Hahn-O'Meara [5], μ is a product of $\frac{1}{2}(\dim U)$ totally degenerate commuting Eichler transformations. So we may assume that $T \neq 0$. Let $w_1 \in T$ be non-zero. If $\mu w_1 = w_1$, then w_1 is in the fixed space $U^\perp$ of μ. But this implies that $w_1 \in U \cap U^\perp = \operatorname{Rad} U$, a contradiction. So $\mu w_1 - w_1$ is a non-zero vector in $\operatorname{Rad} U$. Put $\mu w_1 = u_1 + w_1$ with $u_1 \in \operatorname{Rad} U$. Note that $\mu(Fu_1 \perp Fw_1) = Fu_1 \perp Fw_1$ and (because μ is unipotent) that the restriction of μ to this plane has determinant 1. Let $\alpha_1 = B(w_1, w_1)^{-1}$ and consider the Eichler transformation $\Sigma_{u_1,\alpha_1 w_1}$. Check that $\Sigma_{u_1,\alpha_1 w_1}(u_1) = u_1$ and $\Sigma_{u_1,\alpha_1 w_1}(w_1) =$

$w_1 + u_1$. Observe that $\mu\Sigma^{-1}_{u_1,\alpha_1 w_1}\big|_{(Fu_1 \perp Fw_1)} = 1_{(Fu_1 \perp Fw_1)}$. By 8.2.16 of [5], μ commutes with $\Sigma^{-1}_{u_1,\alpha_1 w_1}$. Put $\mu_1 = \Sigma^{-1}_{u_1,\alpha_1 w_1}\mu$. By general facts, $U_1 \subseteq (Fu_1 \perp Fw_1) + U \subseteq U$. Because μ_1 fixes w_1 while μ does not, the fixed space $U_1^{\perp}$ of μ_1 strictly contains the fixed space $U^{\perp}$ of μ. It follows that $\dim U_1 = \dim U - 2$. Because $\mu = \Sigma_{u_1,\alpha_1 w_1} \cdot \mu_1$, $U \subseteq (Fu_1 \perp Fw_1) + U_1$. By dimensions, $U = (Fu_1 \perp Fw_1) \oplus U_1$. Since $\Sigma_{u_1,\alpha_1 w_1}$ commutes with μ, it commmutes with μ_1. Therefore, $U = (Fu_1 \perp Fw_1) \perp U_1$. Note that $\text{Rad } U = Fu_1 \perp \text{Rad } U_1$. Put $U_1 = \text{Rad } U_1 \perp T_1$. Because

$$U = (Fu_1 \perp \text{Rad } U_1) \perp (Fw_1 \perp T_1)$$

is a radical splitting of U we know that $Fw_1 \perp T_1$ is isometric to T and hence that T_1 is anisotropic. The element μ_1 is unipotent because it is a product of two commuting unipotent elements. An induction completes the proof. QED.

Remark: By dimension considerations, the spaces of the Eichler transformations in Theorem 2 (i) are planes with trivial intersection. Because these Eichler transformations commute, these planes are orthogonal. Observe also that $\frac{1}{2}\dim U \geq \dim T$, and hence that $\dim \text{Rad } U \geq \dim T$.

The next two results will show that the limitations that Theorem 2 imposes on the components μ and ρ of the Zassenhaus splitting of a long element σ are considerable.

Let $p(X) = a_k X^k + \cdots + a_1 X + a_0$ be a polynomial in $F[X]$. We call $p(X)$ *symmetric* if the two sequences of coefficients $a_k, \ldots, a_0$ and $a_0, \ldots, a_k$ are identical.

Theorem 3. Let $\sigma \in \Omega_n(V)$ be long. Then the prime decomposition of the minimal polynomial of σ has the form

$$(X-1)^m \, p_1(X) \, \cdots \, p_j(X)$$

where $0 \leq m \leq 3$ and the $p_i(X)$ are distinct, monic, symmetric, and irreducible.

Proof: If σ is an involution or totally degenerate, this is clear. So assume that Theorem 2 applies to σ. Consider $\rho_{|_R}$. Because R is anisotropic, any non-zero subspace W of R is non-degenerate. It follows that $R = W_1 \perp \cdots \perp W_j$ where each W_i is invariant under ρ, but $\rho_{|_{W_i}}$ has no non-trivial invariant subspaces. By applying the results of Huppert, e.g., Satz 2.4 of [8] and Satz 4.1 of [9], (also see the references to Cikunov in Milnor [14]), we see that the minimal polynomial of $\rho_{|_{W_i}}$ is symmetric and irreducible. QED.

Proposition 4. Let i be the Witt index of V. Let $\sigma \in \Omega_n(V)$ be a unipotent element with minimal polynomial $(X-1)^m$.

i) If V is hyperbolic, then $m \leq 2i - 1$.

ii) If V is not hyperbolic, then $m \leq 2i + 1$.

In either case, there exist unipotent elements such that equality holds.

Proof: The inequalities follow by induction on i. The case $i = 0$ is the anisotropic case, where we already know that 1_V is the only unipotent element. So assume that $i \geq 1$. Now refer to case (2) of the proof of Theorem 2.4 of [4] and in particular to the unipotent element $\tau = \sigma_Y \in \Omega_{n-2}(X)$. Let k be the degree of the minimal polynomial of τ. Notice that $V = Z \perp X$ with Z a hyperbolic plane. So the Witt index of X is $i - 1$. Applying the induction hypothesis to τ, provides the inequality $k \leq 2(i-1) - 1 = 2i - 3$ if X is hyperbolic and $k \leq 2(i-1) + 1 = 2i - 1$ if not. It follows from the way σ and τ are related that $m \leq k + 2$. This completes the proof of the inequalities.
The construction of the required elements also follows inductively. If V is a hyperbolic plane, then 1_V is the only unipotent element and it satisfies the equality trivially. Note next that a hyperbolic space of Witt index i contains a non-degenerate space of Witt index $i - 1$ that is not hyperbolic. This implies that it suffices to carry out the construction of the required unipotent element in case (ii). So suppose that V is not hyperbolic with Witt index i. To get the induction off the ground, take $i = 1$. Let $\sigma = \Sigma_{u,v}$ be a non-degenerate Eichler transformation. Then $m = 3 = 2i + 1$ as required. It is also easy to check that $(\sigma - 1_V)^2 V = Fu$. Because V is spanned by isotropic vectors, there is an isotropic vector w in V such that $(\sigma - 1_V)^2 w = u$. Because $i = 1$, it follows that $B(\sigma(\sigma - 1_V)^2 w, w) = B(u, w) \neq 0$.
Suppose that $i \geq 2$ and let $V = H \perp W$ with H a hyperbolic plane. Note that W is not hyperbolic and that it has Witt index $i - 1$. For the induction hypothesis, assume that τ is a unipotent element in $\Omega_{n-2}(W)$ and that the minimal polynomial of τ is $(X - 1)^k$ with $k = 2(i-1) + 1 = 2i - 1$. Assume further that $(\tau - 1_W)^{k-1} W$ is a line spanned by $(\tau - 1_W)^{k-1} w$ with w isotropic and $B(\tau(\tau - 1)^{k-1} w, w) \neq 0$. Put $H = Fu \oplus Fv$ with u and v isotropic and $B(u, v) = 1$. To complete the proof, we will show that $\sigma = \Sigma_{u,w} \cdot (1_H \perp \tau)$ is a unipotent element in $\Omega_n(V)$ that satisfies all the properties of τ with $k + 2$ in place of k. ¿From the defining equation of $\Sigma_{u,w}$ we see that $\sigma u - u = 0, \sigma v - v = w$, and that

$$\sigma x - x = \tau x - x + B(\tau x, w) u \quad \text{for all } x \in W\,.$$

This formula and an induction shows that

$$(\sigma - 1_V)^j x = (\tau - 1_W)^j x + B(\tau(\tau - 1_W)^{j-1} x, w) u \quad \text{for all} \quad x \in W \text{ and } j \geq 1\,.$$

We claim that σ has minimal polynomial $(X - 1)^{k+2}$, that $(\sigma - 1_V)^{k+1} V$ is spanned by $(\sigma - 1_V)^{k+1} v$, and that $B(\sigma(\sigma - 1_V)^{k+1} v, v) \neq 0$. To see this, observe first that $(\sigma - 1_V)^{k+1} x = 0$ for all $x \in W$. Because $(\sigma - 1_V) u = 0$, it follows that $(\sigma - 1_V)^{k+1}V$ is spanned by $(\sigma - 1_V)^{k+1}v$. Recall that $(\sigma - 1_V)v = w$. Therefore,

$$\begin{aligned}(\sigma - 1_V)^{k+1}v &= (\sigma - 1_V)^k w \\ &= (\tau - 1_W)^k w + B(\tau(\tau - 1_W)^{k-1}w, w)u \\ &= B(\tau(\tau - 1_W)^{k-1}w, w)u \neq 0 \ .\end{aligned}$$

Because $\sigma u = u$, we see that $(\sigma - 1_V)^{k+2}v = 0$ and hence that σ has minimal polynomial $(X - 1)^{k+2}$. Finally, $B(\sigma(\sigma - 1_V)^{k+1}v, v) = B(B(\tau(\tau - 1_W)^{k-1}w, w)u, v) = B(\tau(\tau - 1_W)^{k-1}w, w)B(u, v) \neq 0$. The proof is complete. QED.

The elements of Theorem 2 can be constructed as follows: Start with a long anisotropic ρ in $\Omega_n(V)$. Choose a subspace $U = \text{Rad}\ U \perp T$ in $R^\perp$ such that $T \perp R$ is anisotropic and $\dim \text{Rad}\ U \geq \dim T$. Split U into an orthogonal sum of degenerate planes. For each plane choose an Eichler transformation that has the plane as its space. Let μ be the product of these Eichler transformations and set $\sigma = \mu\rho$. This - by its uniqueness property - is the Zassenhaus decomposition of σ. Therefore, the description of the long elements of $\Omega_n(V)$ has been reduced to the following two problems:

A. Classify all long anisotropic elements ρ in $\Omega_n(V)$ and compute their lengths in the case of an anisotropic V.

B. Determine which of the elements in Theorem 2 are actually long.

Notice that the conjugates of a long element in $\Omega_n(V)$ are long elements of $\Omega_n(V)$. If the long element is anisotropic, then the conjugates are also anisotropic. Thus, the problem of classifying long elements calls for the classification of their conjugacy classes.

4. Local Fields.

The study of the arithmetic theory of quadratic forms flows classically via the progression

$$\mathbb{C},\ \mathbb{R},\ \text{finite fields},\ \text{local fields},\ \text{and global fields}$$

from the easy situations to the hard ones. The theory over local fields makes use of that over finite fields (via the residue class field) and the theory over global fields - in characteristic zero these are the finite extensions of $\mathbb{Q}$ - is based via local/global principles on the theory over local fields and $\mathbb{C}$ and $\mathbb{R}$.

The benefit of hindsight, namely that the length problem that is being considered depends on the arithmetic of the field, suggests that its analysis should proceed along the same path. Theorem 4 below is a routine application of

Theorem 1. See Hahn [6] for the details. Observe that it applies at once to $\mathbb{C}$, $\mathbb{R}$, and finite fields.

Theorem 4. Suppose that $\operatorname{card} \overset{*}{F}/\overset{*}{F}^2 \leq 2$. Then the totally degenerate elements σ are the only long elements in $\Omega_n(V)$, and for these $\ell(\sigma) = \frac{1}{2} \dim\ S + 1$.

Let's turn next to the case of a local field. $\mathbb{R}^+$ be the set of positive real numbers. A *local field* is a field F that has a valuation

$$|\ \ | : F \longrightarrow \mathbb{R}^+ \cup \{0\}$$

which satisfies the strong triangle inequality and with respect to which $|\overset{*}{F}|$ is discrete and F is complete. Let

$$\mathfrak{o} = \{\alpha \in F \mid |\alpha| \leq 1\}$$

be the *valuation ring* of F and $\mathfrak{p} = \{\alpha \in F \mid |\alpha| < 1\}$ its unique maximal ideal. As part of the definition of local field, the *residue class field* $\mathfrak{o}/\mathfrak{p}$ is assumed to be finite. We continue the assumption that char $F \neq 2$. Denote by $\mathfrak{u} = \{\varepsilon \in \mathfrak{o} \mid |\varepsilon| = 1\}$ the group of invertible elements of $\mathfrak{o}$. Because the maximal ideal $\mathfrak{p}$ is principal, $\mathfrak{p} = \mathfrak{o}\pi$ for some $\pi \in \mathfrak{o}$. Any such π is a *prime* element in $\mathfrak{o}$. Note that $|\pi|$ is the largest value such that $|\pi| < 1$.

We refer to O'Meara [15] for the notation and the basic properties of local fields, their quadratic forms and orthogonal groups. Two important facts about quadratic forms over local fields are these: any non-degenerate quadratic space of dimension five or more is isotropic, and there is, up to isometry, a unique anisotropic four dimensional quadratic space.

It will be necessary to distinguish non-dyadic local fields from dyadic local fields. The local field F is *non-dyadic* if 2 is invertible in $\mathfrak{o}$ and *dyadic* if not. So F is non-dyadic if $|2| = 1$ and dyadic if $|2| < 1$. If V is the unique 4-dimensional anisotropic quadratic space, then $\Omega_4(V)$ has index two in $O'_4(V)$ if F is non-dyadic, and $\Omega_4(V) = O'_4(V)$ if F is dyadic.

Consider a long element $\sigma \in \Omega_n(V)$ that is neither totally degenerate nor an involution and return to the properties of the Zassenhaus decomposition $\sigma = \mu\rho$ provided by Theorem 2. The fact that ρ is long implies that $\dim R \geq 4$. Therefore,

$$4 \leq \dim\ R \leq \dim\,(T \perp R) \leq 4\,.$$

So $T = 0$, and μ is totally degenerate. In reference to Theorem 3, it follows that the bound on m is $0 \leq m \leq 2$. Also, $\dim R = 4$ and R is the unique 4-dimensional anisotropic space over F. What else can be said about ρ?

Proposition 5. Let ρ in $\Omega_n(V)$ be anisotropic with dim $R = 4$. Then $\rho_{|_R} \in O'_4(R)$, and

i) If F is non-dyadic, then ρ is long if and only if $n \geq 5$ and $\rho_{|_R} \in O'_4(R) - \Omega_4(R)$.

ii) If F is dyadic, then ρ is long if and only if $\rho_{|_R} \in O'_4(R) = \Omega_4(R)$ is long.

This initial answer to Question A is proved in [6]. Suppose that F is non-dyadic. Then Proposition 5 together with Theorem 2 tell us that $\ell(\sigma) = \frac{1}{2}\text{dim } S+1$ for any long element σ. Proposition 5 also provides a complete answer to Question B. See Theorem 3 of [6]. It asserts that the long elements in $\Omega_n(V)$ that are not involutions and not totally degenerate are those of Theorem 2, namely they are precisely the elements with Zassenhaus decomposition $\sigma = \mu\rho$, where $\mu = 1_V$ or μ is totally degenerate and ρ satisfies (i) above. If F is dyadic, then Question B is as yet not resolved. However, it is known that $\ell(\sigma) = \frac{1}{2}\text{dim } S + 1$ for all long elements σ.

Let V be the anisotropic 4-dimensional space over F. In view of Proposition 5 we will analyze the elements σ in $O'_4(V)$ that satisfy

(A) σ in $O'_4(V) - \Omega_4(V)$ if F is non-dyadic, and

(B) σ a long element in $O'_4(V) = \Omega_4(V)$ if F is dyadic.

Is there a criterion that pinpoints when an element in $O'_4(V)$ satisfies (A) or (B)? A theorem of Milnor [14] tells us where to look.

Theorem 5. Let V be an n-dimensional, non-degenerate quadratic space over a local field F. Let $m(X)$ be a monic, irreducible polynomial in $F[X]$ and let deg $m(X) = k$. Assume that $m(X)$ is neither $X - 1$ nor $X + 1$.

i) $m(X)$ is the minimal polynomial of an element of $O_n(V)$ if and only if k is even and divides n, $m(X)$ is symmetric, and disc $V = (m(1)m(-1))^{\frac{n}{k}}\overset{*}{F}{}^2$.

ii) Given such a polynomial $m(X)$ there is precisely one conjugacy class of elements in $O_n(V)$ with minimal polynomial $m(X)$.

Milnor's result no longer holds when $m(X)$ is reducible. Any Eichler transformation that is not totally degenerate has minimal polynomial $(X - 1)^3$ and provides an example showing that (i) no longer holds. The nontrivial totally degenerate elements - all of which have minimal polynomial $(X - 1)^2$ - show that (ii) fails. Let μ and μ_1 be totally degenerate elements. Then μ is conjugate to μ_1 if and only if their respective spaces U and U_1 have the same dimension. This follows from the conjugacy criterion given by the Wall form and the fact U_μ and U_{μ_1} are both alternating.

Let $\sigma \in O_4'(V)$ and let $m(X)$ be its minimal polynomial. Milnor's theorem suggests that it should be possible to look at $m(X)$ and decide whether σ satisfies condition (A) or (B) or not. In reference to Theorem 3, we are interested in precise information about the product $p_1(X)\cdots p_k(X)$. In the discussion that follows, only the arithmetic aspects of the proofs will be provided.

Proposition 6. Let V be anisotropic with dim $V = 4$ and let $\sigma \in O_4'(V)$. Then σ satisfies (A) or (B) if and only if σ satisfies the *long criterion*:

$$Q(\sigma x - x) = -\beta_x^2 Q(x) \text{ for all } x \in V \text{ and some } \beta_x \in \overset{*}{F}.$$

Suppose that $m(X)$ has a factor of the form $X - a$. So $\sigma x = ax$ for some non-zero x in V. Because x is anisotropic, $a = \pm 1$. Since $\sigma x = x$ violates the long criterion, we must have $\sigma x = -x$. So $a = -1$. Therefore, $X + 1$ is the only possible monic linear factor of $m(X)$. If $(X + 1)^2$ is a factor of $m(X)$, then $-\sigma$ is a non-trivial unipotent element on some subspace of V. But this is impossible, because V is anisotropic.

1. Suppose deg $m(X) = 1$. This implies that $m(X) = X + 1$ and hence that $\sigma = -1_V$. Because disc $V = 1$, $-1_V \in O_4'(V)$. Check that $\sigma = -1_V$ satisfies the long criterion precisely when $-1 \in \overset{*}{F}{}^2$.

2. Suppose deg $m(X) = 2$. Observe that $m(X)$ must be irreducible. By Theorem 5, $m(X) = X^2 - cX + 1$ for some $c \in F$. Notice that $c \neq \pm 2$. Every line of V contains a plane that is invariant under σ. Let W be any such plane. By Theorem 5, disc $W = -(c-2)(c+2)\overset{*}{F}{}^2$. Because V is anisotropic, W is not a hyperbolic plane, and therefore, $(c-2)(c+2) \notin \overset{*}{F}{}^2$. Again by Theorem 5, there is precisely one conjugacy class of such elements σ for a given c. A spinor norm computation shows that $\Theta(\sigma) = (c-2)^2\overset{*}{F}{}^2 = \overset{*}{F}{}^2$. So any σ with minimal polynomial of this form is in $O_4'(V)$. It turns out that σ satisfies the long criterion if and only if $c - 2 \in \overset{*}{F}{}^2$.

3. Suppose deg $m(X) = 3$. By Theorem 5, $m(X)$ is reducible. It follows that $m(X) = (X + 1)(X^2 - cX + 1)$ with $X^2 - cX + 1$ irreducible. Again, $c \neq \pm 2$. Let $p_1(X) = X + 1$ and $p_2(X) = X^2 - cX + 1$. Put $U = p_2(\sigma)V$ and $W = p_1(\sigma)V$. Observe that U and W are planes that are invariant under σ, that $V = U \perp W$, that $\sigma_{|U} = -1_U$, and that $\sigma_{|W}$ has minimal polynomial $X^2 - cX + 1$. As in the previous case, disc $W = -(c-2)(c+2)\overset{*}{F}{}^2$ and $(c-2)(c+2) \notin \overset{*}{F}{}^2$. By Theorem 63:20 of [15], there are two isometry classes of anisotropic planes of a given

discriminant. An application of Theorem 5 implies that there are two conjugacy classes of σ for a given c. By a spinor norm computation, $\Theta(\sigma) = -(c-2)\overset{*}{F}{}^2$. So $\sigma \in O_4'(V)$ if and only $-(c-2) \in \overset{*}{F}{}^2$ and $-(c+2) \notin \overset{*}{F}{}^2$. The analysis of the long criterion for $\sigma \in O_4'(V)$ will follow shortly. We will see that if it holds, then $-1 \in \overset{*}{F}{}^2$ and $c-2 \in 4\mathfrak{u}^2$. This implies in turn that $c \in 2\mathfrak{u}$. If F is non-dyadic, the converse is true. Namely, $-1 \in \overset{*}{F}{}^2$ and $c-2 \in \mathfrak{u}^2$ together imply the long criterion.

4. Suppose $\deg m(X) = 4$. In this case, either

   $m(X) = (X^2 - cX + 1)(X^2 - dX + 1)$ with distinct irreducible factors,

   or

   $m(X) = X^4 - cX^3 - dX^2 - cX + 1$ is irreducible.

   The first case is very similar to case (3). The second seems complicated and is as yet not completely understood.

We now return to case (3) and to the analysis of the long criterion. Let $\dot{U}$ and $\dot{W}$ denote the non-zero elements of U and W and let

$$C = Q(\dot{W})/Q(\dot{U})\,.$$

The set C is closed under multiplication by squares and hence under taking inverses.

Assume that the long criterion holds. Applying it to U and W we get that

$$\text{(i)} \qquad -1 \ \text{ and } \ c-2 \ \text{ are both in } \ \overset{*}{F}{}^2.$$

Put $-1 = i^2$ and $c-2 = s^2$ and let $t = -2is^{-1}$. Applying the long criterion to the vectors $x = u + w$ with $u \in U$ and $w \in W$, tells us that

$$\text{(ii)} \qquad \frac{1+\gamma t^2}{1+\gamma} \in \overset{*}{F}{}^2 \quad \text{for all } \ \gamma \in C\,.$$

Conversely, the long criterion is equivalent to the combination of (i) and (ii).

We assume that (i) and (ii) hold and consider the consequences for the constant c. We show first that $t \in \mathfrak{u}$. It follows from the discussion in paragraph 63.C of O'Meara [15] that C contains a prime element π. Therefore C contains π^i for any odd i either positive or negative. Put $t = \delta\pi^k$ with $\delta \in \mathfrak{u}$. Taking $\gamma = \pi$ we get,

$$\frac{1+\gamma t^2}{1+\gamma} = \frac{1+\delta^2\pi^{2k+1}}{1+\pi}\,.$$

If $k < 0$, then $2k+1 < 0$, and $\frac{|1+\delta^2\pi^{2k+1}|}{|1+\pi|} = |\pi|^{2k+1}$ by the Principle of Domination. Because $2k+1$ is odd, the element above cannot be a square. This contradicts (ii). If $k > 0$, a similar contradiction is obtained by taking $\gamma = \pi^{-1}$. Therefore, $k = 0$ and $t \in \mathfrak{u}$ as required. Let $\varepsilon = -it^{-1} \in \mathfrak{u}$. Because $s = 2\varepsilon$, we get

$$c - 2 = 4\varepsilon^2.$$

Therefore, $c - 2 \in 4\mathfrak{u}^2$ as asserted earlier. Note that $c = 2(2\varepsilon^2+1)$. If F is dyadic, then by domination, $c \in 2\mathfrak{u}$. This is also true in the non-dyadic case. If $c \notin \mathfrak{u}$, then $c \in \mathfrak{p}$. But this would imply by Hensel's Lemma that $X^2 - cX + 1$ is reducible.

We now explore the converse. Assume both $-1 \in \overset{*}{F}{}^2$ and $c - 2 \in 4\mathfrak{u}^2$. Does $\sigma \in O_4'(V)$ with such a c satisfy the long criterion, or equivalently, conditions (i) and (ii)? Condition (i) holds trivially, so the focus is on (ii). Put $c-2 = 4\varepsilon^2$ with $\varepsilon \in \mathfrak{u}$. Set $s = 2\varepsilon$ and $t = -2is^{-1} = -i\varepsilon^{-1}$. Notice that $t \in \mathfrak{u}$. Because C is closed under taking inverses, (ii) is equivalent to $1 + \frac{t^2-1}{1+\gamma} \in \overset{*}{F}{}^2$ for all γ in C. Check that $t^2 - 1 = -\frac{4+s^2}{s^2} = -\frac{c+2}{c-2} = -\frac{c+2}{4\varepsilon^2}$. So the question is this: Is it the case that

$$\text{(iii)} \qquad 1 \; - \; \frac{c+2}{4\varepsilon^2(1+\gamma)} \; \in \; \overset{*}{F}{}^2$$

for all $\gamma \in C$?

The first step toward the answer is the observation that $\{|1+\gamma| \mid \gamma \in C\}$ is bounded below by $|4|$. For suppose that $|1+\gamma| \leq |4\pi|$ for some $\gamma \in C$. Then $1+\gamma = 4\alpha\pi$ for some $\alpha \in \mathfrak{o}$. But this means that $-\gamma = 1 - 4\alpha\pi \in \overset{*}{F}{}^2$ by the Local Square Theorem. Because C is closed under multiplication by squares, $-\gamma \in C$. But this implies that the intersection $Q(\dot{U}) \cap Q(\dot{W})$ is not empty. This would mean that V contains a plane of discriminant $\overset{*}{F}{}^2 = -\overset{*}{F}{}^2$, i.e., a hyperbolic plane. This is not possible because V is anisotropic. Now assume that $|c+2| < |4|^3$. Given the bound just established, $|\frac{c+2}{4\varepsilon^2(1+\gamma)}| < |4|$ for all $\gamma \in C$. Therefore by another application of the Local Square Theorem,

$$1 \; - \; \frac{c+2}{4\varepsilon^2(1+\gamma)} \; \in \; \overset{*}{F}{}^2$$

for all γ in C.

We conclude the discussion of the converse by assuming that F is non-dyadic. In this case (iii) is satisfied for any c (such that $c-2 \in 4\mathfrak{u}^2$). Because $|4| = 1$, we already know that (iii) holds when $c+2 \in \mathfrak{p}$. Since $c+2 \in \mathfrak{o}$, only the case $c+2 \in \mathfrak{u}$ remains. Instead of (iii), we will verify the equivalent condition (ii). Recall from the beginning of the analysis of case (3) that disc $W = -(c-2)(c+2)\overset{*}{F}^2$ and $-(c+2) \notin \overset{*}{F}^2$. So disc $W = -(c+2)\overset{*}{F}^2$ and, by an application of Example 63:15 of [15], $C = \pi\mathfrak{u}\overset{*}{F}^2$. Let $\gamma = \pi\delta\alpha^2 \in C$ with $\delta \in \mathfrak{u}$ and $\alpha \in \overset{*}{F}$ be arbitrary. If $|\alpha| \leq 1$, then $|\gamma| < 1$. So $1+\gamma$ and $1+t^2\gamma$ are both in $\overset{*}{F}^2$ by the Local Square Theorem. Therefore (ii) holds. Suppose $|\alpha| > 1$. Now $|\gamma| > 1$ and the Local Square Theorem tells us that $1+\gamma^{-1}$ and $1+t^{-2}\gamma^{-1}$ are both squares. So $\frac{1+t^{-2}\gamma^{-1}}{1+\gamma^{-1}}$ is a square. Therefore $\frac{1+\gamma t^2}{1+\gamma}$ is a square as well and (ii) holds in this case also. The proof of the converse in the non-dyadic case is complete. The dyadic situation is much more delicate and is not completely settled.

5. Global Fields.

Let F be a global field, let V be a non-degenerate quadratic space over F, and consider the group $\Omega_n(V)$. Not much is known about the length question in this situation, but it is clear that local-global considerations are relevant. Let $\mathfrak{p}$ be a prime - Archimedean or not - and consider the completion $V_\mathfrak{p}$. The first indication is the theorem that tells us that $\sigma \in \Omega_n(V)$ if and only if $\sigma_\mathfrak{p} \in \Omega_n(V_\mathfrak{p})$ for all $\mathfrak{p}$. Another is the fact (analogous to what was observed in the local case) that the analysis of the anisotropic long elements in $\Omega_n(V)$ reduces to the 4-dimensional anisotropic long elements. This is true not only in the situation where F is a function field or a totally complex number field (in these situations there are no anisotropic spaces of dimension 5 or more) but in general. More precisely, if σ is an anisotropic long element, then

$$\sigma = \omega_1 \cdots \omega_k \sigma_1,$$

where all ω_i are elementary commutators of hyperplane reflections, the space $S = W_1 \oplus \cdots \oplus W_k \oplus S_1$, and σ_1 is long with dim $S_1 = 4$.

Bibliography

1. E. Artin, *Geometric Algebra,* Wiley Interscience, New York, 1966.
2. E. W. Ellers and J. Malzan, Products of reflections in the kernel of the spinorial norm, *Geom. Ded.* 36 (1990), 279-285.
3. M. Dyer, On minimal length of expressions of Coxeter group elements as products of reflections, *Proceedings of the AMS*, to appear.
4. A. J. Hahn, Unipotent elements and the spinor norms of Wall and Zassenhaus, *Archiv Math.* (Basel) 32 (1979), 114-122.
5. A. J. Hahn and O. T. O'Meara, *The Classical Groups and K-Theory,* Grundlehren der Mathematik, Springer-Verlag, 1989.

6. A. J. Hahn, The elements of the orthogonal group $\Omega_n(V)$ as products of commutators of symmetries, *J. Algebra* 184 (1996), 927-944.

7. J. E. Humphreys, *Reflection Groups and Coxeter Groups,* Cambridge University Press, 1990.

8. B. Huppert, Isometrien von Vektorräumen I, *Archiv Math.* (Basel) 35 (1980), 164-176.

9. B. Huppert, Isometrien von Vektorräumen II, *Math. Z.* 175 (1980), 5-20.

10. F. Knüppel, Products of simple isometries of given conjugacy types, *Forum Math.* 5 (1993), 441-458.

11. T. Y. Lam, *The Algebraic Theory of Quadratic Forms,* Benjamin, Reading MA, 1973.

12. G. A. Margulis, Explicit constructions of graphs without short cycles and low density codes, *Combinatorica* 2 (1) 1982, 71-78.

13. G. A. Margulis, Explicit group-theoretical constructions of combinatorial schemes and their application to the design of expanders and concentrators, *Communication Network Theory* 1988, 39-46.

14. J. Milnor, On Isometries of Inner Product Spaces, *Inventiones Math.* 8 (1969), 83-97.

15. O. T. O'Meara, *Introduction to Quadratic Forms,* Classics in Mathematics, Springer-Verlag, 2000, a reprint of the 1973 Springer Grundlehren edition.

16. J. Rosenthal and P. O. Vontobel, Construction of LDPC codes using Ramanujan graphs and ideas from Margulis, *Proceedings 38th Allerton Conference on Communication, Control and Computing, October 4-6, 2000*, to appear.

17. H. Zassenhaus, On the spinor norm, *Archiv Math.* (Basel) 13 (1962), 434-451.

Alexander Hahn
Department of Mathematics
University of Notre Dame
Notre Dame, IN 46556
USA
hahn.1@nd.edu

Dimensions of Anisotropic Indefinite Quadratic Forms, I

Detlev W. Hoffmann

Received: May 29, 2001
Revised: November 10, 2001

Communicated by Ulf Rehmann

Abstract. By a theorem of Elman and Lam, fields over which quadratic forms are classified by the classical invariants dimension, signed discriminant, Clifford invariant and signatures are exactly those fields F for which the third power I^3F of the fundamental ideal IF in the Witt ring WF is torsion free. We study the possible values of the u-invariant (resp. the Hasse number $\tilde{u}$) of such fields, i.e. the supremum of the dimensions of anisotropic torsion (resp. anisotropic totally indefinite) forms, and we relate these invariants to the symbol length λ, i.e. the smallest integer n such that the class of each product of quaternion algebras in the Brauer group of the field can be represented by the class of a product of $\leq n$ quaternion algebras. The nonreal case has been treated before by B. Kahn. Here, we treat the real case which turns out to be considerably more involved.

1991 Mathematics Subject Classification: 11E04, 11E10, 11E81, 12D15
Keywords and Phrases: quadratic form, indefinite quadratic form, torsion quadratic form, real field, u-invariant, Hasse number, symbol length

1. Introduction

Let F be a field of characteristic $\neq 2$. An important topic in the algebraic theory of quadratic forms over F is the determination of the supremum of the dimensions of certain types of anisotropic quadratic forms over F. For a general survey on this problem, see [H4]. In the present article, we focus on the u-invariant and the Hasse number $\tilde{u}$ of F, where $u(F)$ (resp. $\tilde{u}(F)$) is defined as the supremum of the dimensions of anisotropic forms which are torsion in the Witt ring of F (resp. totally indefinite, i.e. indefinite with respect to each ordering on F). By Pfister's local-global principle, torsion forms are exactly those forms which have signature 0 with respect to each ordering, they are in particular totally indefinite (or t.i. for short). Hence, $u(F) \leq \tilde{u}(F)$. In the absence of orderings, i.e. for nonreal fields, every form is a torsion form and the two definitions coincide with what was originally called the u-invariant, namely the supremum of the dimensions of anisotropic forms over F.

We will relate these two invariants to another one, the so-called *symbol length* λ, which is defined to be the smallest n (if such an n exists) such that any tensor product of quaternion algebras over F is Brauer-equivalent to a tensor product of $\leq n$ quaternion algebras. $\lambda(F) \leq 1$ is equivalent to saying that the classes of quaternion algebras form a subgroup of the Brauer group $\mathrm{Br}(F)$. In this case, the field is called *linked*. It should be remarked that by Merkurjev's theorem [M1], the classes of products of quaternion algebras are exactly the elements in $\mathrm{Br}_2(F)$, i.e. the elements of exponent ≤ 2 in the Brauer group $\mathrm{Br}(F)$.

Perhaps the first result relating the u-invariant and the Hasse number to the symbol length is due to Elman and Lam [EL2], [E] who determined the values of u and $\tilde{u}$ for linked fields. Their result reads as follows.

Theorem 1.1. *Let F be a linked field. Then $u(F) = \tilde{u}(F) \in \{0, 1, 2, 4, 8\}$. In particular, $I_t^4 F = 0$. Furthermore, for $1 \leq n \leq 3$, $u(F) = \tilde{u}(F) \leq 2^{n-1}$ iff $I_t^n F = 0$.*

In the wake of Merkurjev's construction of fields with $u = 2n$ for any positive integer n ([M2]) which is based on his index reduction results and its consequences (see Lemma 2.2(iii)) and on a simple fact concerning *Albert forms* (see Lemma 2.2(i)), it has been noted by Kahn that for nonreal fields, a lower bound for u can easily be given in terms of λ. More precisely, Kahn [Ka, Th. 2] shows the following.

Theorem 1.2. *Let F be a nonreal field. Then*

(i) *$\lambda(F) = 0$ iff $u(F) \leq 2$.*
(ii) *If $\lambda(F) \geq 1$ then $u(F) \geq 2\lambda(F) + 2$.*
(iii) *If $\lambda(F) \geq 1$ and $I^3 F = 0$, then $u(F) = 2\lambda(F) + 2$.*

(In Kahn's original statement, it was implicitly assumed that $\lambda(F) \geq 1$, and only parts (ii) and (iii) were stated.)

The aim of the present paper is to generalize this result to real fields, in particular to real fields with $I_t^3 F = 0$. Since the quaternion algebra $(-1, -1)_F$ will always be a division algebra over any given real field F, we will always have $\lambda(F) \geq 1$. By Elman and Lam's theorem 1.1 we know for real F that $\lambda(F) = 1$ implies $u(F) = \tilde{u}(F) \in \{0, 2, 4, 8\}$ and that in this case $u(F) = \tilde{u}(F) \in \{0, 2, 4\}$ iff $I_t^3 F = 0$. Thus, we are mainly interested in the case F real and $\lambda(F) \geq 2$.

Now fields with $I_t^3 F = 0$ are also interesting from a different point of view as by another theorem of Elman and Lam [EL3] these are exactly the fields over which quadratic forms can be classified by the classical invariants dimension, signed discriminant, Clifford invariant, and signatures.

Our first main result is the analogue for real fields of Kahn's theorem above, but now in terms of the Hasse number.

Theorem 1.3. *Let F be a real field with $\lambda = \lambda(F) \geq 2$. Then the following holds.*

(i) $\tilde{u}(F) \geq 2\lambda + 2$.

(ii) *If $I_t^3 F = 0$ and $\tilde{u}(F) < \infty$, then $\tilde{u}(F) = 2\lambda + 2$.*

The situation for the u-invariant seems to be more complicated. We could prove an analogue of Kahn's theorem only under invoking rather restrictive additional hypotheses on the space of orderings X_F of the field. Recall that the reduced stability index $st(F)$ of a real F can be defined as follows : $st(F) = 0$ if F is uniquely ordered; otherwise, $st(F)$ is the smallest integer $s \geq 0$ such that for each basic clopen set $H(a_1, \cdots, a_n) \subset X_F$ there exist $b_i \in F^*$, $1 \leq i \leq s$, such that $H(a_1, \cdots, a_n) = H(b_1, \cdots, b_s)$. $st(F) \leq 1$ is equivalent to F being SAP (cf. [KS, Kap. 3, § 7, Satz 3]).

Theorem 1.4. *Let F be a real field with $\lambda = \lambda(F) \geq 2$.*

(i) *If $st(F) \leq 1$ then $u(F) \geq 2\lambda$.*

(ii) *If $I_t^3 F = 0$ and $st(F) \leq 2$, then $u(F) \leq 2\lambda + 2$.*

These results will be shown in the next section.

In [M2], Merkurjev constructed to each $n \geq 1$ fields with $u(F) = 2n$ and $I^3 F = 0$. It has been shown by Hornix [Hor, Th. 3.5] and Lam [L2] that for each $n \geq 3$ there exist real fields F, F' such that $u(F) = \tilde{u}(F) = 2n$ and $u(F') + 2 = \tilde{u}(F') = 2n$. Note that in [L2], it was in addition shown that there exist such fields which are uniquely ordered, but nothing was said about $I_t^3 F$, whereas in [Hor] it was shown that one can construct such fields with $I_t^3 = 0$, but there were no statements made on the space of orderings of such fields.

For the reader's convenience, we will give a proof of these results by Hornix resp. Lam in section 3. Our constructions are slightly different from those given by Hornix and Lam but, just as theirs, rely heavily on Merkurjev's index reduction results as stated in Lemma 2.2. In our constructions, we will also combine the properties of F having $I_t^3 F = 0$ and of F being uniquely ordered in the case $\tilde{u} < \infty$.

In fact, we will put these results into a larger context where we classify all realizable values for the invariants λ, u and $\tilde{u}$ (and their interdependences) for real fields with $I_t^3 F = 0$ which are SAP. Since the values of u and $\tilde{u}$ for fields (real or not) with $\lambda \leq 1$ are covered by Elman and Lam's theorem 1.1 (note that these fields are always SAP since for them $\tilde{u}$ will be finite, [EP, Theorem 2.5]), and since the case of nonreal fields is treated in Kahn's Theorem 1.2, we will only consider the case of real SAP fields with $I_t^3 F = 0$ and $\lambda(F) \geq 2$.

Theorem 1.5. *Let $\mathcal{M} = \{(n, 2n, 2n+2),\ (n, 2n+2, 2n+2);\ n \geq 2\} \cup \{(n, 2n, \infty),\ (n, 2n+2, \infty); n \geq 2\} \cup \{(\infty, \infty, \infty)\}$.*

(i) *Let F be a real SAP field such that $\lambda(F) \geq 2$ and $I_t^3 F = 0$. Then $(\lambda(F), u(F), \tilde{u}(F)) \in \mathcal{M}$.*

(ii) *Let $(\lambda, u, \tilde{u}) \in \mathcal{M}$. Then there exists a real SAP field F with $I_t^3 F = 0$ and $(\lambda(F), u(F), \tilde{u}(F)) = (\lambda, u, \tilde{u})$. In the case where $\tilde{u} < \infty$ or $\lambda = \infty$, there exist such fields which are uniquely ordered.*

As a consequence, we obtain

Corollary 1.6. *Let F be a real field with $I_t^3F = 0$. Then*

$$(u(F), \tilde{u}(F)) \in \{(2n, 2n);\ n \geq 0\} \cup \{(2n, \infty);\ n \geq 0\} \cup \{(2n, 2n+2);\ n \geq 2\}\ .$$

All pairs of values on the right hand side can be realized as pairs $(u(F), \tilde{u}(F))$ for suitable real F.

As far as notation, terminology and basic results from the algebraic theory of quadratic forms is concerned, we refer to the books by Lam [L1] and Scharlau [S]. In particular, $\varphi \cong \psi$ (resp. $\varphi \sim \psi$) denotes isometry (resp. equivalence in the Witt ring) of the forms φ and ψ. $\sum F^2$ denotes all nonzero sums of squares in F. The signed discriminant (resp. Clifford invariant) of a form φ will be denoted by $d_{\pm}\varphi$ (resp. $c(\varphi)$), and we write φ_{an} for the anisotropic part of φ. An n-fold Pfister form is a form of type $\langle 1, -a_1\rangle \otimes \cdots \otimes \langle 1, -a_n\rangle$, $a_i \in F^*$, and we write $\langle\langle a_1, \cdots, a_n\rangle\rangle$ for short. The set of forms isometric (resp. similar) to n-fold Pfister forms will be denoted by P_nF (resp. GP_nF). I_t^nF is the torsion part of I^nF, the n-th power of the fundamental ideal IF of classes of even-dimensional forms in the Witt ring WF of the field F. The space of orderings of a real field F will be denoted by X_F. General references for the SAP property and the reduced stability index are the book by Knebusch and Scheiderer [KS], and the articles [P], [ELP], [EP]. Another property in this context is the so-called ED property (*effective diagonalization*). It is known that ED implies SAP (but not conversely in general), and that fields with finite $\tilde{u}$ have the ED property. Cf. [PW] for more details on ED.

2. Fields with torsion-free I^3

Definition 2.1. (i) Let A be a central simple algebra over F (CSA/F) such that its Brauer class $[A]$ is in $\mathrm{Br}_2(F)$. The *symbol length* $t(A)$ of A is defined as

$$t(A) = \min\{n \mid \exists \text{ quaternion algebras } Q_i/F,\ 1 \leq i \leq n, \text{ s.t. } [A] = [\textstyle\bigotimes_{i=1}^n Q_i]\}\ .$$

(ii) The *symbol length* $\lambda(F)$ of the field F is defined as

$$\lambda(F) = \sup\{t(A) \mid A \text{ CSA}/F,\ [A] \in \mathrm{Br}_2(F)\}\ .$$

(iii) Let φ be a form over F. Let A be a CSA/F such that $c(\varphi) = [A] \in \mathrm{Br}_2(F)$, where $c(\varphi)$ denotes the Clifford invariant of φ. Then $t(\varphi) := t(A)$.

The following lemma compiles some well known results and some special cases of Merkurjev's index reduction theorem which we will use in this and the following section. We refer to [M2], [T] for details (see also [L1, Sect. 3, Ch. V] for basic results on Clifford invariants and how to compute them).

Lemma 2.2. (i) *Let $Q_i = (a_i, b_i)$, $1 \leq i \leq n$, be quaternion algebras over F with associated norm forms $\langle\langle a_i, b_i\rangle\rangle \in P_2F$. Let $A = \bigotimes_{i=1}^n Q_i$ (over F). Then there exist $r_i \in F^*$, $1 \leq i \leq n$, and a form $q \in I^2F$, $\dim q = 2n+2$ such that $c(q) = [A] \in \mathrm{Br}_2 F$ and $q \sim \sum_{i=1}^n r_i\langle\langle a_i, b_i\rangle\rangle$ in WF. (We will* *call such a form q an* Albert form *associated with A.) Furthermore, if $t(A) = n$ (in particular if A is a division algebra), then every Albert form associated with A is anisotropic.*

(ii) *If q is a form over F with either $\dim q = 2n+2$ and $q \in I^2F$, or $\dim q = 2n+1$, or $\dim q = 2n$ and $d_{\pm} q \neq 1$, then there exist quaternion algebras $Q_i = (a_i, b_i)$, $1 \leq i \leq n$, such that for $A = \bigotimes_{i=1}^n Q_i$ we have $c(q) = [A]$, and there exists an Albert form φ associated with A such that $q \subset \varphi$.*

(iii) *If A as in (i) is a division algebra and if ψ is a form over F of one of the following types:*
  (a) $\dim \psi \geq 2n+3$,
  (b) $\dim \psi = 2n+2$ *and* $d_{\pm}\psi \neq 1$,
  (c) $\dim \psi = 2n+2$, $d_{\pm}\psi = 1$ *and* $c(\psi) \neq [A] \in \mathrm{Br}_2 F$,
  (d) $\psi \in I^3F$,
  then A stays a division algebra over $F(\psi)$.

The next result will be used in the proofs of Theorem 1.4(ii) and of Lemma 2.4(ii), which in turn will be used in the proof of Theorem 1.3(ii).

Lemma 2.3. *Let $n \geq 1$ and suppose that $I_t^{n+1}F = 0$. Let φ be a form over F of dimension $> 2^n$. Suppose that either*

- *$\varphi \in I_t^n F$, or*
- *φ is t.i. and F is ED.*

If there exists $\rho \in GP_nF$ such that $\rho \subset \varphi$, then φ is isotropic.

Proof. Write $\varphi \cong \rho \perp \psi$. By assumption, $\dim \psi \geq 1$. After scaling, we may assume that $\rho \in P_nF$. Note that $\mathrm{sgn}_P\, \rho \in \{0, 2^n\}$ for all $P \in X_F$. Let $Y = \{P \in X_F \mid \mathrm{sgn}_P(\rho) = 2^n\}$.

If ρ is torsion, i.e. if Y is empty, then for any x represented by ψ we have that $\rho \otimes \langle\langle -x\rangle\rangle \in P_{n+1}F \cap W_tF \subset I_t^{n+1}F = 0$. Thus, the Pfister neighbor $\rho \perp \langle x\rangle$ is isotropic. Hence, φ is isotropic as it contains $\rho \perp \langle x\rangle$ as subform.

So assume that $Y \neq \emptyset$. First, suppose that $\varphi \in I_t^nF$. Then we have $\mathrm{sgn}_P\, \psi = -2^n$ for all $P \in Y$ and hence $\dim \psi \geq 2^n$. Now $\langle 1,1\rangle \otimes \varphi \in I_t^{n+1}F = 0$, hence $\langle 1,1\rangle \otimes \rho \sim -\langle 1,1\rangle \otimes \psi$ in WF. By β-decomposition (cf. [EL1, p. 289]), we can write $\psi \cong \gamma \perp \sigma$ with $\langle 1,1\rangle \otimes \gamma \sim 0$ (in particular, $\gamma \in W_tF$), $\dim \sigma = \dim \rho = 2^n$ and $\langle 1,1\rangle \otimes \rho \cong -\langle 1,1\rangle \otimes \sigma$. Comparing signatures, we see that $\mathrm{sgn}_P\, \rho = -\mathrm{sgn}_P\, \sigma \in \{0, 2^n\}$. Now let $x \in F^*$ be any element represented by σ. The above shows that $x <_P 0$ for all $P \in Y$. For all other $P \in X_F$, ρ is indefinite. This yields that $\rho \perp \langle x\rangle$ is t.i. and a Pfister neighbor of $\rho \otimes \langle\langle -x\rangle\rangle$ which is therefore torsion. We conclude as before that φ is isotropic.

Finally, suppose that φ is t.i. and that F is ED. Since ρ is positive definite at all orderings $P \in Y$, and since $\varphi \cong \rho \perp \psi$ is t.i., ED implies that ψ represents an $x \in F^*$ such that $x <_P 0$ for all $P \in Y$. Then $\rho \perp \langle x\rangle$ is t.i. and a Pfister neighbor contained in φ, and we conclude as before that φ is isotropic. □

For later purposes, we now state some useful facts on u and $\tilde{u}$ of real fields with $I_t^3F = 0$.

Lemma 2.4. *Let F be a field with $I_t^3F=0$. Then the following holds.*

(i) *If $2<u(F)<\infty$, then there exists an anisotropic form $\varphi\in I_t^2F$ such that* $\dim\varphi=u(F)$.

(ii) *If $\tilde{u}(F)<\infty$, then $\tilde{u}(F)$ is even. Furthermore, if $2<\tilde{u}(F)<\infty$, then there exists an anisotropic t.i. form $\varphi\in I^2F$ such that* $\dim\varphi=\tilde{u}(F)$ *and* $\mathrm{sgn}_P(\varphi)\in\{0,4\}$ *for all $P\in X_F$.*

Proof. (i) See [EL1, Prop. 1.4].

(ii) See [ELP, Th. H] for a proof that $\tilde{u}(F)$ is even if it is finite. Now suppose φ is anisotropic, t.i. and $\dim\varphi=\tilde{u}(F)\geq 4$. Since $\tilde{u}(F)$ is finite, F has ED and one easily sees that φ contains a 3-dimensional t.i. subform τ'. Then τ' is a Pfister neighbor of some anisotropic torsion $\tau\in P_2F$. Thus, if $\tilde{u}(F)=4$, this τ is the desired form. So we may assume that $\tilde{u}(F)\geq 6$.

Since F is SAP, we may scale φ so that $\mathrm{sgn}_P\,\varphi\geq 0$ for all $P\in X_F$. Consider the clopen set $Y=\{P\in X_F\mid \mathrm{sgn}_P\,\varphi\geq 5\}$. Since F is SAP, there exists a 3-fold Pfister form π such that $\mathrm{sgn}_P\,\pi=8$ for all $P\in Y$ and $\mathrm{sgn}_P\,\pi=0$ otherwise. Consider $\varphi_1=x(\varphi\perp -\pi)_{\mathrm{an}}$, where $x\in F^*$ is chosen so that $\mathrm{sgn}_P\,\varphi_1\geq 0$ for all $P\in X_F$. By construction, $0\leq \mathrm{sgn}_P\,\varphi_1\leq\max\{4,|\,\mathrm{sgn}_P\,\varphi-8|\}<\dim\varphi$. If $\dim\varphi_1>\dim\varphi$, then φ_1 would be an anisotropic t.i. form of dimension $\geq \tilde{u}(F)+2$, clearly a contradiction. If $\dim\varphi_1<\dim\varphi$, then φ and π would contain a common 5-dimensional subform which, being a Pfister neighbor, would in turn contain a subform $\rho\in GP_2F$. Since F is ED as $\tilde{u}(F)<\infty$, Lemma 2.3 then implies that φ is isotropic, a contradiction. It follows that $\dim\varphi_1=\dim\varphi$. By repeating this construction, we get a sequence of anisotropic t.i. forms $\varphi_0=\varphi,\varphi_1,\cdots,\varphi_r$ such that for $i\geq 1$ we have $\dim\varphi_i=\dim\varphi$, $0\leq\mathrm{sgn}_P\,\varphi_i\leq\max\{4,|\,\mathrm{sgn}_P\,\varphi_{i-1}-8|\}$ and $0\leq\mathrm{sgn}_P\,\varphi_r\leq 4$ for all $P\in X_F$.

Hence, we may assume that φ is anisotropic t.i., $\dim\varphi=\tilde{u}(F)$ and $0\leq\mathrm{sgn}_P\,\varphi\leq 4$ for all $P\in X_F$. Let $d=d_\pm\varphi$ and consider $\psi=(\varphi\perp\langle 1,-d\rangle)_{\mathrm{an}}$. Note that $\psi\in I^2F$ and therefore $\mathrm{sgn}_P\,\psi\equiv 0 \bmod 4$. Since $0\leq\mathrm{sgn}_P\,\varphi\leq 4$ and $\mathrm{sgn}_P\,\langle 1,-d\rangle\in\{0,\pm 2\}$ for all $P\in X_F$, it follows readily that $\mathrm{sgn}_P\,\psi\in\{0,4\}$. We also have that $\dim\varphi-2\leq\dim\psi\leq\dim\varphi+2$.

If $\dim\psi=\dim\varphi+2$, then $\psi\cong\varphi\perp\langle 1,-d\rangle$ would be an anisotropic t.i. form of dimension $\tilde{u}(F)+2$, clearly a contradiction.

If $\dim\psi=\dim\varphi-2$, then $\varphi\cong\psi\perp\langle d,-1\rangle$. Since $\mathrm{sgn}_P\,\psi\geq 0$ for all $P\in X_F$ and because of ED, we have that ψ represents some $a\in\sum F^2$. Then $\psi\perp -a\psi$ is a torsion form in I^3F and thus hyperbolic. But $\psi\perp -a\psi$ contains the subform $\psi\perp\langle -1\rangle$ which by dimension count must be isotropic. Hence φ is isotropic, a contradiction. Thus $\dim\psi=\dim\varphi=\tilde{u}(F)$ and ψ is the desired form. □

Remark 2.5. (i) If $u(F)=\infty$, then there exist anisotropic torsion forms in I^2F of arbitrarily large dimension. Indeed, let $\varphi\in W_tF$ be anisotropic of dimension $\geq 2n+2$. Let $d=d_\pm\varphi$ and consider $\psi=(\varphi\perp\langle 1,-d\rangle)_{\mathrm{an}}$. Then one readily checks that $\dim\psi\geq 2n$ and $\psi\in I_t^2F$.

(ii) If $\tilde{u}(F) = \infty$, then there exist anisotropic t.i. forms in I^2F of arbitrarily large dimension. Indeed, let φ be any anisotropic t.i. form of dimension $4n+3$ for any $n \geq 1$ (such φ exists by [ELP, Th. A]). Let d be such that $\varphi \perp \langle d \rangle \in I^2F$. Let $\psi = (\varphi \perp \langle d \rangle)_{\mathrm{an}}$. Then $\psi \in I^2F$ and $\dim \psi \in \{4n+2, 4n+4\}$. If $\dim \psi = 4n+4$ then $\psi \cong \varphi \perp \langle d \rangle$ is t.i.. If $\dim \psi = 4n+2$, then $\mathrm{sgn}_P\, \psi \equiv 0 \bmod 4$ for all $P \in X_F$ as $\psi \in I^2F$, and therefore $|\mathrm{sgn}_P\, \psi| \leq 4n < 4n+2 = \dim \psi$ for all $P \in X_F$. Again, ψ is t.i..

Let us now turn to the proof of part (ii) of Theorem 1.4 where we assume that $I_t^3F = 0$ and $st(F) \leq 2$. In [KS, Kap. 3, § 7, Korollar], one finds different characterizations of F having reduced stability index $\leq s$ for an integer $s \geq 1$. The one we are interested in is the following : $st(F) \leq s$ is equivalent to $(I^{s+1}F)_{red} = 2(I^sF)_{red}$, i.e. for each form $\varphi \in I^{s+1}F$ there exists a form $\psi \in I^sF$ such that $\mathrm{sgn}_P\, \varphi = \mathrm{sgn}_P(\langle 1,1 \rangle \otimes \psi)$ for all $P \in X_F$. If $I_t^{s+1}F = 0$, then $st(F) \leq s$ is therefore equivalent to $I^{s+1}F = 2I^sF$. By [Kr, Prop. 1], we thus get

Lemma 2.6. *Let $s \geq 1$ be an integer and let F be a real field with $I_t^{s+1}F = 0$. Then the following are equivalent :*

(i) *$st(F) \leq s$;*
(ii) *$I^{s+1}F = 2I^sF$;*
(iii) *$I^{s+1}F(\sqrt{-1}) = 0$.*

Now $I^{s+1}F(\sqrt{-1}) = 0$ implies $I_t^{s+1}F = 0$, [Kr, Prop. 1], and in view of this lemma, we may replace the hypotheses $I_t^3F = 0$ plus $st(F) \leq 2$ by $I^3F(\sqrt{-1}) = 0$. We then get the following result which holds for *any* field (not just for real fields) and which implies the second part of Theorem 1.4.

Theorem 2.7. *Suppose that $I^3F(\sqrt{-1}) = 0$. Then*

$$u(F) \leq \min\{4\lambda(F(\sqrt{-1})) + 2, 2\lambda(F) + 2\}\ .$$

Proof. First, we prove that $u(F) \leq 2\lambda(F) + 2$. If the level $s(F)$ of F is finite, i.e. F is nonreal, then this follows from Kahn's theorem 1.2.

So assume that F is a real field with $I^3F(\sqrt{-1}) = 0$. We will show that if $\varphi \in I_t^2F$ with $t(\varphi) = t$, then $\dim \varphi > 2t + 2$ implies that φ is isotropic. This then implies readily $u(F) \leq 2\lambda(F) + 2$. Indeed, this follows from the fact that there always exists an anisotropic form in I_t^2F of dimension $u(F)$ if $u(F)$ is finite (Lemma 2.4(i)), resp. of arbitrarily large dimension if $u(F)$ is infinite (Remark 2.5(i)), and the fact that in the case of a real F with $I_t^3F = 0$, $st(F) \leq 2$ is equivalent to $I^3F(\sqrt{-1}) = 0$ by Lemma 2.6.

Now let $\varphi \in I_t^2F$ with $t(\varphi) = t$ and $\dim \varphi > 2t + 2$. We will prove by induction on t that φ is isotropic. If $t = 0$ then $\varphi \in I_t^3F = 0$ and φ is in fact hyperbolic. If $t = 1$ then there exists (an anisotropic) $\tau \in P_2F$ such that $c(\varphi) = c(\tau)$. By Merkurjev's theorem, $\varphi \equiv \tau \bmod I^3F$. Since $\mathrm{sgn}_P\, \tau \in \{0, 4\}$ and $0 = \mathrm{sgn}_P\, \varphi \equiv \mathrm{sgn}_P\, \tau \bmod 8$ for all $P \in X_F$, we see that $\tau \in W_tF$, hence $\varphi \equiv \tau \bmod I_t^3F$ and thus $\varphi \sim \tau \in WF$ as $I_t^3F = 0$. Hence $\dim \varphi > \dim \varphi_{\mathrm{an}} = \dim \tau = 4$ and φ is isotropic.

Let $t \geq 2$. By Lemma 2.2(i), there exists an anisotropic $(2t+2)$-dimensional form $\tau \in I^2F$ such that $\varphi \equiv \tau \bmod I^3F$. Let $d \in F^*$ such that $\tau_{F(\sqrt{d})}$ is isotropic. Let $\tau' \in I^2F(\sqrt{d})$ such that $(\tau_{F(\sqrt{d})})_{\mathrm{an}} \cong \tau'$. Then $\dim \tau' \leq 2t$. Hence, $t(\tau') \leq t-1$. (In fact, one can readily show that $\dim \tau' = 2t$ and $t(\tau') = t-1$, but we won't need this here.) Also, $\varphi_{F(\sqrt{d})} \equiv \tau' \bmod I^3F(\sqrt{d})$. By [EL3, Th. 3] and [EL4, Cor. 4.6], we have that $I_t^3F(\sqrt{d}) = 0$ and $I^3F(\sqrt{d})(\sqrt{-1}) = 0$. By induction hypothesis, we have $\dim(\varphi_{F(\sqrt{d})})_{\mathrm{an}} \leq 2(t-1)+2 = 2t$. Now $\dim \varphi \geq 2t+4$, hence there exist $a, b \in F^*$ such that $\langle a, b\rangle \otimes \langle 1, -d\rangle \subset \varphi$ (cf. [L1, Ch. VII, Lemma 3.1]). Now $\langle a, b\rangle \otimes \langle 1, -d\rangle \in GP_2F$, and by Lemma 2.3, φ is isotropic.

Let us now show that $u(F) \leq 4\lambda(F(\sqrt{-1}))+2$. This is trivially true for $s(F) = 1$ as in this case we have $F = F(\sqrt{-1})$ and already $u(F) \leq 2\lambda(F)+2$.

So suppose that $s(F) \geq 2$. We put $L = F(\sqrt{-1})$ and we may assume that $\lambda = \lambda(L) < \infty$. Since $I_t^3F = 0$, we have $\langle 1,1\rangle I_t^2F = 0$. Hence, $\mathrm{ann}(\langle 1,1\rangle) \cap I^2F = \mathrm{ann}(\langle 1,1\rangle) \cap I_t^2F = I_t^2F$. Consider the Scharlau transfer $s_* : WL \to WF$ induced by the F-linear map $L \to F$ defined by $1 \mapsto 0$ and $\sqrt{-1} \mapsto 1$. Note that for any form ρ over L there exists a form σ over F such that $\dim \sigma \leq 2\dim \rho$ and $s_*(\rho) \sim \sigma$ in WF.

By [AEJ, Prop. 1.24], we have $s_*(I^2L) = \mathrm{ann}(\langle 1,1\rangle) \cap I^2F$ and thus $s_*(I^2L) = I_t^2F$. Now let ψ be any form in I^2L. By Lemma 2.2(i), there exists a form $\eta \in I^2L$ such that $\dim \eta \leq 2\lambda+2$ and $c(\psi) = c(\eta) \in \mathrm{Br}_2 L$. After scaling, we may assume that $\eta \cong \langle 1\rangle \perp \eta'$. In particular, there exists a form $\gamma \in I^3L$ such that $\eta \sim \psi + \gamma$ in WL. Now $s_*(\gamma) \in I_t^3F = 0$. Hence $s_*(\psi) = s_*(\eta) = s_*(\langle 1\rangle) + s_*(\eta') \sim \sigma$ for some form σ over F with $\dim \sigma \leq 2\dim \eta' \leq 4\lambda+2$.

Now let $\varphi \in I_t^2F$. Since $s_*(I^2L) = I_t^2F$, the above shows that $\varphi \sim \mu$ in WF for some form μ over F with $\dim \mu \leq 4\lambda+2$. Hence, if φ is anisotropic we necessarily have $\dim \varphi \leq 4\lambda+2$.

Suppose $u(F) = \infty$. Then there exists some anisotropic form $\tau \in W_tF$ with $\dim \tau \geq 4\lambda+6$ and $\dim \tau$ even. Let $d = d_{\pm}\tau$. Then one easily sees that $\tau \perp \langle 1, -d\rangle \in I_t^2F$, and its anisotropic part must therefore be of dimension $\leq 4\lambda+2$, a contradiction to τ being anisotropic and $\dim \tau \geq 4\lambda+6$.

Hence $u(F) < \infty$. Then Lemma 2.4(i) and the above imply that $u(F) \leq 4\lambda+2$. □

Remark 2.8. Let F be such that $s(F) \geq 2$ and let $L = F(\sqrt{-1})$. Define $u'(F) = \sup\{\dim \varphi \mid \varphi \text{ anisotropic form}/F \text{ and } \langle 1,1\rangle \otimes \varphi = 0 \in WF\}$. It was shown in [Pf, Ch. 8, Th. 2,12] that $u'(F) \leq 2u(L)-2$. Now if $I_t^3F = 0$, then one readily verifies that $u(F) = u'(F)$ (see also [Pf, Ch. 8, Prop. 2.6]). Hence, this would imply that $u(F) \leq 2u(L)-2$. Note, however, that I^3L need not be zero and that therefore $u(L) > 2\lambda(L)+2$ might very well be possible (cf. Theorem 1.2), in which case our bound $u(F) \leq 4\lambda(L)+2$ would be better.

Corollary 2.9. (See also [Pf, Ch. 8, Th. 2,12], [EL1, Th. 4.11].) *Let F be a field with $s(F) \geq 2$ and let $L = F(\sqrt{-1})$. Let $n \in \{1,2,3\}$. Then $u(L) \leq 2n$ implies $u(F) \leq 4n-2$. Furthermore, if $u(L) = 1$ then F is real and pythagorean (i.e. $u(F) = 0$).*

Proof. If $u(L) \leq 2n$, $1 \leq n \leq 3$, then $I^3L = 0$ and thus $I_t^3F = 0$ (cf. [Kr, Prop. 1]). Theorem 1.2 yields $\lambda(L) \leq n-1$. Hence $u(F) \leq 4\lambda(L)+2 \leq 4n-2$. The second part is left to the reader. □

To prove Theorems 1.3 and 1.4(i), we will need the following lemma.

Lemma 2.10. *Let $n \geq 1$ and suppose that F is SAP.*

(i) *Let $\pi_i \in P_nF$, $1 \leq i \leq r$. Then there exists a form $\varphi \in I^nF$ such that $\operatorname{sgn}_P \varphi \in \{0, 2^n\}$ for all $P \in X_F$, and $\varphi \equiv \sum_{i=1}^r \pi_i \bmod I^{n+1}F$.*

(ii) *If $I_t^{n+1}F = 0$, and if $\varphi \in I^nF$ such that $\operatorname{sgn}_P \varphi \in \{0, 2^n\}$ for all $P \in X_F$, then $\varphi \cong \varphi_t \perp \varphi_0$ with $\varphi_t \in W_tF$ and $\dim \varphi_0 \in \{0, 2^n\}$.*

(iii) *If $I_t^{n+1}F = 0$, then the form φ in part (i) can be chosen so as to have dimension $\leq r2^n - 2r + 2$.*

Proof. (i) We use induction on r. If $r = 1$ then $\varphi = \pi_1$ will do. So suppose $r \geq 2$. By induction hypothesis, there exists a form ψ such that $\psi \equiv \sum_{i=1}^{r-1} \pi_i \bmod I^{n+1}F$ and $\operatorname{sgn}_P \psi \in \{0, 2^n\}$ for all $P \in X_F$. Let $\hat{\varphi} = \psi \perp -\pi_r$. Since $\operatorname{sgn}_P \pi_r \in \{0, 2^n\}$, we have $\operatorname{sgn}_P \hat{\varphi} \in \{0, \pm 2^n\}$. Since F is SAP, there exists an $x \in F^*$ such that $\varphi = x\hat{\varphi}$ has $\operatorname{sgn}_P \varphi \in \{0, 2^n\}$ for all $P \in X_F$. Clearly, $\varphi \equiv \sum_{i=1}^r \pi_i \bmod I^{n+1}F$.

(ii) Suppose now that $I_t^{n+1}F = 0$. Consider the clopen set $Y = \{P \in X_F \mid \operatorname{sgn}_P \varphi = 2^n\}$ in X_F. If Y is empty then $\varphi \in W_t$ and there is nothing to show. So suppose $Y \neq \emptyset$. Let $\sigma \in P_nF$ be such that $\operatorname{sgn}_P \sigma = 2^n$ if $P \in Y$, and $\operatorname{sgn}_P \sigma = 0$ otherwise. Such σ exists as F is SAP. It follows that $\langle 1,1\rangle \otimes \varphi \equiv \langle 1,1\rangle \otimes \sigma \bmod I_t^{n+1}F$ (both forms are in $I_t^{n+1}F$ and have the same signatures). Now $I_t^{n+1}F = 0$ and thus $\langle 1,1\rangle \otimes \varphi \sim \langle 1,1\rangle \otimes \sigma$. (Note that $\langle 1,1\rangle \otimes \sigma$ is anisotropic because $\operatorname{sgn}_P \langle 1,1\rangle \otimes \sigma = \dim \langle 1,1\rangle \otimes \sigma = 2^{n+1}$ for all $P \in Y \neq \emptyset$.) Comparing dimensions and using β-decomposition (cf. [EL1, p. 289]), we see that $\varphi \cong \varphi_t \perp \varphi_0$ with $\varphi_t \in W_tF$ and $\dim \varphi_0 = \dim \sigma = 2^n$.

(iii) We use a similar induction argument as in (i), but we assume in addition that the form ψ there is of dimension $\leq (r-1)2^n - 2(r-1) + 2$. By (ii), we can write $\psi \cong \psi_t \perp \psi_0$ with $\dim \psi_0 \in \{0, 2^n\}$, $\psi_t \in W_tF$, and $\dim \psi_0 = 2^n$ only if there exists some $P \in X_F$ with $\operatorname{sgn}_P \psi = 2^n$. Let $y \in D(\psi_0)$ if $\dim \psi_0 = 2^n$, and let $y \in D(\psi)$ otherwise. One readily checks that $\operatorname{sgn}_P y\psi = \operatorname{sgn}_P \psi \in \{0, 2^n\}$ and that $y\psi \cong \langle 1\rangle \perp \psi'$. Let now $\pi_r \cong \langle 1\rangle \perp \pi_r'$ and let $\varphi' = \psi' \perp -\pi_r'$. Note that $\dim \varphi' \leq r2^n - 2r + 2$. As in the proof of (i), $\operatorname{sgn}_P \varphi' \in \{0, \pm 2^n\}$, and after scaling, we obtain the form φ with $\operatorname{sgn}_P \varphi \in \{0, 2^n\}$ for all $P \in X_F$, $\dim \varphi = \dim \varphi' \leq r2^n - 2r + 2$, and $\varphi \equiv \sum_{i=1}^r \pi_i \bmod I^{n+1}F$. □

Proof of Theorem 1.3. (i) If F is not SAP, then $\tilde{u}(F) = \infty$ and there is nothing to show. So suppose that F is SAP. Let $A = Q_1 \otimes \cdots \otimes Q_t \in Br_2F$, where the Q_i are quaternion algebras such that $t(A) = t \geq 2$, and consider the norm forms

$\pi_i \in P_2F$ associated with Q_i. By Lemma 2.10(i), there exists an anisotropic form $\varphi \in I^2F$ such that $\varphi \equiv \sum_{i=1}^t \pi_i \bmod I^3F$ and $\mathrm{sgn}_P\,\varphi \in \{0,4\}$. Note that $c(\varphi) = [A]$. If $\dim\varphi \leq 2t$ then, by Lemma 2.2(ii), $c(\varphi)$ could be represented by a product of fewer than t quaternion algebras, a contradiction to $t(A) = t$. Hence $\dim\varphi \geq 2t+2$. Note that φ is t.i. provided $t \geq 2$.

If $\lambda(F) = \infty$, then for any $t \geq 1$ there exists an $A \in Br_2F$ with $t(A) = t$, and the above shows that $\tilde{u}(F) = \infty$. If $\lambda(F) < \infty$, then choose A as above such that $t(A) = \lambda(F)$. The above shows that $\tilde{u}(F) \geq 2\lambda(F) + 2$.

(ii) By Lemma 2.4(ii), we may assume that there exists an anisotropic t.i. form $\varphi \in I^2F$ with $\dim\varphi = \tilde{u}(F)$ and $\mathrm{sgn}_P\,\varphi \in \{0,4\}$ for all $P \in X_F$. Let $t(\varphi) = t \leq \lambda$ and let $c(\varphi) = Q_1 \otimes \cdots \otimes Q_t \in Br_2F$. With π_i the norm forms associated with Q_i, we get $\varphi \equiv \sum_{i=1}^t \pi_i \bmod I^3F$.

By Lemma 2.10(iii), there exists a form $\psi \in I^2F$, $\dim\psi \leq 2t+2$ such that $\mathrm{sgn}_P\,\psi \in \{0,4\}$ for all $P \in X_F$ and such that $\varphi \equiv \psi \bmod I^3F$. Since $\mathrm{sgn}_P\,\varphi \equiv \mathrm{sgn}_P\,\psi \bmod 8$, this readily yields $\varphi \perp -\psi \in I_t^3F = 0$. The anisotropy of φ then shows that $\tilde{u}(F) = \dim\varphi \leq 2t+2 \leq 2\lambda(F)+2$, which together with (i) yields $\tilde{u}(F) = 2\lambda(F) + 2$. □

Proof of Theorem 1.4(i). Let $A = Q_1 \otimes \cdots \otimes Q_t \in Br_2F$, where the Q_i are quaternion algebras such that $t(A) = t \geq 2$. As in part (i) of the proof of Theorem 1.3, there exists an anisotropic form $\varphi \in I^2F$ such that $c(\varphi) = [A]$, $\mathrm{sgn}_P\,\varphi \in \{0,4\}$, $\dim\varphi \geq 2t+2$.

Now let $\pi \in P_2F$ be such that $\mathrm{sgn}_P\,\varphi = \mathrm{sgn}_P\,\pi$ for all $P \in X_F$. (Such π exists as F is SAP and $\mathrm{sgn}_P\,\varphi \in \{0,4\}$.) Consider $\psi = (\varphi \perp -\pi)_{\mathrm{an}}$. By construction, $\psi \in I_t^2F$ and $\dim\psi \geq \dim\varphi - 4 = 2t-2$. Suppose that $\dim\psi = \dim\varphi - 4$. Then $\varphi \cong \psi \perp \pi$ and we have $\psi, \pi \in I^2F$, $c(\varphi) = c(\psi)c(\pi)$. By dimension count and Lemma 2.2(ii), we have $t(\psi) \leq t-2$, $t(\pi) \leq 1$, and therefore $t(\varphi) = t(A) = t \leq t(\psi) + t(\pi) \leq t-1$, a contradiction. Hence, $\dim\psi \geq \dim\varphi - 2 = 2t$.

If $\lambda(F) = \infty$, then for any $t \geq 1$ there exists an $A \in Br_2F$ with $t(A) = t$, and the above shows that $u(F) = \infty$.

If $\lambda(F) < \infty$, then choose A as above such that $t(A) = \lambda(F)$. The above then shows that $u(F) \geq 2\lambda(F)$. □

Since fields with finite $\tilde{u}$ are always SAP, the following is an immediate consequence of Theorems 1.3, 1.4.

Corollary 2.11. *Let F be a real field with $I_t^3F = 0$ and $\tilde{u}(F) < \infty$. Then $\tilde{u}(F) = 2\lambda(F) + 2 \in \{u(F), u(F)+2\}$.*

Example 2.12. The condition in Theorem 1.4(i) that F be SAP seems to be quite restrictive. However, we will certainly need some sort of additional assumption on F besides $I_t^3F = 0$ to get the lower bound $u(F) \geq 2\lambda(F)$. To see what can go wrong when one drops the assumption that F is SAP, consider the following example. Let $F = \mathbb{R}((t_1))\cdots((t_n))$ be the iterated power series field in n variables over the reals. Then, by Springer's theorem, $u(F) = 0$. In particlar, $I_t^3F = 0$. For $n \geq 2$, F is not SAP as for example $\langle 1, t_1, t_2, -t_1t_2\rangle$ is not weakly isotropic. However, one can show that $\lambda(F) = [n/2] + 1$. The value $t(A) = [n/2] + 1$ can be realized, for example, by the multiquaternion division algebra $A = (-1, -1) \otimes (t_1, t_2) \otimes \cdots (t_{m-1}, t_m)$ where $m = [n/2]$ (i.e. $n \in \{2m, 2m + 1\}$).

As for the upper bound for $u(F)$ for a field with $I_t^3F = 0$, we proved in Theorem 1.4 that $u(F) \leq 2\lambda(F) + 2$ under the assumption that $st(F) \leq 2$. We believe that this additional assumption is in fact superfluous, but we were unable to get this upper bound without it.

Conjecture 2.13. Let F be real with $I_t^3F = 0$. Then $u(F) \leq 2\lambda(F) + 2$.

In support of this conjecture, we can prove that it holds for small values of $\lambda(F)$.

Proposition 2.14. *Let F be real with $I_t^3F = 0$. If $\lambda = \lambda(F) \leq 4$ then $u(F) \leq 2\lambda + 2$.*

Proof. We will show that if φ is an anisotropic form in I_t^2F with $1 \leq t(\varphi) = t \leq 4$, then $\dim \varphi \leq 2t + 2$ and thus $\dim \varphi = 2t + 2$ by Lemma 2.2(ii), which by Lemma 2.4(i) immediately yields the desired result. (Note that $t(\varphi) = 0$ implies that $\varphi \in I_t^3F = 0$, i.e. φ is hyperbolic.)

So let $\varphi \in I_t^2F$ and suppose that $1 \leq t(\varphi) = t \leq 4$ and $\dim \varphi \geq 2t + 4$. By Lemma 2.2(i), there exists a form $\psi \in I^2F$ with $\dim \psi = 2t + 2$ such that $\varphi \equiv \psi \bmod I^3F$. Now $\langle 1, 1\rangle \otimes \varphi \in I_t^3F = 0$ and $\langle 1, 1\rangle \otimes (\varphi \perp -\psi) \in I^4F$, hence $\langle 1, 1\rangle \otimes \psi \in I^4F$. We have $\dim \langle 1, 1\rangle \otimes \psi = 4t + 4 \leq 20$. By the Arason-Pfister Hauptsatz and [H1, Main Theorem], there exists $\rho \in GP_4F$ such that $\langle 1, 1\rangle \otimes \psi \sim \rho$ in WF. After scaling, we may assume that $\rho \in P_4F$. Since ρ is divisible by $\langle 1, 1\rangle$, there exists $\sigma \in P_3F$ such that $\rho \cong \langle 1, 1\rangle \otimes \sigma$. Comparing signatures, we see that $\mathrm{sgn}_P\, \psi = \mathrm{sgn}_P\, \sigma$ for all $P \in X_F$. Thus, $\varphi \perp -\psi \perp \sigma \in I_t^3F = 0$. Thus, in WF we get $\varphi \perp \sigma \sim \psi$. Now $\dim(\varphi \perp \sigma) \geq 2t + 12$ and $\dim \psi = 2t + 2$, hence $i_W(\varphi \perp \sigma) \geq 5$. Therefore, φ contains a 5-dimensional Pfister neighbor of σ. Since 5-dimensional Pfister neighbors always contain a subform in GP_2F, we have that there exists $\tau \in GP_2F$ such that $\tau \subset \varphi$. Thus, φ is isotropic by Lemma 2.3. □

3. Construction of fields with prescribed invariants

We will now focus on the realizability of given triples $(\lambda, u, \tilde{u})$ for nonlinked SAP-fields with $I_t^3 = 0$. Let us restate the corresponding theorem from the introduction, whose proof will take up most of the remainder of this section.

Theorem 3.1. *Let $\mathcal{M} = \{(n, 2n, 2n + 2),\ (n, 2n + 2, 2n + 2);\ n \geq 2\} \cup \{(n, 2n, \infty),\ (n, 2n + 2, \infty); n \geq 2\} \cup \{(\infty, \infty, \infty)\}$.*

(i) *Let F be a real SAP field such that $\lambda(F) \geq 2$ and $I_t^3F = 0$. Then $(\lambda(F), u(F), \tilde{u}(F)) \in \mathcal{M}$.*

(ii) *Let* $(\lambda, u, \tilde{u}) \in \mathcal{M}$. *Then there exists a real SAP field* F *with* $I_t^3 F = 0$ *and* $(\lambda(F), u(F), \tilde{u}(F)) = (\lambda, u, \tilde{u})$. *In the case where* $\tilde{u} < \infty$ *or* $\lambda = \infty$, *there exist such fields which are uniquely ordered.*

Proof. (i) This follows immediately from Theorems 1.3, 1.4.

(ii) We fix once and for all a real field F_0. Our constructions will be divided into three cases : Finite λ and finite $\tilde{u}$, finite λ and infinite $\tilde{u}$, and infinite λ.

$\boxed{\textit{The case } 2 \leq \lambda < \infty \textit{ and } \tilde{u} < \infty}$

Put $n = \lambda + 1$. We have to construct fields F, F' with $(\lambda(F), u(F), \tilde{u}(F)) = (n-1, 2n, 2n)$ and $(\lambda(F'), u(F'), \tilde{u}(F')) = (n-1, 2n-2, 2n)$.

Let $F_1 = F_0(x_1, x_2, \cdots, y_1, y_2, \cdots)$ be the rational function field in an infinite number of variables x_i, y_j over F_0. Consider the multiquaternion algebras $A_n = (1 + x_1^2, y_1) \otimes \cdots \otimes (1 + x_{n-1}^2, y_{n-1})$ and $B_n = A_{n-1} \otimes (-1, -1)$, $n \geq 2$, which are division algebras (cf. [H2, Lemma 2(iv)]). Let ψ_n be a $2n$-dimensional Albert form of A_n such that $\psi_n \sim \sum_{i=1}^{n-1} c_i \langle\langle 1 + x_{i-1}^2, y_{i-1} \rangle\rangle$ in WF_1 for suitable $c_i \in F_1^*$, and let ψ'_n be a $2n$-dimensional Albert form of B_n such that $\psi'_n \sim \langle\langle -1, -1 \rangle\rangle + c\psi_{n-1}$ for suitable $c \in F_1^*$. Since $\operatorname{sgn}_P \langle\langle 1 + x_{i-1}^2, y_{i-1} \rangle\rangle = 0$ and $\operatorname{sgn}_P \langle\langle -1, -1 \rangle\rangle = 4$ for each $P \in X_{F_1}$, we have $\operatorname{sgn}_P \psi_n = 0$ and $\operatorname{sgn}_P \psi'_n = 4$ for all $P \in X_{F_1}$. Now fix any ordering $P_1 \in X_{F_1}$.

Suppose that L is a field such that $(A_n)_L$ (resp. $(B_n)_L$) is a division algebra and such that P_1 extends to an ordering $P \in X_L$. Consider the following classes of forms over L :

$$\begin{aligned}
\mathcal{C}_1(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \dim \alpha = 2n+1,\ \alpha \text{ indefinite at } P\} \\
\mathcal{C}_2(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \alpha \in I^3 L,\ \operatorname{sgn}_P \alpha = 0\} \\
\mathcal{C}_3(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \dim \alpha = 2n,\ \operatorname{sgn}_P \alpha = 0\}
\end{aligned}$$

We construct an infinite tower of fields $F_1 \subset F_2 \subset \cdots$ and $F_1 = F'_1 \subset F'_2 \subset \cdots$ as follows. Suppose we have constructed F_i (resp. F'_i), $i \geq 1$ such that $(A_n)_{F_i}$ (resp. $(B_n)_{F'_i}$) are division algebras and such that P_1 extends to an ordering $P_i \in X_{F_i}$ (resp. X'_{F_i}).

Let F_{i+1} (resp. F'_{i+1}) be the compositum of all function fields $F_i(\alpha)$ (resp. $F'_i(\alpha)$) where $\alpha \in \mathcal{C}_1(F_i) \cup \mathcal{C}_2(F_i)$ (resp. $\mathcal{C}_1(F'_i) \cup \mathcal{C}_2(F'_i) \cup \mathcal{C}_3(F'_i)$).

Since an ordering P of a field L extends to an ordering of the function field $L(\alpha)$ of a form α over L if and only if α is indefinite at P, we see that there exists an ordering on F_{i+1} (resp. F'_{i+1}) extending the ordering P_i since we only take function fields of forms in the $\mathcal{C}_i$, and all these forms are indefinite at P_i (cf. [ELW, Th. 3.5 and Rem. 3.6]). We will fix such an ordering and call it P_{i+1}. Note that no other ordering on F_i (resp. F'_i) will extend to F_{i+1} (resp. F'_{i+1}). Indeed, let Q be any ordering on F_{i+1} (resp. F'_{i+1}) and let $b \in F_i^*$ (resp. $F_i'^*$) be such that $b <_{P_i} 0$ and $b >_Q 0$. Then $2n \times \langle 1 \rangle \perp \langle b \rangle$ is in $\mathcal{C}_1$ and definite at Q, which shows that Q will not extend.

Next, we show that A_n (resp. B_n) stays a division algebra over F_{i+1} (resp. F'_{i+1}). If $\alpha \in \mathcal{C}_1(L) \cup \mathcal{C}_2(L)$ and A_n (resp. B_n) is division over L, then it follows immediately from Lemma 2.2(iii), parts (a) and (d) that A_n (resp. B_n) stays division over $L(\alpha)$. In particular, this shows that $(A_n)_{F_{i+1}}$ will be division.

To show that B_n stays a division algebra over F'_{i+1}, it remains to show that if P_1 extends to an ordering P on L and B_n is division over L, then B_n stays a division algebra over $L(\alpha)$ for $\alpha \in \mathcal{C}_3(L)$. If $d_\pm\alpha \neq 1$, this follows from Lemma 2.2(iii), part (b). If $d_\pm\alpha = 1$, then $\alpha \in I^2L$, and by Lemma 2.2(iii), part (c) it suffices to show that $c(\alpha) \neq [(B_n)_L]$ in $\mathrm{Br}_2\, L$. Suppose $c(\alpha) = [(B_n)_L]$. Since $c((\psi'_n)_L) = [(B_n)_L]$, we have by Merkurjev's theorem that $\alpha \equiv (\psi'_n)_L \bmod I^3L$ and hence $0 = \mathrm{sgn}_P\,\alpha \equiv \mathrm{sgn}_P(\psi'_n)_L \equiv 4 \bmod 8$, clearly a contradiction.

With the F_i and their orderings P_i constructed for all i, we now put $F = \bigcup_{i=1}^\infty F_i$ (resp. $F' = \bigcup_{i=1}^\infty F'_i$) and $P = \bigcup_{i=1}^\infty P_i$. P will then be the unique ordering on F (resp. F') (see also the proof of [H3, Th. 2]). It is also obvious from our construction that $I_t^3F = 0$ and that indefinite forms of dimension $2n+1$ will be isotropic. The latter implies by [ELP, Th. A] that $\tilde{u}(F), \tilde{u}(F') \leq 2n$. Also, A_n (resp. B_n) will stay a division algebra over F (resp. F'). In the case of F, this means that the form $(\psi_n)_F$ will be a $2n$-dimensional torsion form which is anisotropic by Lemma 2.2(i). Hence $u(F) \geq 2n$ and thus $u(F) = \tilde{u}(F) = 2n$. In the case of F', we have by a similar reasoning that $(\psi'_n)_{F'}$ is a $2n$-dimensional indefinite anisotropic form (recall that $\dim(\psi'_n)_{F'} = 2n \geq 6 > 4 = \mathrm{sgn}_P(\psi'_n)_{F'}$). Hence $\tilde{u}(F') = 2n$. However, by construction, torsion forms of dimension $2n$ will be isotropic and thus $u(F') \leq 2n-2$. On the other hand, $B_n = A_{n-1} \otimes (-1,-1)$ will stay a division algebra over F' and thus also A_{n-1}. Hence, just as before, we will now have the anisotropic $(2n-2)$-dimensional torsion form $(\psi_{n-1})_{F'}$, which shows that $u(F') = 2n-2$.

The fact that $\lambda(F) = \lambda(F') = n-1$ follows from Corollary 2.11.

The case $2 \leq \lambda < \infty$ and $\tilde{u} = \infty$

With F_0 as above, we let now $F_1 = F_0(x_1, x_2, \cdots, y_1, y_2, \cdots)((t))$, but we keep the definitions of A_n, B_n, ψ_n, ψ'_n from above. Let L be any extension of F_1 such that all orderings of F_1 extend to L and such that A_n (resp. B_n) is division over L. This time, we consider the following classes of quadratic forms, where $n = \lambda + 1 \geq 3$.

$$\begin{aligned}
\mathcal{C}_1(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \dim\alpha \geq 2n+2,\\
&\qquad\qquad \alpha \cong \alpha_0 \perp \alpha_t,\ \alpha_t \in W_tL,\ \dim\alpha_0 \in \{0,4\}\}\\
\mathcal{C}_2(L) &= \{\alpha \,|\, \alpha = \langle 1,1\rangle \otimes \langle 1,x,y,-xy\rangle,\ x,y \in L^*\}\\
\mathcal{C}_3(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \alpha \in I_t^3L\}\\
\mathcal{C}_4(L) &= \{\alpha \,|\, \alpha \text{ form}/L,\ \dim\alpha = 2n,\ \alpha \in W_tL\}
\end{aligned}$$

Again, we construct infinite towers of fields $F_1 \subset F_2 \subset \cdots$ and $F_1 = F'_1 \subset F'_2 \subset \cdots$. Suppose we have constructed F_i resp. F'_i, $i \geq 1$. Then we let F_{i+1} (resp. F'_{i+1}) be the compositum of all function fields $F_i(\alpha)$ (resp. $F'_i(\alpha)$) where $\alpha \in \mathcal{C}_1(F_i) \cup \mathcal{C}_2(F_i) \cup \mathcal{C}_3(F_i)$ (resp. $\mathcal{C}_1(F'_i) \cup \mathcal{C}_2(F'_i) \cup \mathcal{C}_3(F'_i) \cup \mathcal{C}_4(F'_i)$).

We then put $F = \bigcup_{i=1}^\infty F_i$ (resp. $F' = \bigcup_{i=1}^\infty F'_i$). Note that since we only take function fields of t.i. forms, all orderings of F_1 extend to F, resp. F'. In particular, F, F' will be real.

Now a field F is SAP if and only if all forms of type $\langle 1, a, b, -ab\rangle$ are weakly isotropic, i.e. there exists an n such that the n-fold orthogonal sum $n \times \langle 1, a, b, -ab\rangle$ is isotropic (cf. [P, Satz 3.1], [ELP, Th. C]). Thus, taking function fields of forms of type $\langle 1, 1\rangle \otimes \langle 1, x, y, -xy\rangle$ assures that F (resp. F') is SAP. Taking function fields of forms in I_t^3 yields that $I_t^3 F = 0$ (resp. $I_t^3 F' = 0$).

We now show that $(B_n)_K$ is a division algebra for $K = F, F'$. This then implies that $\lambda(K) \geq n-1$. Let L be an extension of F_1 such that all orderings of F_1 extend to L and suppose we have that $(B_n)_L$ is division. Then B_n stays division over $L(\alpha)$ for $\alpha \in \mathcal{C}_j(L)$, $j = 1, 3, 4$, by a reasoning similar to above after invoking Lemma 2.2(iii). Also, B_n stays division over $K = L(\langle\langle -1, -x, -y\rangle\rangle)$ for all $x, y \in L^*$ by part (d) of Lemma 2.2(iii). Now $\alpha = \langle 1, 1\rangle \otimes \langle 1, x, y, -xy\rangle$ contains the Pfister neighbor $\langle 1, 1\rangle \otimes \langle 1, x, y\rangle$ of $\langle\langle -1, -x, -y\rangle\rangle$, therefore α becomes isotropic over K, hence $K(\alpha)/K$ is purely transcendental and B_n stays division over $K(\alpha) = L(\langle\langle -1, -x, -y\rangle\rangle)(\alpha)$ and therefore over $L(\alpha)$.

This shows that $(B_n)_K$ is a division algebra for $K = F, F'$. Hence, $\lambda(K) \geq n-1$. By a similar reasoning, $(A_n)_F$ is a division algebra.

Suppose that $\lambda(K) \geq n$. Then there exists $C \in \mathrm{Br}_2(K)$ such that $t(C) = n$. Now K is SAP and $I_t^3 K = 0$. Hence, by Lemma 2.2(i) and Lemma 2.10(iii), there exists an anisotropic Albert form α of dimension $2n+2$ associated with C such that $\alpha \cong \alpha_0 \perp \alpha_t$ with $\alpha_t \in W_t F$ and $\dim \alpha_0 \in \{0, 4\}$. But such an α is by construction isotropic (consider the forms in $\mathcal{C}_1$ above !), a contradiction. Hence $\lambda(K) = n-1$. By Theorem 1.4, we get $u(K) \in \{2n-2, 2n\}$.

Now over F', we have by construction that all torsion forms of dimension $2n$ are isotropic (consider the forms in $\mathcal{C}_4$ above !). Thus, $u(F') = 2n-2 = 2\lambda(F')$.

We already remarked that $(A_n)_F$ is a division algebra. Hence, its associated Albert form $(\psi_n)_F$ is anisotropic and torsion. Therefore, $u(F) \geq 2n$ and we necessarily have $u(F) = 2n$.

It remains to show that $\tilde{u}(F) = \tilde{u}(F') = \infty$. Let m be a positive integer and let $\mu_m = m \times \langle 1\rangle \perp t\langle 1, -(1+x_1^2)\rangle$ over F_1. Since $m \times \langle 1\rangle$ and $\langle 1, -(1+x_1^2)\rangle$ are anisotropic over $F_0(x_1, x_2, \cdots, y_1, y_2, \cdots)$, it follows from Springer's theorem [L1, Ch. VI, Prop. 1.9] that μ_m is anisotropic. Furthermore, μ_m is t.i. as $\langle 1, -(1+x_1^2)\rangle$ is a binary torsion form. Thus, if we can show that μ_m stays anisotropic over F (resp. F') for all m, then $\tilde{u}(F), \tilde{u}(F') \geq 2m+2$ for all m and thus $\tilde{u}(F) = \tilde{u}(F') = \infty$.

We now construct a tower of fields $L_1 \subset L_2 \subset \cdots$ such that L_i will be the power series field in the variable t over some L_i', $L_i = L_i'((t))$, such that $F_i \subset L_i$ (resp. $F_i' \subset L_i$), and $(\mu_m)_{L_i}$ anisotropic for all $m \geq 0$ and all $i \geq 1$. This then shows that $(\mu_m)_{F_i}$ (resp. $(\mu_m)_{F_i'}$) is anisotropic for all $m \geq 0$, $i \geq 1$, and therefore $(\mu_m)_F$ (resp. $(\mu_m)_{F'}$) will be anisotropic for all $m \geq 0$.

Suppose we have constructed $L_i = L_i'((t))$. Note that necessarily L_i is real as $(\mu_m)_{L_i}$ is anisotropic for all $m \geq 0$. Let $P_i \in X_{L_i'}$ be any ordering and M_i' be the compositum over L_i' of the function fields of all forms (defined over L_i') in

$$\mathcal{C}'(L_i') = \{\alpha \,|\, \alpha \text{ indefinite at } P_i, \dim \alpha \geq 3\} \; .$$

Let $M_i = M_i'((t))$.

Now let $\rho \in \mathcal{C}_j(F_i)$ (resp. $\mathcal{C}_j(F_i')$), $1 \leq j \leq 4$. By Springer's theorem, $\rho_{L_i} \cong \rho_1 \perp t\rho_2$ with ρ_k, $k = 1, 2$, defined over L_i'. We will show that $\rho_k \in \mathcal{C}'(L_i')$ for at least one $k \in \{1, 2\}$.

First, note that forms in $\mathcal{C}_j(F_i)$ (resp. $\mathcal{C}_j(F_i')$), $1 \leq j \leq 4$, are of dimension ≥ 6 (recall that $2 \leq \lambda = n-1$). Thus, $\dim \rho_k \geq 3$ for at least one $k \in \{1, 2\}$. If ρ_{L_i} is isotropic, then over L_i' we have $\langle 1, -1\rangle \subset \rho_k$ for at least one $k \in \{1, 2\}$, and since $\langle 1, -1\rangle \cong t\langle 1, -1\rangle$, we may "shift" the hyperbolic plane from one ρ_k to the other if necessary to obtain the desired result, namely that $\rho_k \in \mathcal{C}'(L_i')$ for at least one $k \in \{1, 2\}$.

Let us therefore assume that ρ_{L_i} is anisotropic.

Suppose $\rho \in \mathcal{C}_1(F_i)$ (resp. $\mathcal{C}_1(F_i')$). Then $\dim \rho \geq 8$, and we can write $\rho \cong \eta \perp \tau$ over F_i, with τ torsion and $\dim \tau \geq 4$. Now $\tau_{L_i} \cong \tau_1 \perp t\tau_2$ with τ_k, $k = 1, 2$, defined over L_i'. Since τ is torsion, we have that τ_1 and τ_2 are torsion. Now $\tau_k \subset \rho_k$ over L_i' by Springer's theorem as ρ_{L_i} is anisotropic, and a simple dimension count shows that there exists at least one $k \in \{1, 2\}$ such that $\dim \tau_k \geq 2$ and $\dim \rho_k \geq 4$, which implies that for this k we have $\rho_k \in \mathcal{C}'(L_i')$.

Suppose $\rho \cong \langle 1, 1\rangle \otimes \langle 1, x, y, -xy\rangle \in \mathcal{C}_2(F_i)$ (resp $\mathcal{C}_2(F_i')$). Then either ρ_{L_i} is already defined over L_i', in which case it is a t.i. form of dimension 8 and thus in $\mathcal{C}'(L_i')$. Or there exist $a, b \in L_i'^*$ such that $\rho_{L_i} \cong \langle 1, 1\rangle \otimes \langle 1, a\rangle \perp bt\langle 1, 1\rangle \otimes \langle 1, -a\rangle$. then either $\langle 1, 1\rangle \otimes \langle 1, a\rangle$ is indefinite at P_i and thus in $\mathcal{C}'(L_i')$, or $\langle 1, 1\rangle \otimes \langle 1, -a\rangle$ is indefinite at P_i and thus in $\mathcal{C}'(L_i')$.

Finally, if $\rho \in \mathcal{C}_j(F_i)$ (resp. $\rho \in \mathcal{C}_j(F_i')$), $j = 3, 4$, then ρ is already torsion of dimension ≥ 6 (for $j = 3$ this follows from the Arason-Pfister Hauptsatz), but then ρ_1 and ρ_2 are torsion over L_i', and since at least one of them is necessarily of dimension ≥ 4, the result follows.

Thus, each $\rho \in \mathcal{C}_j(F_i)$ (resp. $\mathcal{C}_j(F_i)$), $1 \leq j \leq 4$ has the property that $\rho_{L_i} \cong \rho_1 \perp t\rho_2$ with ρ_k, $k = 1, 2$, defined over L_i' and $\rho_k \in \mathcal{C}'(L_i')$ for at least one k. But then, $(\rho_k)_{M_i'}$ is isotropic by construction, hence also ρ_{M_i}. In particular, $M_i(\rho)/M_i$ is a purely transcendental extension.

Let us now show that $(\mu_m)_F$ is anisotropic for all m. Let N_i be the compositum of the function fields of all forms α_{M_i} with $\alpha \in \mathcal{C}_1(F_i) \cup \mathcal{C}_2(F_i) \cup \mathcal{C}_3(F_i)$. By the above, N_i/M_i is purely transcendental. Let B be a transcendence basis so that $N_i = M_i(B) = M_i'((t))(B)$. We now put $L_{i+1}' = M_i'(B)$ and $L_{i+1} = L_{i+1}'((t)) = M_i'(B)((t))$. There are obvious inclusions $F_{i+1} \subset N_i = M_i'((t))(B) \subset M_i'(B)((t)) = L_{i+1}$. Since M_i' is obtained from L_i' by taking function fields of forms indefinite at P_i, we see that P_i extends to an ordering on M_i' and thus clearly also to orderings on L_{i+1}'.

To show that $(\mu_m)_F$ is anisotropic, it thus suffices to show that if μ_m is anisotropic over L_i, then it stays anisotropic over L_{i+1}. Now $m \times \langle 1\rangle$ is clearly anisotropic over the real field L_{i+1}'. Also, $\langle 1, -(1 + x_1^2)\rangle$, which is anisotropic over L_i' by assumption, stays anisotropic over L_{i+1}' as L_{i+1}' is obtained by taking function fields of forms of of dimension ≥ 3 over L_i' followed by a purely transcendental extension. By Springer's theorem, $(\mu_m)_{L_{i+1}} = (m \times \langle 1\rangle \perp t\langle 1, -(1 + x_1^2)\rangle)_{L_{i+1}}$ is anisotropic.

The proof for F' is the same as above except that we have to take N_i above to be the compositum of the function fields of all forms α_{M_i} with $\alpha \in \mathcal{C}_1(F'_i) \cup \mathcal{C}_2(F'_i) \cup \mathcal{C}_3(F'_i) \cup \mathcal{C}_4(F'_i)$.

$\boxed{\textit{The case } \lambda = u = \tilde{u} = \infty}$

This can be done by the same type of construction as above, but this time we only consider function fields of forms of the following types :

$$\begin{array}{l} \mathcal{C}_1(L) = \{\alpha \,|\, \alpha = \langle 1,1\rangle \perp \langle 1,x,y,-xy\rangle,\ x,y \in L^*\} \\ \mathcal{C}_2(L) = \{\alpha \,|\, \alpha \text{ form}/L,\ \alpha \in I_t^3 L\} \end{array}$$

The field F we will obtain has, just as before, the property SAP and $I_t^3 F = 0$. Furthermore, the algebra A_n will stay a division algebra over F for all $n \geq 3$. Hence $\lambda(F) = \infty$ and it follows immediately that $u(F) = \tilde{u}(F) = \infty$. (Note that for each $n \geq 2$ the form $(\psi_n)_F$ will be an anisotropic $2n$-dimensional torsion form.) ☐

Now we can prove Corollary 1.6 from the introduction, which we restate in a more detailed version for the reader's convenience.

Corollary 3.2. *Let F be a real field with $I_t^3 F = 0$. Then*

$$(u(F), \tilde{u}(F)) \in \{(2n, 2n); n \geq 0\} \cup \{(2n, \infty); n \geq 0\} \cup \{(2n, 2n+2); n \geq 2\}\ .$$

All pairs of values on the right hand side can be realized as pairs $(u(F), \tilde{u}(F))$ for suitable real F with $I_t^3 F = 0$. Furthermore, there exist such F which are SAP with the only exceptions being the pairs $(0, \infty)$, $(2, \infty)$.

Proof. Let us first show that no other values are possible. By Lemma 2.4, u and $\tilde{u}$ are always even or infinite. If F in non-SAP, then $\tilde{u}(F) = \infty$. So suppose that F is SAP. If $u(F) \leq 2$, then $\tilde{u}(F) \leq 2$ by [ELP, Theorems E,F], and it follows readily that $u(F) = \tilde{u}(F) \in \{0, 2\}$. Note that this also shows that $(0, \infty)$, $(2, \infty)$ cannot be realized by SAP-fields. If F is linked, then by Theorem 1.1, $u(F) = \tilde{u}(F) \in \{0, 2, 4, 8\}$. If, however F is non-linked, then Theorem 1.5 (3.1) shows that there can be no other pairs $(u, \tilde{u})$ than the ones in the statement of the corollary.

The pairs $(u, \tilde{u}) = (0, 0)$ (resp. $(2, 2)$) can be realized by $\mathbb{R}$ (resp. the rational function field in one variable over the reals, $\mathbb{R}(X)$). Real global fields have $(u, \tilde{u}) = (4, 4)$. $(u, \tilde{u}) = (0, \infty)$ is realized by $\mathbb{R}((X))((Y))$, see also Example 2.12. Examples of fields with $(u, \tilde{u}) = (2, \infty)$ can be found in [EP, Cor. 5.2]. All other combinations have been realized in Theorem 1.5 (3.1) by SAP-fields. ☐

References

[AEJ] Arason, J.Kr.; Elman, R.; Jacob, B. *The graded Witt ring and Galois cohomology* I. Quadratic and hermitian forms (eds. I. Hambelton, C.R. Riehm), CMS Conf. Proc. Vol. **4** (1984), 17–50.

[E] Elman, R.: *Quadratic forms and the u-invariant*, III. Proc. of Quadratic Forms Conference (ed. G. Orzech). Queen's Papers in Pure and Applied Mathematics No. **46** (1977), 422–444.

[EL1] Elman, R.; Lam, T.Y.: *Quadratic forms and the u-invariant*, I. Math. Z. **131** (1973), 283–304.

[EL2] Elman, R.; Lam, T.Y.: *Quadratic forms and the u-invariant*, II. Invent. Math. **21** (1973), 125–137.

[EL3] Elman, R.; Lam, T.Y.: *Classification theorems for quadratic forms over fields.* Comm. Math. Helv. **49** (1974), 373–381.

[EL4] Elman, R.; Lam, T.Y.: *Quadratic forms under algebraic extensions.* Math. Ann. **219** (1976), 21–42.

[ELP] Elman, R.; Lam, T.Y.; Prestel, A.: *On some Hasse principles over formally real fields.* Math. Z. **134** (1973), 291–301.

[ELW] Elman, R.; Lam, T.Y.; Wadsworth, A.R.: *Orderings under field extensions.* J. Reine Angew. Math. **306** (1979), 7–27.

[EP] Elman, R.; Prestel, A.: *Reduced stability of the Witt ring of a field and its Pythagorean closure.* Amer. J. Math. **106** (1983), 1237–1260.

[H1] Hoffmann, D.W.: *On the dimensions of anisotropic quadratic forms in I^4.* Invent. Math. **131** (1998), 185–198.

[H2] Hoffmann, D.W.: *On Elman and Lam's filtration of the u-invariant.* J. Reine Angew. Math. **495** (1998), 175–186.

[H3] Hoffmann, D.W.: *Pythagoras numbers of fields.* J. Amer. Math. Soc. **12** (1999), 839–848.

[H4] Hoffmann, D.W.: *Isotropy of quadratic forms and field invariants.* Cont. Math. **272** (2000), 73–101.

[Hor] Hornix, E.A.M.: *Formally real fields with prescribed invariants in the theory of quadratic forms.* Indag. Math., New Ser. **2** (1991), 65–78.

[Ka] Kahn, B.: *Quelques remarques sur le u-invariant.* Sém. de Th. des Nombres, Bordeaux **2** (1990), 155–161. Erratum : Sém. de Th. des Nombres, Bordeaux **3** (1991), 247.

[KS] Knebusch, M.; Scheiderer, C.: Einführung in die reelle Algebra. Braunschweig, Wiesbaden: Vieweg 1989.

[Kr] Krüskemper, M.: *Annihilators in graded Witt rings and Milnor's K-theory.* Contemp. Math. **155** (1994), 307–320.

[L1] Lam, T.Y.: The Algebraic Theory of Quadratic Forms. Reading, Massachusetts: Benjamin 1973 (revised printing 1980).

[L2] Lam, T.Y.: *Some consequences of Merkurjev's work on function fields.* Unpublished manuscript (1989).

[M1] Merkurjev, A.S.: *On the norm residue symbol of degree 2.* Dokladi Akad. Nauk. SSSR **261** (1981), 542–547. (English translation: Soviet Math. Doklady **24** (1981), 546–551.)

[M2] Merkurjev, A.S.: *Simple algebras and quadratic forms.* Izv. Akad. Nauk. SSSR **55** (1991), 218–224. (English translation: Math. USSR Izvestiya **38** (1992), 215–221.)

[Pf] Pfister, A.: Quadratic forms with applications to algebraic geometry and topology. London Math. Soc. Lect. Notes **217**, Cambridge University Press 1995.

[P] Prestel, A: *Quadratische Semi-Ordnungen und quadratische Formen.* Math. Z. **133** (1973), 319–342.

[PW] Prestel, A; Ware, R.: *Almost isotropic quadratic forms.* J. London Math. Soc. **19** (1979), 241–244.

[S] Scharlau, W.: Quadratic and Hermitian Forms. Grundlehren **270**, Berlin, Heidelberg, New York, Tokyo: Springer 1985.

[T] Tignol, J.-P.: *Réduction de l'indice d'une algèbre simple centrale sur le corps des fonctions d'une quadrique.* Bull. Soc. Math. Belgique Sér. A **42** (1990), 735–745.

Detlev W. Hoffmann
Laboratoire de Mathématiques
UMR 6623 du CNRS
Université de Franche-Comté
16 route de Gray
F-25030 Besançon Cedex
detlev@math.univ-fcomte.fr

Quadratic Quaternion Forms, Involutions and Triality

Max-Albert Knus and Oliver Villa

Received: May 31, 2001

Communicated by Ulf Rehmann

Abstract. Quadratic quaternion forms, introduced by Seip-Hornix (1965), are special cases of generalized quadratic forms over algebras with involutions. We apply the formalism of these generalized quadratic forms to give a characteristic free version of different results related to hermitian forms over quaternions:
1) An exact sequence of Lewis
2) Involutions of central simple algebras of exponent 2.
3) Triality for 4-dimensional quadratic quaternion forms.

1991 Mathematics Subject Classification: 11E39, 11E88
Keywords and Phrases: Quadratic quaternion forms, Involutions, Triality

1. Introduction

Let F be a field of characteristic not 2 and let D be a quaternion division algebra over F. It is known that a skew-hermitian form over D determines a symmetric bilinear form over any separable quadratic subfield of D and that the unitary group of the skew-hermitian form is the subgroup of the orthogonal group of the symmetric bilinear form consisting of elements which commute with a certain semilinear mapping (see for example Dieudonné [3]). Quadratic forms behave nicer than symmetric bilinear forms in characteristic 2 and Seip-Hornix developed in [9] a complete, characteristic-free theory of quadratic quaternion forms, their orthogonal groups and their classical invariants. Her theory was subsequently (and partly independently) generalized to forms over algebras (even rings) with involution (see [11], [10], [1], [8]).
Similitudes of hermitian (or skew-hermitian) forms induce involutions on the endomorphism algebra of the underlying space. To generalize the case where only similitudes of a quadratic form are considered, the notion of a quadratic pair was worked out in [6]. Relations between quadratic pairs and generalized quadratic forms were first discussed by Elomary [4].

The aim of this paper is to apply generalized quadratic forms to give a characteristic free presentation of some results on forms and involutions. After briefly recalling in Section 2 the notion of a generalized quadratic form (which, following the standard literature, we call an (ε, σ)-quadratic form) we give in Section 3 a characteristic-free version of an exact sequence of Lewis (see [7], [8, p. 389] and the appendix to [2]), which connects Witt groups of quadratic and quaternion algebras. The quadratic quaternion forms of Seip-Hornix are the main ingredient. Section 4 describes a canonical bijective correspondence between quadratic pairs and (ε, σ)-quadratic forms and Section 5 discusses the Clifford algebra. In particular we compare the definitions given in [10] and in [6]. In Section 6 we develop triality for 4-dimensional quadratic quaternion forms whose associated forms (over a separable quadratic subfield) are 3-Pfister forms. Any such quadratic quaternion form θ is an element in a triple $(\theta_1, \theta_2, \theta_3)$ of forms over 3 quaternions algebras D_1, D_2 and D_3 such that $[D_1][D_2][D_3] = 1$ in the Brauer group of F. Triality acts as permutations on such triples.

2. Generalized quadratic forms

Let D be a division algebra over a field F with an involution $\sigma : x \mapsto \overline{x}$. Let V be a finite dimensional right vector space over D. An F-bilinear form

$$k : V \times V \to D$$

is *sesquilinear* if $k(xa, yb) = \overline{a}k(x, y)b$ for all x, $y \in V$, a, $b \in D$. The additive group of such maps will be denoted by $\mathrm{Sesq}_\sigma(V, D)$. For any $k \in \mathrm{Sesq}_\sigma(V, D)$ we write

$$k^*(x, y) = \overline{k(y, x)}.$$

Let $\varepsilon \in F^\times$ be such that $\varepsilon\overline{\varepsilon} = 1$. A sesquilinear form k such that $k = \varepsilon k^*$ is called *ε-hermitian* and the set of such forms on V will be denoted by $\mathrm{Herm}_\sigma^\varepsilon(V, D)$. Elements of

$$\mathrm{Alt}_\sigma^\varepsilon(V, D) = \{g = f - \varepsilon f^* \mid f \in \mathrm{Sesq}_\sigma(V, D)\}.$$

are *ε-alternating forms.* We obviously have $\mathrm{Alt}_\sigma^{-\varepsilon}(V, D) \subset \mathrm{Herm}_\sigma^\varepsilon(V, D)$. We set

$$\mathrm{Q}_\sigma^\varepsilon(V, D) = \mathrm{Sesq}_\sigma(V, D) / \mathrm{Alt}_\sigma^\varepsilon(V, D)$$

and refer to elements of $\mathrm{Q}_\sigma^\varepsilon(V, D)$ as *(ε, σ)-quadratic forms.* We recall that (ε, σ)-quadratic forms were introduced by Tits [10], see also Wall [11], Bak [1] or Scharlau [8, Chapter 7]. For any algebra A with involution τ, let $\mathrm{Sym}^\varepsilon(A, \tau) = \{a \in A \mid a = \varepsilon\tau(a)\}$ and $\mathrm{Alt}^\varepsilon(A, \tau) = \{a \in A \mid a = c - \varepsilon\tau(c),\ c \in A\}$. To any class $\theta = [k] \in \mathrm{Q}_\sigma^\varepsilon(V, D)$, represented by $k \in \mathrm{Sesq}_\sigma(V, D)$, we associate a quadratic map

$$q_\theta : V \to D / \mathrm{Alt}^\varepsilon(D, \sigma), \quad q_\theta(x) = [k(x, x)]$$

where $[d]$ denotes the class of d in $D / \mathrm{Alt}_\sigma^\varepsilon(D)$. The ε-hermitian form

$$b_\theta(x, y) = k(x, y) + \varepsilon k^*(x, y) = k(x, y) + \varepsilon\overline{k(y, x)}$$

depends only on the class θ of k in $\mathrm{Q}^{\varepsilon}_{\sigma}(V,D)$. We say that b_θ is the *polarization* of q_θ.

Proposition 2.1. *The pair (q_θ, b_θ) satisfies the following formal properties:*

$$\begin{array}{lrcl} & q_\theta(x+y) & = & q_\theta(x)+q_\theta(y)+[b_\theta(x,y)] \\ (1) & q_\theta(xd) & = & \overline{d}q_\theta(x)d \\ & b_\theta(x,x) & = & q_\theta(x)+\varepsilon\overline{q_\theta(x)} \end{array}$$

for all $x,\ y \in V$, $d \in D$. Conversely, given any pair (q,b), $q : V \to D/\operatorname{Alt}^{\varepsilon}(D,\sigma)$, $b \in \operatorname{Herm}^{\varepsilon}_{\sigma}(V,D)$ satisfying (1)*, there exist a unique $\theta \in \mathrm{Q}^{\varepsilon}_{\sigma}(V,D)$ such that $q = q_\theta$, $b = b_\theta$.*

Proof. The formal properties are straightforward to verify. For the converse see [11, Theorem 1]. □

Example 2.2. Let $D = F$, $\sigma = Id_F$ and $\varepsilon = 1$. Then sesquilinear forms are F-bilinear forms, $\operatorname{Alt}^{\varepsilon}(D,\sigma) = 0$ and a (σ,ε)-quadratic form is a (classical) quadratic form. We denote the set of bilinear forms on V by $\operatorname{Bil}(V,F)$. Accordingly we speak of ε-symmetric bilinear forms instead of ε-hermitian forms.

Example 2.3. Let D be a division algebra with involution σ and let D be a finite dimensional (right) vector space over D. We use a basis of V to identify V with D^n and $\operatorname{End}_D(V)$ with the algebra $M_n(D)$ of $(n \times n)$-matrices with entries in D. For any $(n \times m)$-matrix $x = (x_{ij})$, let $x^* = \overline{x}^t$, where t is transpose and $\overline{x} = (\overline{x}_{ij})$. In particular the map $a \mapsto a^*$ is an involution of $A = M_n(D)$. If we write elements of D^n as column vectors $x = (x_1, \dots, x_n)^t$ any sesquilinear form k over D^n can be expressed as $k(x,y) = x^* a y$, with $a \in M_n(D)$, and $k^*(x,y) = x^* a^* y$. We write $\operatorname{Alt}_n(D) = \{a = b - \varepsilon b^*\} \subset M_n(D)$, so that $\mathrm{Q}^{\varepsilon}_{\sigma}(V,D) = M_n(D)/\operatorname{Alt}_n(D)$.

Example 2.4. Let D be a quaternion division algebra, i.e. D is a central division algebra of dimension 4 over F. Let K be a maximal subfield of D which is a quadratic Galois extension of F and let $\sigma : x \mapsto \overline{x}$ be the nontrivial automorphism of K. Let $j \in K \setminus F$ be an element of trace 1, so that $K = F(j)$ with $j^2 = j + \lambda$, $\lambda \in F$. Let $\ell \in D$ be such that $\ell x \ell^{-1} = \overline{x}$ for $x \in K$, $\ell^2 = \mu \in F^{\times}$. The elements $\{1, j, \ell, \ell j\}$ form a basis of D and $D = K \oplus \ell K$ is also denoted $[K,\mu)$. The F-linear map $\sigma : D \to D$, $\sigma(d) = \operatorname{Trd}_D(d) - d = \overline{d}$ is an involution of D (the "conjugation") which extends the automorphism σ of K. The element $\mathrm{N}(d) = d\sigma(d) = \sigma(d)d$ is the reduced norm of d. We have $\operatorname{Alt}^{-1}_{\sigma}(D) = F$ and $(\sigma,-1)$-quadratic forms correspond to the quadratic quaternion forms introduced by Seip-Hornix in [9]. Accordingly we call $(\sigma,-1)$-quadratic forms *quadratic quaternion forms*.

The restriction of the involution τ to the center Z of A is either the identity (*involutions of the first kind*) or an automorphism of order 2 (*involutions of the second kind*). If the characteristic of F is different from 2 or if the involution is of second kind there exists an element $j \in Z$ such that $j + \sigma(j) = 1$. Under such conditions the theory of (σ,ε)-quadratic forms reduces to the theory of ε-hermitian forms:

PROPOSITION 2.5. *If the center of D contains an element j such that $j+\sigma(j) = 1$, then $\operatorname{Herm}_\sigma^{-\varepsilon}(V,D) = \operatorname{Alt}_\sigma^{\varepsilon}(V,D)$ and a (σ,ε)-quadratic form is uniquely determined by its polar form b_θ.*

Proof. If $k = -\varepsilon k^* \in \operatorname{Herm}_\sigma^{-\varepsilon}(V,D)$, then $k = 1k = jk + \overline{j}k = jk - \overline{j}\varepsilon k^* \in \operatorname{Alt}_\sigma^{\varepsilon}(V,D)$. The last claim follows from the fact that polarization induces an isomorphism $\operatorname{Sesq}_\sigma(V,D)/\operatorname{Herm}_\sigma^{-\varepsilon}(V,D) \xrightarrow{\sim} \mathrm{Q}_\sigma^{\varepsilon}(V,D)$. □

For any left (right) D-space V we denote by ${}^\sigma V$ the space V viewed as right (left) D-space through the involution σ. If ${}^\sigma x$ is the element x viewed as an element of ${}^\sigma V$, we have ${}^\sigma x d = {}^\sigma\big(\sigma(d)x\big)$. Let V^* be the dual ${}^\sigma\operatorname{Hom}_D(V,D)$ as a right D-module, i.e., $({}^\sigma f d)(x) = {}^\sigma(\overline{d}f)(x)$, $x \in V$, $d \in D$. Any sesquilinear form $k \in \operatorname{Sesq}_\sigma(V,D)$ induces a D-module homomorphism $\widehat{k} : V \to V^*$, $x \mapsto k(x,-)$. Conversely any homomorphism $g : V \to V^*$ induces a sesquilinear form $k \in \operatorname{Sesq}_\sigma(V,D)$, $k(x,y) = g(x)(y)$ and the additive groups $\operatorname{Sesq}_\sigma(V,D)$ and $\operatorname{Hom}_D(V,V^*)$ can be identified through the map $h \mapsto \widehat{k}$. For any $f : V \to V'$, let $f^* : V'^* \to V^*$ be the transpose, viewed as a homomorphisms of right vector spaces. We identify V with V^{**} through the map $v \mapsto v^{**}$, $v^{**}(f) = \overline{f(v)}$. Then, for any $f \in \operatorname{Hom}_D(V,V^*)$, f^* is again in $\operatorname{Hom}_D(V,V^*)$ and $\widehat{k}^* = \widehat{k^*}$. A (σ,ε)-quadratic form q_θ is called *nonsingular* if its polar form b_θ induces an isomorphism $\widehat{b_\theta}$. A pair (V,q_θ) with q_θ nonsingular is called a (σ,ε)-*quadratic space*. For any vector space W, the *hyperbolic space* $V = W \oplus W^*$ equipped with the quadratic form q_θ, $\theta = [k]$ with

$$k\big((p,q),(p',q')\big) = q(p'),$$

is nonsingular. There is an obvious notion of orthogonal sum $V \perp V'$ and a quadratic space decomposes whenever its polarization does. Most of the classical theory of quadratic spaces extends to (σ,ε)-quadratic spaces. For example Witt cancellation holds and any (σ,ε)-quadratic space decomposes uniquely (up to isomorphism) as the orthogonal sum of its anisotropic part with a hyperbolic space. Moreover, if we exclude the case $\sigma = 1$ and $\varepsilon = -1$, any (σ,ε)-quadratic space has an orthogonal basis. A *similitude* of (σ,ε)-quadratic spaces $t : (V,q) \xrightarrow{\sim} (V',q')$ is a D-linear isomorphism $V \xrightarrow{\sim} V'$ such that $q'(tx) = \mu(t)q(x)$ for some $\mu(t) \in F^\times$. The element $\mu(t)$ is called the *multiplier* of the similitude. Similitudes with multipliers equal to 1 are *isometries*. As in the classical case there is a notion of Witt equivalence and corresponding Witt groups are denoted by $W^\varepsilon(D,\sigma)$.

3. AN EXACT SEQUENCE OF LEWIS

Let D be a quaternion division algebra. We fix a representation $D = [K,\mu) = K \oplus \ell K$, with $\ell^2 = \mu$, as in (2.4). Let V be a vector space over D. Any sesquilinear form $k : V \times V \to D$ can be decomposed as

$$k(x,y) = P(x,y) + \ell R(x,y)$$

with $P : V \times V \to K$ and $R : V \times V \to K$. The following properties of P and R are straightforward.

LEMMA 3.1. *1) $P \in \mathrm{Sesq}_\sigma(V,K)$, $R \in \mathrm{Sesq}_1(V,K) = \mathrm{Bil}(V,K)$.*
2) $k^* = P^* - \ell R^t$, where $P^*(x,y) = \overline{P(y,x)}$ and $R^t(x,y) = R(y,x)$.

The sesquilinearity of k implies the following identities:

$$\begin{array}{llll} R(x\ell,y) & = -P(x,y), & R(x,y\ell) & = \overline{P(x,y)} \\ P(x\ell,y) & = -\mu R(x,y), & P(x,y\ell) & = \mu\overline{R(x,y)} \\ P(x\ell,y\ell) & = -\mu\overline{P(x,y)}, & R(x\ell,y\ell) & = -\mu\overline{R(x,y)} \end{array} \tag{2}$$

Let V^0 be V considered as a (right) vector space over K (by restriction of scalars) and let $T : V^0 \to V^0, x \mapsto x\ell$. The map T is a K-semilinear automorphism of V^0 such that $T^2 = \mu$. Conversely, given a vector space U over K, together with a semilinear automorphism T such that $T^2 = \mu \in F^\times$, we define the structure of a right D-module on U, $D = [K,\mu)$, by putting $x\ell = T(x)$.

LEMMA 3.2. *Let V be a vector space over D. 1) Let $f_1 : V^0 \times V^0 \to K$ be a sesquilinear form over K. The form*

$$f(x,y) = f_1(x,y) - \ell\mu^{-1} f_1(Tx,y)$$

is sesquilinear over D if and only if $f_1(Tx,Ty) = -\mu\overline{f_1(x,y)}$.
2) Let $f_2 : V^0 \times V^0 \to K$ be a bilinear form over K. The form

$$f(x,y) = -f_2(Tx,y) + \ell f_2(x,y)$$

is sesquilinear over D if and only if $f_2(Tx,Ty) = -\mu\overline{f_2(x,y)}$.

Proof. The two claims follow from the identities (2). □

Let f be a bilinear form on a space U over K and let $\lambda \in K^\times$. A semilinear automorphism t of U such that $f(tx,ty) = \lambda\overline{f(x,y)}$ for all $x \in U$ is a *semilinear similitude* of (U,f), with *multiplier* λ. In particular $Tx = x\ell$ is a semilinear similitude of R on V^0, such that $T^2 = \mu$ and with multiplier $-\mu$. The following nice observation of Seip-Hornix [9, p. 328] will be used later:

PROPOSITION 3.3. *Let R be a K-bilinear form over U and let T be a semilinear similitude of U with multiplier $\lambda \in K^\times$ and such that $T^2 = \mu$. Then:*
1) $\mu \in F$,
2) For any $\xi \in K$ and $x \in U$, let $\rho_\xi(x) = x\xi$. There exists $\nu \in K^\times$ such that $T' = \rho_\nu \circ T$ satisfies ${T'}^2 = \mu'$ and $R(T'x,T'y) = -\mu'\overline{R(x,y)}$.

Proof. The first claim follows from $\mu = \lambda\overline{\lambda}$. For the second we may assume that $\lambda \neq \mu$ (if $\lambda = \mu$ replace T by $T \circ \rho_k$ for an appropriate k). For $\nu = (1 - \mu\lambda^{-1})$ we have $\mu' = 2\mu - \lambda - \overline{\lambda}$. □

Assume that $k \in \mathrm{Sesq}_\sigma(V, D)$ defines a (σ, ε)-quadratic space $[k]$ on V over D. It follows from (3.1) that P defines a (σ, ε)-quadratic space $[P]$ on V^0 over K and R a $(Id, -\varepsilon)$-quadratic space $[R]$ on V^0 over K. Let $K = F(j)$ with $j^2 = j + \lambda$. Let $r(x, y) = R(x, y) - \varepsilon R(y, x)$ be the polar of R.

PROPOSITION 3.4. *1)* $q_{[P]}(x) = \overline{\varepsilon} j[r(x, Tx)]$
2) $q_{[k]}(x) = \overline{\varepsilon} j[r(x, Tx)] + \ell q_{[R]}(x)$
3) The map T is a semilinear similitude of $\big(q_{[R]}, V^0\big)$ with multiplier $-\mu$.

Proof. It follows from the relations (2) that

$$\overline{P(x,x)} + \varepsilon P(x, x) = R(x, Tx) - \varepsilon R(Tx, x) = r(x, Tx) \tag{3}$$

and obviously this relation determines $P(x, x)$ up to a function with values in $\mathrm{Sym}^{-\varepsilon}(K, \sigma)$. Since $\mathrm{Sym}^{-\varepsilon}(K, \sigma) = \mathrm{Alt}^{+\varepsilon}(K, \sigma)$ by (2.5), $[P]$ is determined by (3). Since $\overline{r(x, Tx)} = \overline{\varepsilon} r(x, Tx)$ by (2), we have $\overline{\overline{\varepsilon} jr(x, Tx)} + \varepsilon\big(\overline{\varepsilon} jr(x, Tx)\big) = r(x, Tx)$ and 1) follows. The second claim follows from 1) and 3) is again a consequence of the identities (2). $\square$

COROLLARY 3.5. *Any pair $\big([R], T\big)$ with $[R] \in \mathrm{Q}_1^\varepsilon(U, K)$ and T a semilinear similitude with multiplier $-\mu \in F^\times$ and such that $T^2 = \mu$, determines the structure of a (σ, ε)-quadratic space on U over $D = [K, \mu)$.*

PROPOSITION 3.6. *The assignments $h \mapsto P$ and $h \mapsto R$ induce homomorphisms of groups $\pi_1 : W^\varepsilon(D, -) \to W^\varepsilon(K, -)$ and $\pi_2 : W^{-\varepsilon}(D, -) \to W^\varepsilon(K, Id)$.*

Proof. The assignments are obviously compatible with orthogonal sums and Witt equivalence. $\square$

We recall that $W^\varepsilon(K, -)$ can be identified with the corresponding Witt group of ε-hermitian forms (apply (2.5)). However, it is more convenient for the following computations to view ε-hermitian forms over K as (σ, ε)- quadratic forms. Let $i \in K^\times$ be such that $\sigma(i) = -i$ (take $i = 1$ if $\mathrm{Char}\, F = 2$). The map $k \mapsto ik$ induces an isomorphism $s : W^\varepsilon(K, -) \xrightarrow{\sim} W^{-\varepsilon}(K, -)$ ("scaling"). For any space U over K, let $U_D = U \otimes_K D$. We identify U_D with $U \oplus U\ell$ through the map $u \otimes (x + \ell y) \mapsto (ux, u\overline{y}l)$ and get a natural D-module structure on $U_D = U \oplus U\ell$. Any K-sesquilinear form k on U extends to a D-sesquilinear form k_D on U_D through the formula

$$k_D(x \otimes a, y \otimes b) = \overline{a} k(x, y) b$$

for $x,\, y \in U$ and $a,\, b \in D$.

LEMMA 3.7. *The assignment $k \mapsto (ik)_D$ induces a homomorphism*

$$\beta : W^\varepsilon(K, -) \to W^{-\varepsilon}(D, -)$$

Proof. Let $\widetilde{k} = (ik)_D$. We have $(\widetilde{k})^* = -\widetilde{k^*}$. $\square$

Theorem 3.8 (Lewis). *With the notations above, the sequence*

$$W^\varepsilon(D,-) \xrightarrow{\pi_1} W^\varepsilon(K,-) \xrightarrow{\beta} W^{-\varepsilon}(D,-) \xrightarrow{\pi_2} W^\varepsilon(K,Id)$$

is exact.

Proof. This is essentially the proof given in Appendix 2 of [2] with some changes due to the use of generalized quadratic forms, instead of hermitian forms. We first check that the sequence is a complex. Let $[k] \in \mathrm{Q}^\varepsilon_\sigma(V,D)$ and let $V^0 = U$. We write elements of $U_D = U \oplus U\ell$ as pairs $(x, y\ell)$ and decompose $k_D = P + \ell R$. By definition we have $\beta\pi_1([k]) = [\beta(P)]$ and

$$\begin{aligned}\beta(P)\big((x_1,y_1),(x_2,y_2)\big) \quad &= \quad i\big(P(x_1,x_2) + P(x_1,y_2)\ell + \ell P(y_1,x_2)\\ &\qquad +\ell P(y_1,y_2)\ell\big).\end{aligned}$$

Let $(x\ell, x\ell) \in U \oplus U\ell$. We get $\beta(P)\big((x\ell,x\ell),(x\ell,x\ell)\big) = 0$ hence $W = \{(x\ell,x\ell)\} \subset U \oplus U\ell$ is totally isotropic. It is easy to see that $W \subset W^\perp$, so that $[\beta(P)]$ is hyperbolic and $\beta \circ \pi_1 = 0$. Let $[g] \in \mathrm{Q}^\varepsilon_\sigma(U,K)$. The subspace $W = \{(x,0) \in U \oplus U\ell\}$ is totally isotropic for $\pi_2\beta([g])$ and $W \subset W^\perp$. Hence $\pi_2\beta([g]) = 0$. We now prove exactness at $W^\varepsilon(K,-)$. Since the claim is known if Char $\neq 2$, we may assume that Char $= 2$ and $\varepsilon = 1$. Let $[g] \in \mathrm{Q}^\varepsilon_\sigma(U,K)$ be anisotropic such that $\beta([g]) = 0 \in W^{-\varepsilon}(D,-)$. In particular $\beta([g]) \in \mathrm{Q}^{-\varepsilon}_\sigma(U_D,D)$ is isotropic. Hence the exist elements $x_1,\ x_2 \in U$ such that $[\widetilde{g}]\big((x_1,x_2\ell),(x_1,x_2\ell)\big) = 0$. This implies (in Char 2) that

$$g(x_1,x_1) + \mu\overline{g(x_2,x_2)} \in F, \quad g(x_1,x_2)\ell + \ell g(x_2,x_1) = 0. \tag{4}$$

Let V_1 be the K-subspace of V generated by x_1 and x_2. Since $[g]$ is anisotropic, $[g] = [g_1] \perp [g_2]$ with $g_1 = g|_{V_1}$. We make V_1 into a D-space by putting

$$(x_1a_1 + x_2a_2)\ell = \mu x_2\overline{a}_1 + x_1\overline{a}_2$$

To see that the action is well-defined, it suffices to show that $\dim_K V_1 = 2$. The elements x_1 and x_2 cannot be zero since $[g]$ is anisotropic, so assume $x_2 = x_1c,\ c \in K^\times$. Then (4) implies $g(x_1,x_1) + \mu c\overline{c}\overline{g(x_1,x_1)} \in F$, which contradicts the fact that g is anisotropic. Let $g_1(x_1,x_1) + \mu\overline{g_1(x_2,x_2)} = z \in F$. Let $f \in \mathrm{Sesq}_\sigma(V_1,K)$. Replacing g_1 by $g_1 + f + f^*$ defines the same class in $\mathrm{Q}^\varepsilon_\sigma(V_1,K)$ (recall that Char $F = 2$). Choosing f as

$$f(x_1,x_1) = jz,\ f(x_2,x_2) = 0,\ f(x_1,x_2) = f(x_2,x_1) = 0,$$

we may assume that

$$g_1(x_1,x_1) + \mu\overline{g_1(x_2,x_2)} = 0, \quad g_1(x_1,x_2)\ell + \ell g_1(x_2,x_1) = 0. \tag{5}$$

By (3.2) we may extend g_1 to a sesquilinear form

$$g'(x,y) = g_1(x,y) + \ell\mu^{-1}g_1(x\ell,y)$$

over D if g_1 satisfies

$$g_1(x\ell,y\ell) = -\mu\overline{g_1(x,y)}$$

This can easily be checked using (5) (and the definition of $x\ell$). Then g_1 is in the image of π_1. Exactness at $W^\varepsilon(K,-)$ now follows by induction on the dimension of U. We finally check exactness at $W^{-\varepsilon}(D,-)$. Let $[k]$ be anisotropic such that $\pi_2([k]) = 0$ in $W^{-\varepsilon}(K, Id)$. In particular $\pi_2([k])$ is isotropic; let $x \neq 0$ be such that $\pi_2 k(x,x) = 0$ and let W be the D-subspace of V generated by x. Since $[k]$ is anisotropic, $[k'] = [k|_W]$ is nonsingular and $[k] = [k'] \perp [k'']$. The condition $\pi_2 k(x,x) = 0$ implies $k(x,x) \in K$. Let W_1 be the K-subspace of W generated by x. Define $g : W_1 \times W_1 \to K$ by $g(xa, xb) = k(xa, xb)i^{-1}$ for a, $b \in K$. Then clearly $[g]$ defines an element of $W^{\varepsilon}(K,-)$ and $\beta(g) = k'$. Once again exactness follows by induction on the dimension of V. $\square$

4. Involutions on central simple algebras

Let D be a central division algebra over F, with involution σ and let $b : V \times V \to D$ be a nonsingular ε-hermitian form on a finite dimensional space over D. Let $A = \operatorname{End}_D(V)$. The map $\sigma_b : A \to A$ such that $\sigma_b(\lambda) = \sigma(\lambda)$ for all $\lambda \in F$ and

$$b\big(\sigma_b(f)(x), y\big) = b\big(x, f(y)\big)$$

for all x, $y \in V$, is an involution of A, called the involution *adjoint to* b. We have $\sigma_b(f) = \widehat{b}^{-1} f^* \widehat{b}$, where $\widehat{b} : V \xrightarrow{\sim} V^*$ is the adjoint of b. Conversely, any involution of A is adjoint to some nonsingular ε-hermitian form b and b is uniquely multiplicatively determined up to a σ-invariant element of $F^\times$.
Any automorphism ϕ of A compatible with σ_b, i.e., $\sigma_b\big(\phi(a)\big) = \phi\big(\sigma_b(a)\big)$, is of the form $\phi(a) = uau^{-1}$ with $u : V \xrightarrow{\sim} V$ a similitude of b. We say that an involution τ of A is a *q-involution* if τ is adjoint to the polar b_θ of a (σ,ε)-quadratic form θ. We write $\tau = \sigma_\theta$. Two algebras with q-involutions are *isomorphic* if the isomorphism is induced by a similitude of the corresponding quadratic forms. Over fields q-involutions differ from involutions only in characteristic 2 and for symplectic involutions. In view of possible generalizations (for example rings in which $2 \neq 0$ is not invertible) we keep to the general setting of (σ,ε)-quadratic forms. Let F_0 be the subfield of F of σ-invariant elements and let T_{F/F_0} be the corresponding trace.

LEMMA 4.1. *The symmetric bilinear form on A given by* $\operatorname{Tr}(x,y) = T_{F/F_0}\big(\operatorname{Trd}_A(xy)\big)$ *is nonsingular and* $\operatorname{Sym}(A,\tau)^\perp = \operatorname{Alt}(A,\tau)$.

Proof. If τ is of the first kind $F_0 = F$ and the claim is (2.3) of [6]. Assume that τ is of the second kind. Since the bilinear form $(x,y) \to \operatorname{Trd}_A(xy)$ is nonsingular, Tr is also nonsingular and it is straightforward that $\operatorname{Alt}(A,\tau) \subset \operatorname{Sym}(A,\tau)^\perp$. Equality follows from the fact that $\dim_{F_0} \operatorname{Alt}(A,\tau) = \dim_{F_0} \operatorname{Sym}(A,\tau) = \dim_F A$. $\square$

PROPOSITION 4.2. *Let (V,θ), $\theta = [k]$ be a (σ,ε)-quadratic space over D and let $h = \widehat{k} + \varepsilon\widehat{k}^* : V \xrightarrow{\sim} V^*$. The F_0-linear form*

$$f_\theta : \operatorname{Sym}(A,\sigma_\theta) \to F_0, \quad f_\theta(s) = \operatorname{Tr}(h^{-1}\widehat{k}s), \ s \in \operatorname{Sym}(A,\sigma_\theta)$$

depends only on the class θ and satisfies $f_\theta\big(x + \sigma_\theta(x)\big) = \operatorname{Tr}(x)$.

Proof. The first claim follows from (4.1) and the fact that if $k \in \mathrm{Alt}^{\varepsilon}_{\sigma}(V,D)$ then $h^{-1}\widehat{k} \in \mathrm{Alt}^{1}_{\sigma_\theta}(V,D)$. For the last claim we have:

$$\begin{aligned} f_\theta\big(x+\sigma_\theta(x)\big) &= \mathrm{Tr}\big(h^{-1}\widehat{k}(x+\sigma_\theta(x))\big) \\ &= \mathrm{Tr}\big(h^{-1}\widehat{k}x\big)+\mathrm{Tr}\big(h^{-1}\widehat{k}h^{-1}x^*h\big) \\ &= \mathrm{Tr}\big(h^{-1}\widehat{k}x\big)+\mathrm{Tr}\big(\widehat{k}h^{-1}x^*\big) \\ &= \mathrm{Tr}\big(h^{-1}\widehat{k}x\big)+\mathrm{Tr}\big(x(h^{-1})^*\widehat{k}^*\big) \\ &= \mathrm{Tr}\big(h^{-1}\widehat{k}x\big)+\mathrm{Tr}\big(h^{-1}\varepsilon\widehat{k}^*x\big)=\mathrm{Tr}(x). \end{aligned}$$

□

Lemma 4.3. *Let τ be an involution of $A=\mathrm{End}_D(V)$ and let f be A F_0-linear form on $\mathrm{Sym}(A,\tau)$ such that $f\big(x+\tau(x)\big)=\mathrm{Tr}(x)$ for all $x\in A$. There exists an element $u\in A$ such that $f(s)=\mathrm{Tr}(us)$ and $u+\tau(u)=1$. The element u is uniquely determined up to additivity by an element of $\mathrm{Alt}(A,\tau)$. We take $u=1/2$ if $\mathrm{Char}\,F\neq 2$.*

Proof. The proof of (5.7) of [6] can easily be adapted. □

Proposition 4.4. *Let τ be an involution of $A=\mathrm{End}_D(V)$ and let f be A F_0-linear form on $\mathrm{Sym}(A,\tau)$ such that $f\big(x+\tau(x)\big)=\mathrm{Tr}(x)$ for all $x\in A$.*
1) There exists a nonsingular (σ,ε)-quadratic form θ on V such that $\tau=\sigma_\theta$ and $f=f_\theta$.
2) $(\sigma_\theta,f_\theta)=(\sigma_{\theta'},f_{\theta'})$ if and only if $\theta'=\lambda\theta$ for $\lambda\in F_0$.
3) If $\tau=\sigma_\theta$ and $f=f_\theta$ with $f_\theta(s)=\mathrm{Tr}(us)$, the class of u in $A/\mathrm{Alt}(A,\sigma_\theta)$ is uniquely determined by θ.

Proof. Here the proof of (5.8) of [6] can adapted. We prove 1) for completeness. Let $\tau(x)=h^{-1}x^*h$, $h=\varepsilon h^*: V\xrightarrow{\sim}V^*$. Let $f(s)=\mathrm{Tr}(us)$ with $u+\tau(u)=1$ and let $k\in\mathrm{Sesq}_\sigma(V,D)$ be such that $\widehat{k}=hu: V\to V^*$. We set $\theta=[k]$. It is then straightforward to check that $h=k+\varepsilon k^*$. □

Proposition 4.5. *Let $\phi:\big(\mathrm{End}_D(V),\sigma_\theta\big)\xrightarrow{\sim}\big(\mathrm{End}_D(V'),\sigma_{\theta'}\big)$ be an isomorphism of algebras with involution. Let $f_\theta(s)=\mathrm{Tr}(us)$ and $f_{\theta'}(s')=\mathrm{Tr}(u's')$. The following conditions are equivalent:*
1) ϕ is an isomorphism of algebras with q-involutions.
2) $f_{\theta'}\big(\phi(s)\big)=f_\theta(s)$ for all $s\in\mathrm{Sym}\big(\mathrm{End}_D(V),\sigma_\theta\big)$.
3) $[\phi(u)]=[u']\in\mathrm{End}_D(V')/\mathrm{Alt}\big(\mathrm{End}_D(V'),\sigma_{\theta'}\big)$.

Proof. The implication 1) $\Rightarrow$ 2) is clear. We check that 2) $\Rightarrow$ 3). Let ϕ be induced by a similitude $t:(V,b_\theta)\xrightarrow{\sim}(V',b_{\theta'})$. Since $f_{\theta'}(\phi s)=f_\theta(s)$, we have $\mathrm{Tr}(t^{-1}u'ts)=\mathrm{Tr}(u'tst^{-1})=\mathrm{Tr}(us)$ for all $s\in\mathrm{Sym}\big(\mathrm{End}_D(V),\sigma_\theta\big)$, hence $[\phi(u)]=[u']$. The implication 3) $\Rightarrow$ 1) follows from the fact that u can be chosen as $h^{-1}\widehat{k}$, $h=\widehat{k}+\varepsilon\widehat{k}^*$. □

REMARK 4.6. We call the pair $(\sigma_\theta, f_\theta)$ a (σ,ε)-*quadratic pair* or simply a quadratic pair. It determines θ up to the multiplication by a σ-invariant scalar $\lambda \in F^\times$. In fact σ_θ determines the polar b_θ up to λ and f_θ determines u. We have $\theta = [\widehat{b_\theta} u]$.

EXAMPLE 4.7. Let $q : V \to F$ be a nonsingular quadratic form. The polar b_q induces an isomorphism $\psi : V \otimes_F V \xrightarrow{\sim} \mathrm{End}_F(V)$ such that $\sigma_q\big(\psi(x \otimes y)\big) = \psi(y \otimes x)$. Thus $\psi(x \otimes x)$ is symmetric and $f_q\big(\psi(x \otimes x)\big) = q(x)$ (see [6, (5.11)]. More generally, if V is a right vector space over D, we denote by ${}^*V$ the space V viewed as a left D-space through the involution σ of D. The adjoint $\widehat{b_\theta}$ of a (σ,ε)-quadratic space (V,θ) induces an isomorphism $\psi_\theta : V \otimes_D {}^\sigma V \xrightarrow{\sim} \mathrm{End}_D(V)$ and $\psi_\theta(xd \otimes x)$ is a symmetric element of $\big(\mathrm{End}_D(V), \sigma_\theta\big)$ for all $x \in V$ and all ε-symmetric $d \in D$. One has $f_\theta\big(\psi(xd \otimes x)\big) = [dk(x,x)]$, where $\theta = [k]$ (see [4, Theorem 7]).

5. CLIFFORD ALGEBRAS

Let σ be an involution of the first kind on D and let θ be a nonsingular (σ,ε)-quadratic form on V. Let σ_θ be the corresponding q-involution on $A = \mathrm{End}_D(V)$. We assume in this section that over a splitting $A \otimes_F \tilde{F} \xrightarrow{\sim} \mathrm{End}_{\tilde{F}}(M)$ of A, $\theta_{\tilde{F}} = \theta \otimes 1_{\tilde{F}}$ is a $(Id, 1)$-quadratic form $\tilde{q}$ over $\tilde{F}$, i.e. $\theta_{\tilde{F}}$ is a (classical) quadratic form. In the terminology of [6] this means that σ_θ is orthogonal if Char $\neq 2$ and symplectic if Char $= 2$. From now on we call such forms over D *quadratic forms over* D, resp. *quadratic spaces* over D if the forms are non-singular.
Classical invariants of quadratic spaces (V,θ) are the dimension $\dim_D V$ and the discriminant $\mathrm{disc}(\theta)$ and the Clifford invariant associated with the Clifford algebra. We refer to [6, §7] for the definition of the discriminant. We recall the definition of the Clifford algebra $\mathrm{Cl}(V,\theta)$, following [10, 4.1]. Given (V,θ) as above, let $\theta = [k]$, $k \in \mathrm{Sesq}_\sigma(V,D)$, $b_\theta = k + \varepsilon k^*$ and $h = \widehat{b_\theta} \in \mathrm{Hom}_D(V,V^*)$. Let $A = \mathrm{End}_D(V)$, $B = \mathrm{Sesq}_\sigma(V,D)$ and $B' = V \otimes_D {}^\sigma V$. We identify A with $V \otimes_D {}^\sigma V^*$ through the canonical isomorphism $(x \otimes {}^\sigma f)(v) = xf(v)$ and B with $V^* \otimes_D {}^\sigma V^*$ through $(f \otimes {}^\sigma g)(x,y) = \overline{g(x)} f(y)$. The isomorphism h can be used to define further isomorphisms:

$$\varphi_\theta : B' = V \otimes_D {}^\sigma V \xrightarrow{\sim} A = \mathrm{End}_D(M),\ \varphi_\theta : x \otimes y \mapsto x \otimes h(y)$$

and the isomorphism ψ_θ already considered in (4.7):

$$\psi_\theta : A \xrightarrow{\sim} B,\ \psi_\theta : x \otimes {}^\sigma f \mapsto h(x) \otimes {}^\sigma f.$$

We use φ_θ and ψ_θ to define maps $B' \times B \to A$, $(b', b) \mapsto b'b$ and $A \times B' \to B'$, $(a, b') \mapsto ab'$:

$$(x \otimes {}^\sigma y)(h(u) \otimes g) = xb(y,u) \otimes {}^\sigma f \text{ and } (x \otimes, {}^\sigma f)(u \otimes, {}^\sigma v) = xf(u) \otimes {}^\sigma h(v)$$

Furthermore, let $\tau_\theta = \varphi_\theta^{-1} \sigma_\theta \varphi_\theta : B' \to B'$ be the transport of the involution σ_θ on A. We have $\tau_\theta(x \otimes {}^\sigma y) = \varepsilon y \otimes {}^\sigma x$. Let $S_1 = \{s_1 \in B' \mid \tau_\theta(s_1) = s_1\}$. We have $S_1 = \big(\mathrm{Alt}^\varepsilon(V,D)\big)^\perp$ for the pairing $B' \times B \to F$, $(b',b) \mapsto \mathrm{Trd}_A(b'b)$. Let Sand be the bilinear map $B' \otimes B' \times B \to B'$ defined by $\mathrm{Sand}(b'_1 \otimes b'_2, b) = b'_2 b b'_1$. The *Clifford algebra* $\mathrm{Cl}(V,\theta)$ of the quadratic space (V,θ) is the quotient of the tensor algebra of the F-module B' by the ideal I generated by the sets

$$\begin{array}{rcl} I_1 & = & \{s_1 - \mathrm{Trd}_A(s_1 k)1,\ s_1 \in S_1\} \\ I_2 & = & \{c - \mathrm{Sand}(c,k) \mid \mathrm{Sand}\big(c, \mathrm{Alt}^\varepsilon(V,D)\big) = 0\}. \end{array}$$

The Clifford algebra $\mathrm{Cl}(V,\theta)$ has a canonical involution σ_0 induced by the map τ. We have $\mathrm{Cl}(V,\theta) \otimes_F \widetilde{F} = \mathrm{Cl}(V \otimes_F \widetilde{F}, \theta \otimes 1_{\widetilde{F}})$ for any field extension $\widetilde{F}$ of F and $\mathrm{Cl}(V,q)$ is the even Clifford algebra $C_0(V,q)$ of (V,q) if $D = F$ ([10, Théorème 2]). The reduction is through Morita theory for hermitian spaces (see for example [5, Chapter I, §9] for a description of Morita theory). In [6, §8] the Clifford algebra $C(A,\sigma_\theta,f_\theta)$ of the triple $(A,\sigma_\theta,f_\theta)$ is defined as the quotient of the tensor algebra $T(A)$ of the F-space A by the ideal generated by the sets

$$\begin{array}{rcl} J_1 & = & \{s - \mathrm{Trd}_A(us),\ s \in \mathrm{Sym}(A,\sigma_\theta)\} \\ J_2 & = & \{c - \mathrm{Sand}'(c,u),\ c \in A \text{ with } \mathrm{Sand}'\big(c, \mathrm{Alt}(A,\sigma_\theta)\big) = 0\} \end{array}$$

where $u = \widehat{b_\theta}^{-1} k$ and $\mathrm{Sand}' : (A \otimes A, A) \to A$ is defined as $\mathrm{Sand}'(a \otimes b, x) = axb$. The two definitions give in fact isomorphic algebras:

PROPOSITION 5.1. *The isomorphism $\varphi_\theta : V \otimes_D {}^\sigma V \xrightarrow{\sim} \mathrm{End}_D(V)$ induces an isomorphism $\mathrm{Cl}(V,\theta) \xrightarrow{\sim} C(A,\sigma_\theta,f_\theta)$.*

Proof. We only check that φ_θ maps I_1 to J_1. By definition of τ and S_1, $s = \varphi_\theta(s_1)$ is a symmetric element of A. On the other hand we have by definition of the pairing $B' \times B \to A$,

$$\begin{array}{rcl} \mathrm{Trd}_A(s_1 k) & = & \mathrm{Trd}_A\big(\varphi_\theta(s_1)\psi_\theta^{-1}(k)\big) \\ & = & \mathrm{Trd}_A\big(sh^{-1}\widehat{k}\big) = \mathrm{Trd}_A(su) = \mathrm{Trd}_A(us), \end{array}$$

hence the claim. □

In particular we have $C\big(\mathrm{End}_F(V),\sigma_q,f_q\big) = C_0(V,q)$ for a quadratic space (V,q) over F. It is convenient to use both definitions of the Clifford algebra of a generalized quadratic space.

Let $D = [K,\mu) = K \oplus \ell K$ be a quaternion algebra with conjugation σ. Let V be a D-module and let V^0 be V as a right vector space over K (through restriction of scalars). Let $T : V^0 \to V^0$, $Tx = x\ell$. We have $\mathrm{End}_D(V) \subset \mathrm{End}_K(V^0)$ and

$$\mathrm{End}_D(V) = \{f \in \mathrm{End}_K(V^0) \mid fT = Tf\}.$$

Let $\theta = [k]$ be a $(\sigma,-1)$-quadratic space and let $k(x,y) = P(x,y) + \ell R(x,y)$ as in Section 3. It follows from (3.1) that R defines a quadratic space $[R]$ on V^0 over K.

PROPOSITION 5.2. *We have $\sigma_{[R]}|_{\mathrm{End}_D(V)} = \sigma_\theta$ and $f_\theta = f_{[R]}|_{\mathrm{End}_D(V)}$.*

Proof. We have an embedding $D \hookrightarrow M_2(K)$, $a+\ell b \mapsto \begin{pmatrix} a & \mu\bar{b} \\ b & \bar{a} \end{pmatrix}$ and conjugation given by $x \mapsto x^* = c^{-1}x^t c$, $c = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$. The choice of a basis of V over D identifies V with D^n, V^0 with K^{2n}, $\mathrm{End}_D(V)$ with $M_n(D)$ and $\mathrm{End}_K(V^0)$ with $M_{2n}(K)$, where $n = \dim_D V$. We further identify V and V^* through the choice of the dual basis. We embed any element $x = x_1 + \ell x_2 \in M_{k,l}(D)$, $x_i \in M_{k,l}(K)$ in $M_{2k,2l}(K)$ through the map $\iota : x \mapsto \xi = \begin{pmatrix} x_1 & \mu\overline{x_2} \\ x_2 & \overline{x_1} \end{pmatrix}$. In particular D^n is identified with a subspace of the space of $(2n \times 2)$-matrices over K. Then $D \subset M_2(K)$ operates on the right through (2×2)-matrices and $M_n(D) \subset M_{2n}(K)$ operates on the left through $(2n \times 2n)$-matrices. With the notations of Example (2.3) we have $\iota(x^*) = \mathrm{Int}(c^{-1})(x^t)$. Any D-sesquilinear form k on D^n can be written as $k(x,y) = x^*ay$, where $a \in M_n(D)$, as in (2.3). Let $a = a_1 + \ell a_2$, $a_i \in M_n(K)$ and let

$$\alpha = \iota(a) = \begin{pmatrix} a_1 & \mu\overline{a_2} \\ a_2 & \overline{a_1} \end{pmatrix}.$$

Let $\eta = \iota(y)$, $y = y_1 + \ell y_2$. We have

$$k(x,y) = x^*ay = \xi^*\alpha\eta = \begin{pmatrix} x_1 & \mu\overline{x_2} \\ x_2 & \overline{x_1} \end{pmatrix}^* \begin{pmatrix} a_1 & \mu\overline{a_2} \\ a_2 & \overline{a_1} \end{pmatrix} \begin{pmatrix} y_1 & \mu\overline{y_2} \\ y_2 & \overline{y_1} \end{pmatrix}.$$

On the other side it follows from $h = P + \ell R$ that $R(x,y) = \xi^t\rho\eta$ with

$$\rho = \begin{pmatrix} a_2 & \overline{a_1} \\ -a_1 & -\mu\overline{a_2} \end{pmatrix}.$$

Assume that $\theta = [k]$, so that σ_θ corresponds to the involution $\mathrm{Int}(\gamma^{-1}) \circ *$, where $\gamma = \alpha - \alpha^*$. Similarly $\sigma_{[R]}$ corresponds to the involution $\mathrm{Int}(\widetilde{\rho}^{-1}) \circ t$ where $\widetilde{\rho} = \rho + \rho^t$. We obviously have $\rho = c\alpha$ with $c = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, so that $\rho^t = \alpha^t c^t = -\alpha^t c = -ca^*$ and $\rho + \rho^t = c(\alpha - \alpha^*)$ or $c\gamma = \widetilde{\rho}$. Now $^* = \mathrm{Int}(c^{-1}) \circ t$ implies $\sigma_{[R]}|_{M_n(D)} = \sigma_\theta$. We finally check that $f_\theta = f_{[R]}|_{\mathrm{Sym}(M_n(D),\sigma_\theta)}$. We have $f_\theta(s) = \mathrm{Trd}_{M_n(D)}(\gamma^{-1}\alpha s)$ and $f_{[R]}(s) = \mathrm{Trd}_{M_{2n}(K)}(\widetilde{\rho}^{-1}\rho s)$, hence the claim, since $\rho = c\alpha$ and $\widetilde{\rho} = c\gamma$ implies $\gamma^{-1}\alpha = \widetilde{\rho}^{-1}\rho$. ◻

Corollary 5.3. *The embedding* $\mathrm{End}_D(V) \hookrightarrow \mathrm{End}_K(V^0)$ *induces*
1) an isomorphism $\big(\mathrm{End}_D(V), \sigma_\theta, f_\theta\big) \otimes K \xrightarrow{\sim} \big(\mathrm{End}_K(V^0), \sigma_{[R]}, f_{[R]}\big)$,
2) an isomorphism $C\big(\mathrm{End}_D(V), \sigma_\theta, f_\theta\big) \otimes K \xrightarrow{\sim} C_0(V^0, [R])$.

In view of (2) the semilinear automorphism $T : V^0 \xrightarrow{\sim} V^0$, $Tx = x\ell$, is a semilinear similitude with multiplier $-\mu$ of the quadratic form $[R]$, such that $T^2 = \mu$.

Lemma 5.4. *The map T induces a semilinear automorphism $C_0(T)$ of $C_0(V^0, R)$ such that*

$$C_0(T)(xy) = (-\mu)^{-1}T(x)T(y) \text{ for } x,\ y \in V^0$$

and $C_0(T)^2 = Id$.

Proof. This follows (for example) as in [6, (13.1)] $\square$

Proposition 5.5.

$$C\big(\mathrm{End}_D(V), \sigma_\theta, f_\theta\big) = \{c \in C_0(V^0, R) \mid C_0(T)(c) = c\}.$$

Proof. The claim follows from the defining relations of $C\big(\mathrm{End}_D(V), \sigma_\theta, f_\theta\big)$ and the fact that

$$\mathrm{End}_D(V) = \{f \in \mathrm{End}_K(V^0) \mid T^{-1} f T = f\}.$$

$\square$

We call $C\big(\mathrm{End}_D(V), \sigma_\theta, f_\theta\big)$ or equivalently $\mathrm{Cl}(V, \theta)$ the *Clifford algebra of the quadratic quaternion space* (V, θ).

Let t be a semilinear similitude of a quadratic space (U, q) of even dimension over K. Assume that $\mathrm{disc}(q)$ is trivial, so that $C_0(U, q)$ decomposes as product of two K-algebras $C^+(U, q)$ and $C^-(U, q)$. We say that t is *proper* if $C_0(t)\big(C^\pm(U, q)\big) \subset C^\pm(U, q)$ and we say that t is *improper* if $C_0(t)\big(C^\pm(U, q)\big) \subset C^\mp(U, q)$. In general we say that t is *proper* if t is proper over some field extension of F which trivializes $\mathrm{disc}(q)$. For any semilinear similitude t, let $d(t) = 1$ is t if proper and $d(t) = -1$ if t is improper.

Lemma 5.6. *Let t_i be a semilinear similitude of (U_i, q_i), $i = 1, 2$. We have $d(t_1 \perp t_2) = d(t_1)d(t_2)$.*

Proof. We assume that $\mathrm{disc}(q_i)$, $i = 1, 2$, is trivial. Let e_i be an idempotent generating the center Z_i of $C_0(q_i)$. We have $t_i(e_i) = e_i$ if t_i is proper and $t_i(e_i) = 1 - e_i$ if t_i is improper. The idempotent $e = e_1 + e_2 - 2e_1e_2 \in C_0(q_1 \perp q_2)$ generates the center of $C_0(q_1 \perp q_2)$ (see for example [5, (2.3), Chap. IV]) and the claim follows by case checking. $\square$

Lemma 5.7. *Let V, θ, V^0, R and T be as above. Let $\dim_K V^0 = 2m$. Then T is proper if m is even and is improper if m is odd.*

Proof. The quadratic space (V, θ) is the orthogonal sum of 1-dimensional spaces and we get a corresponding orthogonal decomposition of $\big(V^0, [R]\big)$ into subspaces (U_i, q_i) of dimension 2. In view of (5.6) it suffices to check the case $m = 1$. Let $\alpha = a = a_1 + \ell a_2 \in D$ and $\rho = \begin{pmatrix} a_2 & \overline{a_1} \\ -a_1 & -\mu\overline{a_2} \end{pmatrix}$. We choose $\mu = 1$, $a_1 = j$ (j as in (2.4)), put $i = 1 - 2j$, so that $\bar{i} = -i$ and choose $a_2 = 0$. Let $x = x_1e_1 + x_2e_2 \in V^0$, so $[R](x_1, x_2) = ix_1x_2$ and $C([R])$ is generated by e_1, e_2 with the relations $e_1^2 = 0$, $e_2^2 = 0$, $e_1e_2 + e_2e_1 = i$. The element $e = i^{-1}e_1e_2$ is an idempotent generating the center. Since $T(x_1e_1 + x_2e_2) = \overline{x}_2e_1 + \overline{x}_1e_2$, we have $C_0(T)(e_1e_2) = -e_2e_1$ and $C_0(T)(e) = 1 - e$. Thus T is not proper. $\square$

Of special interest for the next section are quadratic quaternion forms $[k]$ such that the induced quadratic forms $\pi_2([k])$ are Pfister forms. For convenience we call such forms *Pfister quadratic quaternion forms.* Hyperbolic spaces of dimension 2^n are Pfister forms, hence spaces of the form $\beta([b])$, b a hermitian form over K, are Pfister, in view of the exactness of the sequence of Lewis [7]. It is in fact easy to give explicit examples of Pfister forms using the following constructions:

EXAMPLE 5.8 ($\mathrm{Char}\, F \neq 2$). Let $q =< \lambda_1, \dots, \lambda_n >$ be a diagonal quadratic form on F^n, i.e., $q(x) = \sum \lambda_i x_i^2$. Let $[k]$ on D^n be given by the diagonal form ℓq. Then the corresponding quadratic form $[R]$ on K^{2n} is given by the diagonal form $< 1, -\mu > \otimes q$. In particular we get the 3-Pfister form $<< a, b, \mu >>$ choosing for q the norm form of a quaternion algebra $(a, b)_F$.

EXAMPLE 5.9 ($\mathrm{Char}\, F = 2$). Let $b =< \lambda_1, \dots, \lambda_n >$ be a bilinear diagonal form on F^n, i.e., $b(x, y) = \sum \lambda_i x_i y_i$. Let $k = (j + \ell)b$ on D^n. Then the corresponding quadratic form $[R]$ over $K = R(j)$, $j^2 = j + \lambda$, is given by the form $[R] = b \otimes [1, \lambda]$ where $[\xi, \eta] = \xi x_1^2 + x_1 x_2 + \eta x_2{}^2$. In particular, for $b =< 1, a, c, ac >$, we get the 3-Pfister form $<< a, c, \lambda]]$ with the notations of [6], p. xxi.

6. TRIALITY FOR SEMILINEAR SIMILITUDES

Let $\mathfrak{C}$ be a Cayley algebra over F with conjugation $\pi : x \mapsto \overline{x}$ and norm $\mathfrak{n}\colon x \mapsto x\overline{x}$. The new multiplication $x \star y = \overline{x}\,\overline{y}$ satisfies

$$x \star (y \star x) = (x \star y) \star x = \mathfrak{n}(x)y \tag{6}$$

for x, $y \in \mathfrak{C}$. Further, the polar form $b_{\mathfrak{n}}$ is *associative* with respect to $\star$, in the sense that

$$b_{\mathfrak{n}}(x \star y, z) = b_{\mathfrak{n}}(x, y \star z).$$

PROPOSITION 6.1. *For x, $y \in \mathfrak{C}$, let $r_x(y) = y \star x$ and $\ell_x(y) = x \star y$. The map $\mathfrak{C} \to \mathrm{End}_F(\mathfrak{C} \oplus \mathfrak{C})$ given by*

$$x \mapsto \begin{pmatrix} 0 & \ell_x \\ r_x & 0 \end{pmatrix}$$

induces isomorphisms $\alpha\colon \big(C(\mathfrak{C}, \mathfrak{n}), \tau\big) \xrightarrow{\sim} \big(\mathrm{End}_F(\mathfrak{C} \oplus \mathfrak{C}), \sigma_{\mathfrak{n} \perp \mathfrak{n}}\big)$ and

$$\alpha_0\colon \big(C_0(\mathfrak{C}, \mathfrak{n}), \tau_0\big) \xrightarrow{\sim} \big(\mathrm{End}_F(\mathfrak{C}), \sigma_{\mathfrak{n}}\big) \times \big(\mathrm{End}_F(\mathfrak{C}), \sigma_{\mathfrak{n}}\big), \tag{7}$$

of algebras with involution.

Proof. We have $r_x\big(\ell_x(y)\big) = \ell_x\big(r_x(y)\big) = \mathfrak{n}(x) \cdot y$ by (6). Thus the existence of the map α follows from the universal property of the Clifford algebra. The fact that α is compatible with involutions is equivalent to

$$b_{\mathfrak{n}}\big(x \star (z \star y), u\big) = b_{\mathfrak{n}}\big(z, y \star (u \star x)\big)$$

for all x, y, z, u in $\mathfrak{C}$. This formula follows from the associativity of $b_{\mathfrak{n}}$. Since $C(\mathfrak{C},\mathfrak{n})$ is central simple, the map α is an isomorphism by a dimension count. □

Assume from now on that $\mathfrak{C}$ is defined over a field K which is quadratic Galois over F. Any proper semilinear similitude t of $\mathfrak{n}$ induces a semilinear automorphism $C(t)$ of the even Clifford algebra $\big(C_0(\mathfrak{C},\mathfrak{n}),\tau_0\big)$, which does not permute the two components of the center of $C_0(\mathfrak{C},\mathfrak{n})$. Thus $\alpha_0\circ C_0(t)\circ\alpha_0^{-1}$ is a pair of semilinear automorphisms of $\big(\mathrm{End}_K(\mathfrak{C}),\sigma_{\mathfrak{n}}\big)$. It follows as in (4.5) that, for any quadratic space (V,q), semilinear automorphisms of $(\mathrm{End}_K(V),\sigma_q,f_q)$ are of the form $\mathrm{Int}(f)$, where f is a semilinear similitude of q. The following result is due to Wonenburger [12] in characteristic different from 2:

Proposition 6.2. *For any proper semilinear similitude t_1 of $\mathfrak{n}$ with multiplier μ_1, there exist proper semilinear similitudes t_2, t_2 such that*

$$\alpha_0\circ C_0(t_1)\circ\alpha_0^{-1}=\big(\mathrm{Int}(t_2),\mathrm{Int}(t_3)\big)$$

and

$$\begin{aligned}\mu_3^{-1}t_3(x\star y)&=t_1(x)\star t_2(y),\\ \mu_1^{-1}t_1(x\star y)&=t_2(x)\star t_3(y),\\ \mu_2^{-1}t_2(x\star y)&=t_3(x)\star t_1(y).\end{aligned}\tag{8}$$

Let t_1 be an improper similitude with multiplier μ_1. There exist improper similitudes t_2, t_3 such that

$$\begin{aligned}\mu_3^{-1}t_3(x\star y)&=t_1(y)\star t_2(x),\\ \mu_1^{-1}t_1(x\star y)&=t_2(y)\star t_3(x),\\ \mu_2^{-1}t_2(x\star y)&=t_3(y)\star t_1(x).\end{aligned}$$

The pair (t_2,t_3) is determined by t_1 up to a factor (λ,λ^{-1}), $\lambda\in K^\times$, and we have $\mu_1\mu_2\mu_3=1$.
Furthermore, any of the formulas in (8) *implies the two others.*

Proof. The proof given in [6, (35.4)] for similitudes can also be used for semilinear similitudes. □

Remark 6.3. The class of two of the t_i, $i=1$, 2, 3, modulo $K^\times$ is uniquely determined by the class of the third t_i.

Corollary 6.4. *Let T_1 be a proper semilinear similitude of $(\mathfrak{C},\mathfrak{n})$ such that $T_1^2=\mu_1$, $\mu_1\in K^\times$ and with multiplier $-\mu_1$. There exist elements $a_i\in K^\times$, $i=1$, 2, 3, and proper semilinear similitudes T_i of $(\mathfrak{C},\mathfrak{n})$, with $T_i^2=\mu_i$, $\mu_i\in K^\times$ and with multiplier $-\mu_i$, $i=2$, 3, such that $a_i\overline{a_i}\mu_i=\mu_{i+1}\mu_{i+2}$ and*

$$\begin{aligned}a_3T_3(x\star y)&=T_1(x)\star T_2(y)\\ a_1T_1(x\star y)&=T_2(x)\star T_3(y)\\ a_2T_2(x\star y)&=T_3(x)\star T_1(y\end{aligned}$$

The class of any T_i modulo $K^\times$ determines the two other classes and the μ_i's are determined up to norms from $K^\times$. Furthermore any of the three formulas determines the two others.

Proof. Counting indices modulo 3, we have relations

$$T_i(x) \star T_{i+1}(y) = b_{i+2}T_{i+2}, \quad b_i \in K^\times$$

in view of (6.2). If we replace all T_j by $T_j \circ \rho_{\nu_j}$, $\nu_j \in K^\times$, we get new constants a_i. The claim then follows from (3.3). □

7. Triality for quadratic quaternion forms

Let $D_1 = K \oplus \ell_1 K = [K, \mu_1)$ be a quaternion algebra over F and let (V_1, q_{θ_1}) be a quaternion quadratic space of dimension 4 over D_1. Let $\theta_1 = [h_1]$, $h_1(x,y) = P_1(x,y) + \ell R_1(x,y)$, so that $[R_1] = \pi_2(\theta_1)$ corresponds to a 8-dimensional (classical) quadratic form on V_1^0 over K. The map $T_1 : V_1^0 \to V_1^0$, $T_1(x) = x\ell_1$, is a semilinear similitude of $\big(V_1^0, [R_1]\big)$ with multiplier $-\mu_1$ and such that $T_1^2 = \mu_1$. We recall that by (3.5) it is equivalent to have a quadratic quaternion space (V_1, q_{θ_1}) or a pair $\big(V_1^0, [T_1]\big)$. We assume from now on that the quadratic form $q_{[R_1]}$ is a 3-Pfister form, i.e.,the norm form $\mathfrak{n}$ of a Cayley algebra $\mathfrak{C}$ over K. In view of (6.4) T_1 induces two semilinear similitudes T_2, resp. T_3, with multipliers μ_2, resp. μ_3, which in turn define a quaternion quadratic space (V_2, θ_2) of dimension 4 over $D_2 = [K, \mu_2)$, resp. a quaternion quadratic space (V_3, θ_3) of dimension 4 over $D_3 = [K, \mu_3)$. Let $\mathrm{Br}(F)$ be the Brauer group of F.

Proposition 7.1. *1)* $[D_1][D_2][D_3] = 1 \in \mathrm{Br}(F)$,
2) The restriction of $\alpha : C_0\big(\mathfrak{C}, \mathfrak{n}\big) \xrightarrow{\sim} \mathrm{End}_K(\mathfrak{C}) \times \mathrm{End}_K(\mathfrak{C})$ *to* $C(V_i, D_i, \theta_i)$ *induces isomorphisms*

$$\alpha_i : \big(C(V_i, D_i, \theta_i), \tau\big) \xrightarrow{\sim} \big(\mathrm{End}_{D_{i+1}}(V_{i+1}), \sigma_{\theta_{i+1}}\big) \times \big(\mathrm{End}_{D_{i+2}}(V_{i+2}), \sigma_{\theta_{i+2}}\big)$$

Proof. The first claim follows from the fact that $\mu_1\mu_2 = \mu_3 \operatorname{Nrd}_{D_3}(a_3)$ and the second is a consequence of (5.5), (3.5) and the definition of α. □

Example 7.2. Let $\mathfrak{C}_0$ be a Cayley algebra over F and let $\mathfrak{C} = \mathfrak{C}_0 \otimes_F K$. For any $c \in \mathfrak{C}_0$ such that $c^2 = \mu_1 \in F^\times$, $T_1 : \mathfrak{C} \to \mathfrak{C}$ given by $T_1(k \otimes x) = \overline{k} \otimes xc$ is a semilinear similitude with multiplier $-\mu_1$ such that $T_1^2 = \mu_1$. The Moufang identity $(cx)(yc) = c(xy)c$ in $\mathfrak{C}$ implies that

$$(xc) \star (cy) = \overline{c}(x \star y)\overline{c}.$$

Thus $T_2(k \otimes y) = \overline{k} \otimes cy$ and $T_3(k \otimes z) = i\overline{k} \otimes \overline{c}z\overline{c}$ (where $i \in K^\times$ is such that $\overline{i} = -i$) satisfy (6.4). The corresponding triple of quaternion algebras is $\big([K, \mu_1), [K, \mu_1), [K, i\overline{i}\mu_1^2)\big)$, the third algebra being split.

Example 7.3. Let D_i, $i = 1, , 2, 3$, be quaternion algebras over F such that $[D_1][D_2][D_3] = 1 \in \mathrm{Br}(F)$. We may assume that the D_i contain a common separable quadratic field K and that $D_i = [K, \mu_i)$, $\mu_i \in F^\times$ such that $\mu_1\mu_2\mu_3 \in$

$F^{\times 2}$. In [6, (43.12)] similitudes S_i with multiplier μ_i, $i = 1, 2, 3$, of the split Cayley algebra $\mathfrak{C}_s$ over F are given, such that 1) $\mu_3^{-1} S_3(x \star y) = S_1(x) \star S_2(y)$ and 2) $S_i^2 = \mu_i$. Let $\mathfrak{C} = K \otimes \mathfrak{C}_s$. Let $u \in K^\times$ be such that $\overline{u} = -u$. The semilinear similitudes $T_i(k \otimes x) = u\overline{k} \otimes S_i(x)$, $i = 1, 2, 3$, satisfy

$$a_3 T_3(x \star y) = T_1(x) \star T_2(y)$$

with $a_3 = u\mu_3^{-1}$ (we use the same notation $\star$ in $\mathfrak{C}_s$ and in $\mathfrak{C}$). Thus there exist a triple of quadratic quaternion forms $(\theta_1, \theta_2, \theta_3)$ corresponding to the three given quaternion algebras. We hope to describe the corresponding quadratic quaternion forms in a subsequent paper.

References

[1] A. Bak. *K-Theory of forms*, volume 98 of *Annals of Mathematics Studies.* Princeton University Press, Princeton, N.J., 1981.

[2] E. Bayer-Fluckiger and R. Parimala. Galois cohomology of the classical groups over fields of cohomological dimension $\leq$ 2. *Invent. Math.*, 122(2):195–229, 1995.

[3] J. Dieudonné. Sur les groupes unitaires quaternioniques à deux ou trois variables. *Bull. Sci. Math.*, 77:195–213, 1953.

[4] M. A. Elomary. *Orthogonal sum of central simple algebras with quadratic pairs in characteristic* 2 *and classification theorems.* PhD thesis, Université Catholique de Louvain, 2000.

[5] M.-A. Knus. *Quadratic and Hermitian forms over rings*, volume 294 of *Grundlehren der Mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1991. With a foreword by I. Bertuccioni.

[6] M.-A. Knus, A. A Merkurjev, M. Rost and J.-P. Tignol. *The Book of Involutions.* Number 44 in American Mathematical Society Colloquium Publications. American Mathematical Society, Providence, R.I., 1998.

[7] D. W. Lewis. New improved exact sequences of Witt groups. *J. Algebra*, 74:206–210, 1982.

[8] W. Scharlau. *Quadratic and Hermitian forms*, volume 270 of *Grundlehren der mathematischen Wissenschaften.* Springer-Verlag, Berlin, 1985.

[9] E. A. M. Seip-Hornix. Clifford algebras of quadratic quaternion forms. I, II. *Nederl. Akad. Wetensch. Proc. Ser. A 68 = Indag. Math.*, 27:326–363, 1965.

[10] J. Tits. Formes quadratiques, groupes orthogonaux et algèbres de Clifford. *Invent. Math.*, 5:19–41, 1968.

[11] C. T. C Wall. On the axiomatic foundations of the theory of Hermitian forms. *Proc. Camb. Phil. Soc.*, 67:243–250, 1970.

[12] M. J. Wonenburger. Triality principle for semisimilarities. *J. Algebra*, 1:335–341, 1964.

Max-Albert Knus
ETH Zentrum
CH-8092-Zürich
Switzerland
knus@math.ethz.ch

Oliver Villa
ETH Zentrum
CH-8092-Zürich
Switzerland

Certaines Combinaisons Linéaires de Deux Formes de Pfister et le Problème d'Isotropie

Ahmed Laghribi

Received: May 30, 2001
Revised: September 28, 2001

Communicated by Ulf Rehmann

Abstract. In this paper we treat the isotropy problem of certain linear combinations of two Pfister forms over the function field of a projective quadric. More precisely, we discuss the connection between this problem and some conjectures on a field that appears in the generic splitting tower of a quadratic form associated to a given linear combination. The case of some linear combinations of low dimension will be detailed.

2000 Mathematics Subject Classification: 11E04, 11E81.
Keywords and Phrases: Forme quadratique, Forme de Pfister, Quadrique projective, Déploiement générique d'une forme quadratique, Cohomologie galoisienne.

1. Introduction

Soit F un corps commutatif de caractéristique différente de 2. Dans ce papier on s'intéresse au problème suivant:

Problème 1.1. *Pour φ une F-forme quadratique anisotrope de dimension ≥ 2, quelles sont les F-formes quadratiques ψ pour lesquelles φ devient isotrope sur $F(\psi)$ le corps des fonctions de la quadrique projective d'équation $\psi = 0$?*

Une n-forme de Pfister est une forme de type $\langle 1, -a_1\rangle \otimes \cdots \otimes \langle 1, -a_n\rangle$ avec $a_1, \cdots, a_n \in F^*$, qu'on note $\langle\langle a_1, \cdots, a_n\rangle\rangle$. Une 0-forme de Pfister est une forme de dimension 1. Fixons quelques notations:

Notations 1.2. *Pour $n, m \geq 0$ deux entiers, on note:*

1. $P_n(F)$ *l'ensemble des n-formes de Pfister et* $GP_n(F) = F^*P_n(F)$.
2. $(P_n(F))' = \{\pi' \mid \langle 1\rangle \perp \pi' \in P_n(F)\}$ *et* $(GP_n(F))' = F^*(P_n(F))'$.
3. $L_{n,m}(F) = \{\pi \perp \tau \mid \pi \in GP_n(F),\ \tau \in GP_m(F)\}$.
4. $(L_{n,m}(F))' = \{\pi \perp \tau \mid \pi \in (GP_n(F))',\ \tau \in (GP_m(F))'\}$.

Une forme quadratique ψ est dite une *sous-forme* de φ et on note $\psi \subset \varphi$ s'il existe une forme quadratique ξ telle que $\varphi \cong \psi \perp \xi$ où $\cong$ et $\perp$ désignent respectivement l'isométrie et la somme orthogonale des formes quadratiques.

Le but de ce papier est d'étudier le problème précédent lorsque $\varphi \in L_{n,m}(F)$ ou $(L_{n,m}(F))'$. En général, il est très difficile de faire cela de façon complète. En ce qui concerne les formes de $L_{n,m}(F)$, notre stratégie consiste à associer à une forme $\varphi \in L_{n,m}(F)$ un corps F_ϵ qui apparaît dans la tour de déploiement générique d'une forme liée à φ. Le corps F_ϵ est indépendant de l'écriture de φ et que la forme φ_{F_ϵ} devient une sous-forme d'une forme de $GP_{n+1}(F_\epsilon)$ (propositions 2.2, 2.7). On conjecture que φ_{F_ϵ} est anisotrope et on fait le lien avec d'autres conjectures (propositions 2.12, 2.16). On va discuter de manière générale ces conjectures et ce en répondant au problème 1.1 dès que ces conjectures sont vérifiées (théorème 2.19). Entre autre, on étudie de manière réelle l'isotropie d'une forme de $L_{n,1}(F)$ et d'une forme de $L_{n,m}(F)$ qui est divisible par une $(m-1)$-forme de Pfister, et cette étude est plus détaillée lorsque $n=3$ et $m=2$. Pour ce qui est de l'isotropie d'une forme de $(L_{n,m}(F))'$, l'étude se ramène souvent à celle d'une forme de $L_{n,m}(F)$ qui la contient (proposition 3.2). Mais, en général, l'isotropie des formes de $(L_{n,m}(F))'$ reste plus compliquée que celle des formes de $L_{n,m}(F)$. On se limite à étudier l'isotropie de certaines formes de $(L_{n,m}(F))'$ avec $n \geq 3$ et $m = 1, 2$.

Rappelons que dans la proposition 2.16 et les théorèmes 2.19, 4.1 on suppose dans certains cas que F est de caractéristique 0. Cela est dû au fait qu'on se base sur des résultats d'Orlov-Vishik-Voevodsky [31] qui sont établis en cette caractéristique. Plus précisément, par [38] et [31, Theorem 2.1] on déduit qu'on a le résultat suivant:

(R1) *Pour $m \geq n \geq 1$ et $\pi \in GP_nF$, on a* $\mathrm{Ker}(H^mF \longrightarrow H^mF(\pi)) = e^n(\pi) \cdot H^{m-n}F$

où e^n est le n-ième invariant d'Arason et H^nF est le n-ième groupe de cohomologie galoisienne à coefficients dans $\mathbb{Z}/2$ (voir la section 4 pour plus de détails sur l'invariant e^n). Aussi par [38] et [31, Theorem 2.10] on a un autre résultat:

(R2) *Pour $h \in H^nF$ non nul, il existe K/F une extension telle que h_K soit un symbole non nul*

où un symbole désigne un élément de H^nF de type $(a_1) \cdot \ldots \cdot (a_n)$ avec (a_i) est la classe de $a_i \in F^*$ dans H^1F et $\cdot$ est le cup-produit. Comme conséquence du résultat (R2) on obtient:

(R3) *Pour φ de dimension $> 2^n$, on a* $\mathrm{Ker}(H^nF \longrightarrow H^nF(\varphi)) = \{0\}$.

Aussi on mentionne un autre résultat important [18, fin de la page 166], [30]:

(R4) *Pour $n \geq 0$, e^n induit un isomorphisme entre $I^nF/I^{n+1}F$ et H^nF.*

Lorsque la caractéristique n'est pas nécessairement 0, les résultats (R1), (R3) et (R4) sont vrais comme suit:

1. Le résultat (R1) est vrai pour $m \leq 4$: Arason [1] pour $m = 2, 3$; Kahn-Rost-Sujatha [17, Corollary 2] pour $m = 4$ et $n = 3$; Kahn-Sujatha [19, Theorem 2] pour $m = 4$ et $n = 2$.
2. Le résultat (R3) est vrai pour $n \leq 4$: Arason [1] pour $n \leq 3$; Kahn-Rost-Sujatha [17, Corollary 2] pour $n = 4$.
3. Le résultat (R4) est vrai pour $n \leq 4$: Evident pour $n = 0$; Par la théorie de Kummer pour $n = 1$; Merkur'ev [28] pour $n = 2$; Merkur'ev-Suslin/Rost [29], [33] pour $n = 3$; Rost (non publié) et Szyjewski [37] pour $n = 4$.

On dit qu'une forme φ est *voisine* s'il existe une n-forme de Pfister π telle que $\dim \varphi > 2^{n-1}$ et $a\pi \cong \varphi \perp \xi$ pour certains $a \in F^*$ et ξ une forme quadratique. Dans ce cas, les formes π et ξ sont uniques, et pour toute extension de corps K/F on a que φ_K est isotrope si et seulement si π_K l'est aussi. La forme ξ est appelée la forme *complémentaire* de φ.

Si φ est voisine de $\pi \in P_nF$, en particulier si $\varphi \in L_{n,0}(F)$, alors par le théorème de la sous-forme (théorème 1.4) on répond au problème 1.1 de façon complète:

$$\begin{aligned} \varphi_{F(\psi)} \text{ est isotrope } &\Leftrightarrow \pi_{F(\psi)} \text{ est isotrope} \\ &\Leftrightarrow a\psi \subset \pi \text{ pour un certain } a \in F^* \end{aligned} \tag{1}$$

Ainsi pour la suite de ce papier et dans le cas des formes de $L_{n,m}(F)$, on va considérer uniquement celles de $L_{n,m}(F)$ avec $m \geq 1$.

L'isotropie d'une forme de $L_{1,1}(F)$ a été étudiée par Leep [27] et Shapiro [35]; l'isotropie d'une forme de $L_{2,1}(F)$ a été étudiée par Hoffmann [7] et Izhboldin-Karpenko [13]; l'isotropie d'une forme de $L_{2,2}(F)$ a été étudiée par l'auteur [22], [23].

Plus généralement, le problème précédent a été aussi étudié par Hoffmann pour une forme de dimension 5 [6]; par Leep [27] et Merkur'ev [26] pour une forme d'Albert (c'est-à-dire une forme de dimension 6 et de discriminant à signe -1); par l'auteur [24] et Izhboldin-Karpenko [12] pour des formes de dimension 6 qui ne sont pas nécessairement dans $L_{2,1}(F)$; par l'auteur pour une forme de dimension 8 et de discriminant à signe 1 mais qui n'est pas nécessairement dans $L_{2,2}(F)$, et pour certaines formes de dimension 7.

Si K/F est une extension de corps, alors on notera $W(K/F)$ le noyau de l'homomorphisme $W(F) \longrightarrow W(K)$ induit par l'inclusion $F \subset K$. Pour deux formes φ_1 et φ_2, on note $\varphi_1 \sim \varphi_2$ si $\varphi_1 \perp -\varphi_2$ est hyperbolique. On dit que φ_1 et φ_2 sont semblables si $\varphi_1 \cong a\varphi_2$ pour un certain scalaire $a \in F^*$. La partie anisotrope φ_{an} d'une forme quadratique φ est l'unique forme quadratique anisotrope telle que $\varphi \sim \varphi_{\mathrm{an}}$. On dit que φ est divisible par ψ si on a $\varphi \cong \psi \otimes \rho$ pour une certaine forme quadratique ρ.

On désigne par $C(\varphi)$ (resp. $C_0(\varphi)$) l'algèbre de Clifford de φ (resp. l'algèbre de Clifford paire de φ). L'invariant de Clifford de φ est désigné par $c(\varphi)$. On note $D_F(\varphi) = \{a \in F^* | \exists x \in V, \varphi(x) = a\}$ où V est l'espace vectoriel sous-jacent à

φ, et $G_F(\varphi) = \{a \in F^* | a\varphi \cong \varphi\}$. On rappelle que $D_F(\pi) = G_F(\pi)$ pour toute forme de Pfister π et on dit que dans ce cas que π est *multiplicative*.

Pour A une F-algèbre simple centrale de dimension finie, on désigne par ind A l'indice de Schur de A.

Les deux théorèmes suivants seront utilisés de manière fréquente. On y fera référence respectivement par les noms *"Hauptsatz"* et *"le théorème de la sous-forme"*.

THÉORÈME 1.3. *(Arason-Pfister) Si $\varphi \in I^nF$ anisotrope, alors* $\dim \varphi \geq 2^n$.

THÉORÈME 1.4. *(Cassels-Pfister) Soient φ et ψ deux formes quadratiques anisotropes telles que $1 \in D_F(\psi)$ et que $\varphi_{F(\psi)}$ soit hyperbolique. Alors, pour tout $\alpha \in D_F(\varphi)$ on a $\alpha\psi \subset \varphi$. En particulier,* $\dim \varphi \geq \dim \psi$.

2. LES FORMES QUADRATIQUES DE $L_{n,m}(F)$

Le long de cette section on va fixer les notations suivantes:

$$(*) \qquad \begin{cases} \pi \in P_n(F),\, \tau \in P_m(F) & \text{avec } n \geq m \geq 1 \\ \varphi = a\pi \perp b\tau \in L_{n,m}(F) & \text{avec } a, b \in F^* \\ \eta = \pi \perp -\tau & \\ \pi_0 = \pi \perp ab\pi \in P_{n+1}(F). & \end{cases}$$

Faisons remarquer qu'avec la multiplicativité d'une forme de Pfister, on déduit que si φ est anisotrope alors π_0 est aussi anisotrope.

2.1. QUELQUES RÉSULTATS PRÉLIMINAIRES. Par la théorie générique de Knebusch [20], [21], on associe à une forme quadratique φ non nulle une suite de formes quadratiques et d'extensions de F, appelée *la tour de déploiement générique* de φ, de la manière suivante:

$$F_0 = F, \qquad \varphi_0 = \varphi_{\mathrm{an}}$$

et pour $n \geq 1$, on définit par récurrence

$$F_n = F_{n-1}(\varphi_{n-1}) \quad \text{et} \quad \varphi_n = ((\varphi_{n-1})_{F_n})_{\mathrm{an}}.$$

La hauteur de φ, notée $\mathrm{h}(\varphi)$, est le plus petit entier h tel que $\dim \varphi_h \leq 1$. Pour $j \in \{0, \cdots, h\}$, on note $i_j(\varphi)$ l'indice de Witt de φ_{F_j}. On a $0 \leq i_0(\varphi) < \cdots < i_h(\varphi)$. On appelle $(i_0(\varphi), \cdots, i_h(\varphi))$ *la suite des indices* de déploiement de φ (splitting patterns [10]), et $(\varphi_0, \cdots, \varphi_h)$ (resp. $(F_0, \cdots, F_h)$) *la suite des noyaux* (resp. *la suite des extensions*) de la tour de déploiement générique de φ. Si $\dim \varphi$ est paire, alors φ_{h-1} est semblable à une forme de Pfister $\rho \in P_dF_{h-1}$ qu'on appelle *la forme dominante* de φ (Knebusch [20, Theorem 5.8] et Wadsworth [39]). L'entier d s'appelle le degré de φ, qu'on note $\deg(\varphi)$. Lorsque $\dim \varphi$ est impaire, on dit que φ est de degré 0. Le corps F_h s'appelle *le corps de déploiement générique* de φ.

On commence par rappeler un résultat.

Proposition 2.1. *(Elman-Lam* [3, 4.5]*; Kahn* [15, Remarque, Page 61]*) Soient $\pi \in P_nF$, $\tau \in P_mF$ anisotropes et $a, b \in F^*$. Soit i le plus grand entier tel que π et τ soient divisibles par une i-forme de Pfister. Alors, $i_W(a\pi \perp b\tau) = 0$ ou 2^i. De plus, si $i_W(a\pi \perp b\tau) = 2^i$ alors il existe $\rho \in P_iF$, $\mu \in P_{n-i}F$ et $\nu \in P_{m-i}F$ telles que $\pi \cong \rho \otimes \mu$ et $\tau \cong \rho \otimes \nu$.*

La proposition suivante est liée au déploiement générique de η.

Proposition 2.2. *On garde les mêmes notations que dans $(*)$. Soient $(F_i)_{0\leq i\leq \mathrm{h}(\eta)}$ (resp. $(i_j(\eta))_{0\leq j\leq \mathrm{h}(\eta)}$) la suite des extensions de la tour de déploiement générique de η (resp. la suite des indices de déploiement de η). Alors:*
(1) Il existe $\epsilon \in \{0, \cdots, \mathrm{h}(\eta)\}$ tel que $i_\epsilon(\eta) = 2^m$.
(2) Pour $\epsilon \in \{0, \cdots, \mathrm{h}(\eta)\}$ comme dans l'assertion (1), on a:
(i) $a\varphi_{F_\epsilon} \subset (\pi_0)_{F_\epsilon}$,
(ii) Si $i_W(\eta) = 2^{m-1}$, alors $\epsilon = 1$ c'est-à-dire $F_\epsilon = F(\eta_{an})$,
(iii) Si $i_W(\eta) = 2^m$, alors $\epsilon = 0$ c'est-à-dire $F_\epsilon = F$.
(iv) Si $n > m$, alors l'extension $F_\epsilon(\pi)/F(\pi)$ est transcendante pure.

Démonstration. (1) Voir [11, Theorem 2.8].

(2)(i) Puisque $i_\epsilon(\eta) = 2^m$, on déduit que $\tau_{F_\epsilon} \subset \pi_{F_\epsilon}$. Ainsi, $a\varphi_{F_\epsilon} \subset (\pi_0)_{F_\epsilon}$.
(ii) Il existe $\rho \in P_{m-1}F$, $\mu = \mu' \perp \langle 1\rangle \in P_{n-m+1}F$ et $d \in F^*$ tels que $\pi \cong \rho\otimes\mu$ et $\tau \cong \rho \otimes \langle 1, -d\rangle$. On a $\eta_{\mathrm{an}} = \rho \otimes (\mu' \perp \langle d\rangle)$. Comme $i_W(\eta_{F(\eta_{\mathrm{an}})})$ est une puissance de 2 (proposition 2.1) strictement supérieure à 2^{m-1}, on a $i_1(\eta) = 2^m$. Ainsi, $\epsilon = 1$ c'est-à-dire $F_\epsilon = F(\eta_{\mathrm{an}})$.
(iii) Evident.
(iv) On a $\eta_{F(\pi)} \sim (-\tau)_{F(\pi)}$. Par le théorème de la sous-forme, $\tau_{F(\pi)}$ est anisotrope, et donc $(\eta_{F(\pi)})_{\mathrm{an}} \cong (-\tau)_{F(\pi)}$. Ainsi, $i_W(\eta_{F(\pi)}) = 2^{n-1} \geq 2^m = i_\epsilon(\eta)$. D'après [20, Remark 5.5] on a que $F_\epsilon(\pi)/F(\pi)$ est transcendante pure.

Remarque 2.3. *Avec les notations de la proposition 2.2 et lorsque $n = m$, le corps F_ϵ n'est autre que le corps de déploiement générique de η c'est-à-dire $\epsilon = \mathrm{h}(\eta)$.*

Définition 2.4. *(*[21, Definition 7.7]*) Toute forme de dimension ≤ 1 est dite excellente. Une forme φ de dimension ≥ 2 est dite excellente si elle est voisine et sa forme complémentaire est excellente.*

La condition $i_W(\eta) = 2^m$ peut être vue autrement:

Proposition 2.5. *On garde les mêmes notations que dans $(*)$ et on suppose que φ est anisotrope. On a équivalence entre:*
(1) φ est une voisine d'une $(n+1)$-forme de Pfister,
(2) φ est divisible par une m-forme de Pfister,
(3) φ est excellente.
(4) $\tau \subset \pi$.

Démonstration. Les implications (3) $\Longrightarrow$ (1) et (4) $\Longrightarrow$ (2) sont évidentes.

(1) $\Longrightarrow$ (4) Puisque $a\pi \perp \langle b\rangle$ est une voisine de π_0 contenue dans φ, on déduit que φ est une voisine de π_0. Par multiplicativité $a\varphi \subset \pi_0$. Ainsi, $\tau \subset \pi$.
(2) $\Longrightarrow$ (3) Soit $\rho \in P_m(F)$ divisant φ. Alors, $a\pi_{F(\rho)} \sim -b\tau_{F(\rho)}$ puisque $\varphi_{F(\rho)} \sim 0$. Si $n > m$ on obtient $\pi_{F(\rho)} \sim \tau_{F(\rho)} \sim 0$. Ainsi, $\rho \cong \tau$ et $\pi_{F(\tau)} \sim 0$. Soit $\lambda \in P_{n-m}(F)$ tel que $\pi \cong \tau \otimes \lambda$. On a $\varphi \cong \tau \otimes (a\lambda \perp \langle b\rangle)$ qui est bien une forme excellente. Si $n = m$ on obtient que $\varphi \in GP_{n+1}(F)$ qui est aussi excellente.

DÉFINITION 2.6. *(1) Deux corps K et L contenant F sont dits F-équivalents (au sens de Knebusch) s'il existe une F-place de l'un vers l'autre et inversement.*
(2) Deux suites croissantes de corps $(F_0 = F, \ldots, F_r)$ et $(G_0 = F, \ldots, G_s)$ sont dites F-équivalentes si:
(i) $r = s$,
(ii) Pour tout $i \in \{0, \ldots, r\}$ les corps F_i et G_i sont F-équivalents.

La proposition suivante sera démontrée au début de la section 5.

PROPOSITION 2.7. *On garde les mêmes notations que dans $(*)$. Soient $\zeta \in P_nF$, $\sigma \in P_mF$ et $c, d \in F^*$ de sorte que*

$$\varphi \cong c\zeta \perp d\sigma$$

soit une autre écriture de φ. Soient $\delta = \zeta \perp -\sigma$ et $(F_i, \eta_i)_{0\le i\le \mathrm{h}(\eta)}$ (resp. $(G_i, \delta_i)_{0\le i\le \mathrm{h}(\delta)}$) la tour de déploiement générique de η (resp. la tour de déploiement générique de δ). Soit $\epsilon \in \{0, \cdots, \mathrm{h}(\eta)\}$ (resp. $\nu \in \{0, \cdots, \mathrm{h}(\delta)\}$) tel que $i_W(\eta_{F_\epsilon}) = 2^m$ (resp. $i_W(\delta_{G_\nu}) = 2^m$).
(1) Si $n > m$, alors les suites $(F_0, \ldots, F_\epsilon)$ et $(G_0, \ldots, G_\nu)$ sont F-équivalentes. En particulier, $\epsilon = \nu$ et les corps F_ϵ et G_ν sont F-équivalents.
(2) Si $n = m$, alors les corps F_ϵ et G_ν sont aussi F-équivalents.
Ainsi, à F-équivalence près, le corps F_ϵ ne dépend pas de l'écriture de φ.

DÉFINITION 2.8. *Avec les mêmes notations et hypothèses que dans la proposition 2.2, on appelle F_ϵ le corps de voisinage de φ.*

2.2. ISOTROPIE DES FORMES QUADRATIQUES DE $L_{n,m}(F)$. Soient φ comme dans $(*)$ et F_ϵ son corps de voisinage. Comme on va le voir l'isotropie de φ est liée à la question de savoir si φ_{F_ϵ} reste anisotrope lorsque φ est anisotrope. Sur cette question on pose la conjecture suivante:

CONJECTURE 2.9. *On garde les mêmes notations que dans $(*)$ et on suppose que φ est anisotrope. Soit F_ϵ le corps de voisinage de φ. Alors, $(\pi_0)_{F_\epsilon}$ est anisotrope. En particulier, $\pi_0 \notin W(F_\epsilon/F)$.*

De manière équivalente, la conjecture dit que φ_{F_ϵ} est anisotrope du fait que $a\varphi_{F_\epsilon} \subset (\pi_0)_{F_\epsilon}$ et $2\dim\varphi > \dim\pi_0$.

Plus généralement sur l'ensemble $P_{n+1}F \cap W(F_\epsilon/F)$ on pose la conjecture suivante:

Conjecture 2.10. *Avec les mêmes hypothèses que dans la conjecture 2.9 et pour* $\rho \in P_{n+1}(F)$, *on a:*

$$\rho \in W(F_\epsilon/F) \Leftrightarrow \begin{cases} \rho \cong \pi \perp r\pi \text{ avec } r \in D_F(-\tau) & si\ n > m \\ \rho \perp -(\eta \perp \alpha\eta) \in I^{n+2}F \text{ pour } \alpha \in F^* & si\ n = m. \end{cases}$$

Dans la proposition qui suit on mentionne quelques inclusions qui sont toujours vraies dans la conjecture 2.10.

Proposition 2.11. *Avec les mêmes notations que dans* $(*)$, *on a:*
(1) $\{\pi \perp r\pi \mid r \in D_F(-\tau)\} \subset P_{n+1}F \cap W(F_\epsilon/F)$.
(2) Si $n > m$, *alors* $P_{n+1}F \cap W(F_\epsilon/F) \subset \{\pi \perp r\pi \mid r \in F^*\}$.
(3) Si $n = m$, *alors*
$\{\rho \in P_{n+1}F \mid \rho \perp -(\eta \perp \alpha\eta) \in I^{n+2}F,\ \alpha \in F^*\} \subset P_{n+1}F \cap W(F_\epsilon/F)$.

Démonstration. (1) Si $r \in D_F(-\tau)$, alors on a:

$$\begin{aligned} \eta \perp r\eta &\sim \pi \perp -\tau \perp -r\tau \perp r\pi \\ &\sim \pi \perp -\tau \perp \tau \perp r\pi \\ &\sim \pi \perp r\pi \in I^{n+1}F. \end{aligned}$$

Puisque $\dim(\eta_{F_\epsilon})_{\mathrm{an}} = 2^n - 2^m$, on a $\dim((\eta \perp r\eta)_{F_\epsilon})_{\mathrm{an}} \leq 2^{n+1} - 2^{m+1}$. Par le Hauptsatz , on déduit que $\pi \perp r\pi \in W(F_\epsilon/F)$.
(2) Si $\rho \in P_{n+1}F \cap W(F_\epsilon/F)$, alors $\rho_{F_\epsilon(\pi)} \sim 0$. Par la proposition 2.2(iv) on a que $F_\epsilon(\pi)/F(\pi)$ est transcendante pure, et donc $\rho_{F(\pi)} \sim 0$. D'où le résultat.
(3) C'est une conséquence du Hauptsatz et du fait que $\eta_{F_\epsilon} \sim 0$ (remarque 2.3).

Proposition 2.12. *La conjecture 2.10 implique la conjecture 2.9.*

Voici quelques cas où la conjecture 2.10 est vérifiée:

Proposition 2.13. *La conjecture 2.10 est vraie si* $i_W(\eta) \in \{2^m, 2^{m-1}\}$.

Comme un corollaire immédiat on a:

Corollaire 2.14. *La conjecture 2.10 est vraie pour* $\varphi \in L_{n,1}(F)$.

En caractéristique 0 lorsque $n \geq 4$ et avec le résultat (R4) la conjecture 2.9 dit de manière équivalente que $e^{n+1}(\pi_0) \notin H^{n+1}(F_\epsilon/F)$. Plus généralement, on pose une conjecture sur le noyau $H^{n+1}(F_\epsilon/F)$:

Conjecture 2.15. *Avec les mêmes hypothèses que dans la conjecture 2.9, on a:*

$$H^{n+1}(F_\epsilon/F) = \begin{cases} \{e^n(\pi) \cdot (r) \mid r \in D_F(\tau)\} & si\ n > m \\ e^n(\eta) \cdot H^1F & si\ n = m. \end{cases}$$

Entre les conjectures 2.10 et 2.15 on a les liens suivants:

Proposition 2.16. *On suppose que* F *est de caractéristique* 0 *lorsque* $n \geq 4$. *On a:*
(1) Si $n > m$, *alors les conjectures 2.10 et 2.15 sont équivalentes.*
(2) Si $n = m$, *alors la conjecture 2.15 implique la conjecture 2.10.*

THÉORÈME 2.17. *La conjecture 2.15 est vraie dans les cas suivants:*
(1) $n = m \leq 2$.
(2) $i_W(\eta) \in \{2^{m-1}, 2^m\}$ *en supposant que* F *est de caractéristique* 0 *lorsque* $(n > m$ *et* $n \geq 4)$ *ou* $(n = m \geq 4$ *et* $i_W(\eta) = 2^{m-1})$.

Démonstration. (1) La conjecture a été prouvée dans [1] lorsque $(n = m = 1)$ et $(n = m = 2$ avec $i_W(\eta) = 2)$; et dans [32] lorsque $n = m = 2$ avec $i_W(\eta) = 1$ (en fait dans ce dernier cas la conjecture se déduit de [32] comme cela est fait dans [22, Corollaire 6]).
(2) (i) Si $n > m$ et $i_W(\eta) \in \{2^m, 2^{m-1}\}$, alors la conjecture est une conséquence de la proposition 2.16(1) et la proposition 2.13.
(ii) Si $n = m$ et $i_W(\eta) = 2^m$, alors la conjecture est évidente car $\eta \sim 0$ et donc $F_\epsilon = F$.
(iii) Si $n = m$ et $i_W(\eta) = 2^{m-1}$, alors $\eta_{\mathrm{an}} \in GP_nF$ et par la proposition 2.2 $F_\epsilon = F(\eta_{\mathrm{an}})$. La conjecture est une conséquence du résultat (R1).

On combine les propositions 2.13, 2.16 et le théorème 2.17 pour obtenir:

COROLLAIRE 2.18. *La conjecture 2.10 est vraie dans les cas suivants:*
(1) $n = m \leq 2$;
(2) $i_W(\eta) \in \{2^{m-1}, 2^m\}$ *en supposant que* F *est de caractéristique* 0 *lorsque* $(n > m$ *et* $n \geq 4)$ *ou* $(n = m \geq 4$ *et* $i_W(\eta) = 2^{m-1})$.

Maintenant on énonce nos principaux résultats sur l'isotropie d'une forme quadratique de $L_{n,m}(F)$. Dans le théorème suivant et lorsque $n = m \geq 4$ on suppose que F est de caractéristique 0.

THÉORÈME 2.19. *Soit* φ *comme dans* $(*)$ *et qu'on suppose anisotrope et soit* ψ *une forme quadratique de dimension* $\geq 2^{n+1}$. *On suppose que la conjecture 2.10 est vraie pour* φ *lorsque* $n > m$, *et que la conjecture 2.15 est vraie pour* φ *lorsque* $n = m$. *On a:*
(1) Si $\dim \psi > 2^{n+1}$, *alors* $\varphi_{F(\psi)}$ *est anisotrope.*
(2) Si $n > m$ *et* $\dim \psi = 2^{n+1}$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* $(n+1)$*-forme de Pfister dont* φ *contient une voisine.*
(3) Si $n = m$, $\dim \psi = 2^{n+1}$ *avec* F *de caractéristique* 0 *lorsque* $n \geq 4$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* $(n+1)$*-forme de Pfister dont* φ *contient une voisine, ou* $\varphi \perp \alpha\psi \in I^{n+2}F$ *pour un certain* $\alpha \in F^*$.

Le corollaire suivant se déduit du théorème 2.19 et du corollaire 2.14.

COROLLAIRE 2.20. *Soient* $n \geq 2$ *un entier et* $\varphi \in L_{n,1}(F)$ *anisotrope. Soit* ψ *une forme quadratique de dimension* 2^{n+1}. *Alors, on a équivalence entre:*
(1) $\varphi_{F(\psi)}$ *est isotrope;*
(2) ψ *est voisine d'une* $(n+1)$*-forme de Pfister dont* φ *contient une voisine.*

Comme dans le corollaire 2.20 et lorsque $n = 3$, on obtient:

Corollaire 2.21. *Soit* $\varphi \in L_{3,1}(F)$ *anisotrope et* ψ *une forme quadratique telle que* $11 \leq \dim \psi \leq 16$. *On suppose que* φ *n'est pas voisine et* $\operatorname{ind} C_0(\psi) \leq 2$ *lorsque* $\dim \psi = 11$. *Alors, on a équivalence entre:*
(1) $\varphi_{F(\psi)}$ *est isotrope;*
(2) ψ *est voisine d'une* 4*-forme de Pfister dont* φ *contient une voisine.*

On introduit les notations suivantes:

Notations 2.22. *Soient* $n \geq m \geq 1$ *deux entiers. A un entier* l *tel que* $m \geq l \geq 0$, *on associe les ensembles suivants:*

1. $L_{n,m,l}(F)$ *est l'ensembles des formes* $\alpha\pi \perp \beta\tau$ *anisotropes avec* $\alpha, \beta \in F^*$, $\pi \in P_nF$, $\tau \in P_mF$ *et* $i_W(\pi \perp -\tau) = 2^l$.
2. $(L_{n,m,l}(F))'$ *est l'ensemble des formes* $\alpha\pi' \perp \beta\tau'$ *anisotropes avec* $\alpha, \beta \in F^*$, $\pi = \langle 1 \rangle \perp \pi' \in P_nF$, $\tau = \langle 1 \rangle \perp \tau' \in P_mF$ *et* $i_W(\pi \perp -\tau) = 2^l$.

Pour le cas des formes de $L_{n,m,m-1}(F)$, on combine la proposition 2.13 et les théorèmes 2.17, 2.19 pour obtenir:

Corollaire 2.23. *Soit* $\varphi \in L_{n,m,m-1}(F)$ *anisotrope et soit* ψ *une forme quadratique de dimension* 2^{n+1}. *Alors on a:*
(1) Si $n > m$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* $(n+1)$*-forme de Pfister dont* φ *contient une voisine.*
(2) Si $n = m$ *et* F *est de caractéristique* 0 *lorsque* $n \geq 4$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* $(n+1)$*-forme de Pfister dont* φ *contient une voisine, ou* $\varphi \perp \alpha\psi \in I^{n+2}F$ *pour un certain* $\alpha \in F^*$.

Pour le cas des formes de $L_{3,2,1}(F)$ on obtient:

Théorème 2.24. *On garde les mêmes notations que dans* $(*)$. *On suppose que* $\varphi \in L_{3,2,1}(F)$ *est anisotrope mais non voisine. Soit* ψ *une forme quadratique telle que* $11 \leq \dim \psi \leq 16$.
(1) Si $\dim \psi \geq 13$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* 4*-forme de Pfister dont* φ *contient une voisine.*
(2) Si $\dim \psi = 12$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) ψ *est voisine d'une* 4*-forme de Pfister dont* φ *contient une voisine ou il existe* $c \in F^*$, $r \in D_F(\tau)$ *tels que* $ab\pi \perp -\psi \perp c\eta \perp r\pi \in I^5F$.
(3) Si $\dim \psi = 11$ *et* $\operatorname{ind} C_0(\psi) \leq 2$, *alors on a équivalence entre:*
(i) $\varphi_{F(\psi)}$ *est isotrope,*
(ii) $\varphi_{F(\psi')}$ *est isotrope où* $\psi' = \psi \perp \langle -d_{\pm}\psi \rangle$.

3. Les formes quadratiques de $(L_{n,m}(F))'$

Voici certains cas où l'isotropie d'une forme $\varphi \in (L_{n,m}(F))'$ a été étudiée:

1. Si $m = 1$, alors $\varphi_{F(\sqrt{u})} \in GP_nF(\sqrt{u})$ pour un certain $u \in F^*$. Dans ce cas, l'isotropie de φ a été étudiée par Hoffmann [9].
2. Si $n = m = 2$, alors l'isotropie de φ a été étudiée par Hoffmann [7], l'auteur [24] et Izhboldin-Karpenko [12], [13].

Lemme 3.1. ([8, Lemma 3]) *Soient φ et ψ des formes quadratiques telles que $\psi \subset \varphi$ et $\dim \psi \geq \dim \varphi - i_W(\varphi) + 1$. Alors, ψ est isotrope.*

La proposition suivante précise que dans certains cas l'isotropie d'une forme de $(L_{n,m}(F))'$ se ramène à celle d'une forme de $L_{n,m}(F)$ qui la contient.

Proposition 3.2. *Soient $\varphi' \in (L_{n,m,l}(F))'$ et $\varphi \in L_{n,m,l}(F)$ qui contient φ' comme une sous-forme. Si $l \geq 2$, alors pour toute extension de corps K/F on a φ_K isotrope si et seulement si φ'_K isotrope.*

Démonstration. Puisque $\varphi \in L_{n,m,l}(F)$, alors φ est divisible par une l-forme de Pfister. Ainsi, $i_W(\varphi_{F(\varphi)}) \geq 2^l$. Puisque $\dim \varphi = 2^n + 2^m$, $\dim \varphi' = 2^n + 2^m - 2$ et $l \geq 2$, on vérifie bien que $\dim \varphi' \geq \dim \varphi - i_W(\varphi_{F(\varphi)}) + 1$. Par le lemme 3.1 on a $\varphi'_{F(\varphi)}$ isotrope. Puisque $\varphi_{F(\varphi')}$ est isotrope, le résultat se déduit de [20, Theorem 3.3].

Pour la suite de cette section, on va se limiter à étudier l'isotropie des formes de $(L_{n,2,1}(F))'$ avec $n \geq 3$.

Soient $\pi = \pi' \perp \langle 1 \rangle \in P_nF$ (avec $n \geq 3$), $\tau = \tau' \perp \langle 1 \rangle \in P_2F$ et $b \in F^*$. On suppose $i_W(\pi \perp -\tau) = 2$. Par la proposition 2.1, il existe $d, d' \in F^*$ et $\pi_1 = \langle 1 \rangle \perp \pi'_1 \in P_{n-1}F$ tels que $\pi \cong \langle\langle d \rangle\rangle \otimes \pi_1$ et $\tau \cong \langle\langle d, d' \rangle\rangle$.

On fixe les notations suivantes:

$$(**) \qquad \begin{cases} \varphi = \pi' \perp b\tau' \\ \eta = -bd' \langle 1, -d \rangle \otimes \pi'_1 \perp \langle 1, bd \rangle \\ \pi_0 = \pi \perp -bd'\pi. \end{cases}$$

On suppose que φ est anisotrope.

Lemme 3.3. *Avec les mêmes notations et hypothèses que dans $(**)$, on a:*
(1) $\eta_{F(\pi)}$ est isotrope.
(2) La forme π_0 est anisotrope.
(3) Si φ n'est pas voisine, alors η est anisotrope.

Démonstration. Soit $\xi = -bd' \langle 1, -d \rangle \otimes \pi'_1$.
(1) On a $\xi \subset \eta$ et $-bd'\xi \subset \pi$. Puisque $n \geq 3$, on a $\dim \xi > 2^{n-1}$ et donc $\xi_{F(\pi)}$ est isotrope. Ainsi, $\eta_{F(\pi)}$ est aussi isotrope.
(2) La forme φ est anisotrope et $\pi' \perp -bd' \langle 1, -d \rangle$ est une sous-forme de π_0 et φ, de dimension $> 2^n$. Ainsi, π_0 est anisotrope.
(3) Supposons que φ ne soit pas une voisine et que η soit isotrope. Clairement, on a

$$\begin{cases} \pi_0 = \pi' \perp -bd' \langle 1, -d \rangle \perp \xi \perp \langle 1 \rangle \\ \varphi = \pi' \perp -bd' \langle 1, -d \rangle \perp \langle -bd \rangle \end{cases} \qquad (2)$$

La forme $\xi \perp \langle 1\rangle$ est anisotrope car c'est une sous-forme de π_0. Puisque η est isotrope, on obtient $\langle -bd\rangle \subset \xi \perp \langle 1\rangle$, et par (2) on voit bien que $\varphi \subset \pi_0$. Comme $\dim \varphi > 2^n$ on déduit que φ est une voisine de π_0, une contradiction.

Définition 3.4. *On garde les mêmes notations et hypothşes que dans* $(**)$. *On définit* $S(\varphi)$ *comme étant l'ensemble des scalaires* $\alpha \in F^*$ *pour lesquels les deux formes* $\tau \perp \alpha\pi$ *et* $d'\langle 1, b\rangle \perp (\tau \perp \alpha\pi)_{an}$ *sont isotropes.*

Proposition 3.5. *Avec les mêmes notations et hypothşes que dans* $(**)$, *on a:*
(1) L'ensemble $S(\varphi)$ *est non vide.*
(2) Si $\alpha \in S(\varphi)$, *alors* $\dim(d'\langle 1, b\rangle \perp (\tau \perp \alpha\pi)_{an})_{an} \leq 2^n$.
(3) On a $P_{n+1}F \cap W(F(\eta)/F) = \{\pi \perp \alpha\pi \mid \alpha \in S(\varphi)\}$.

Concernant l'isotropie de φ (φ comme dans $(**)$), on a le théorème suivant:

Théorème 3.6. *On garde les mêmes notations et hypothèses que dans* $(**)$. *Soit* ψ *une forme quadratique de dimension* 2^{n+1}. *Alors, on a équivalence entre:*
(1) $\varphi_{F(\psi)}$ *est isotrope;*
(2) Il existe $s \in S(\varphi)$ *telle que* ψ *soit semblable à* $\pi \perp sbd'\pi$.

Lorsque $n = 3$, on obtient:

Théorème 3.7. *On garde les mêmes notations et hypothèses que dans* $(**)$. *On suppose que* $n = 3$. *Soit* ψ *une forme quadratique telle que* $11 \leq \dim \psi \leq 16$ *et* $\operatorname{ind} C_0(\psi) = 2$ *lorsque* $\dim \psi = 11$. *On a équivalence entre:*
(1) $\varphi_{F(\psi)}$ *est isotrope;*
(2) ψ *est voisine d'une* 4*-forme de Pfister* ρ *telle que* $\varphi_{F(\rho)}$ *soit isotrope.*

4. Quelques résultats cohomologiques

D'après Arason [1] il existe une application $\tilde{e}^n$ de P_nF vers H^nF, définie par $\tilde{e}^n(\langle\langle a_1, \ldots, a_n\rangle\rangle) = (a_1) \cdot \ldots \cdot (a_n)$. L'application $\tilde{e}^n$ se prolonge en un homomorphisme e^n de $I^nF/I^{n+1}F$ vers H^nF pour $n = 0, 1, 2$. On a $e^0(\varphi) = \dim \varphi \in H^0F \simeq \mathbb{Z}/2$, $e^1(\varphi) = d_\pm\varphi \in H^1F \simeq F^*/F^{*2}$, $e^2(\varphi) = c(\varphi) \in H^2F \simeq \mathrm{Br}_2(F)$ où $\mathrm{Br}_2(F)$ est la 2-torsion du groupe de Brauer $\mathrm{Br}(F)$ de F. e^0, e^1 sont des isomorphismes. Lorsque $n = 3, 4$ l'application $\tilde{e}^n$ se prolonge en un homomorphisme de $I^nF/I^{n+1}F$ vers H^nF (Arason [1] pour $n = 3$ et Jacob-Rost [14] pour $n = 4$).

Dans le théorème suivant, on calcule le noyau $H^i(F_\epsilon/F)$ lorsque $i \leq n$:

Théorème 4.1. *On garde les mêmes notations que dans* $(*)$. *On suppose que* φ *est anisotrope et que* F *est de caractéristique* 0 *lorsque* $n \geq 5$. *Soit* F_ϵ *le corps de voisinage de* φ. *Alors* $\mid H^i(F_\epsilon/F) \mid \leq 2$ *pour* $i \leq n$. *Plus précisément:*

$$H^i(F_\epsilon/F) = \begin{cases} \{0\} & si \quad n > i \\ \{0, e^n(\eta)\} & si \quad n = i = m. \end{cases}$$

Si $n = i > m$, *alors* $H^i(F_\epsilon/F) \subset \{0, e^n(\pi)\}$ *et ce noyau est nul si la conjecture 2.9 est vraie.*

Pour la preuve de ce théorème on commence par un lemme préliminaire.

LEMME 4.2. *Soit* η *comme dans* $(*)$ *qu'on suppose non nulle, et soit* $(F_i, \eta_i)_{0 \leq i \leq \mathrm{h}(\eta)}$ *sa tour de déploiement générique. On supposons que* $n = m$ *et que* F *est de caractéristique* 0 *lorsque* $n \geq 5$. *Alors:*
(1) $\deg(\eta) = n$.
(2) Si $\mathrm{h}(\eta) \geq 2$, *alors:*
(i) Pour tout $i \in \{0, \cdots, \mathrm{h}(\eta) - 2\}$ *la forme* $(\eta_i)_{F_i(\tau)}$ *est isotrope.*
(ii) Pour tout $(i, j) \in \{0, \cdots, \mathrm{h}(\eta) - 1\} \times \{0, \cdots, n\}$, *on a*

$$\mathrm{Ker}(H^j F \longrightarrow H^j F_i) = \{0\}.$$

Démonstration. Puisque $\eta \not\sim 0$, on a $h := \mathrm{h}(\eta) \geq 1$.
(1) On a $\eta \in I^n F$ et $\dim \eta_{\mathrm{an}} < 2^{n+1}$ ce qui implique $\deg(\eta) = n$.
(2) (i) Puisque $\deg(\eta) = n$, on obtient $\dim \eta_i > 2^n$ pour tout $i \in \{0, \cdots, h - 2\}$. De la relation $\eta_{F(\tau)} \sim \pi_{F(\tau)}$ on déduit $(\eta_i)_{F_i(\tau)} \sim \pi_{F_i(\tau)}$. Par raison de dimension la forme $(\eta_i)_{F_i(\tau)}$ est isotrope pour $i \in \{0, \cdots, h - 2\}$.
(ii) Soient $(i, j) \in \{0, \cdots, h - 1\} \times \{0, \cdots, n\}$ et $x \in \mathrm{Ker}(H^j F \longrightarrow H^j F_i)$. Le résultat est évident pour $i = 0$. Supposons $i \geq 1$. Par (i) l'extension $F_i(\tau)/F(\tau)$ est transcendante pure. Ainsi, $x \in \mathrm{Ker}(H^j F \longrightarrow H^j F(\tau))$. Si $j < n$ on déduit par (R3) que $x = 0$. Si $j = n$ on déduit par (R1) que $x \in \{0, e^n(\tau)\}$. Si $x = e^n(\tau)$, alors par (R4) et le Hauptsatz $\tau_{F_i} \sim 0$. Par récurrence il suffit de considérer le cas τ_{F_1} hyperbolique. Ceci implique par le théorème de la sous-forme que $\tau \sim 0$ puisque $\mathrm{h}(\eta) \geq 2$ et donc $\dim \eta > 2^n$, une contradiction. Ainsi, $x = 0$.
Démonstration du théorème 4.1. Si $\eta \sim 0$, alors $F_\epsilon = F$ et le théorème est évident. Pour la suite, on suppose que $\eta \not\sim 0$ et donc $\mathrm{h}(\eta) \geq 1$. Soit $(\eta_0 = \eta_{\mathrm{an}}, \dots, \eta_{\mathrm{h}(\eta)})$ la suite des noyaux de la tour de déploiement générique de η. Lorsque $n = m$, on a $\deg(\eta) = n$ par le lemme 4.2(1) et $\epsilon = \mathrm{h}(\eta)$ par la remarque 2.3. Soit $x \in H^i(F_\epsilon/F)$.

(1) Supposons $n > i$.

(i) Si $n > m$, alors $F_\epsilon(\pi)/F(\pi)$ est transcendante pure. Ainsi, $x \in H^i(F(\pi)/F)$. Puisque $\dim \pi = 2^n > 2^i$, on déduit par (R3) que $x = 0$.

(ii) Si $n = m$. On a $\eta_{\epsilon-1} \in GP_n(F_{\epsilon-1})$. Puisque $F_\epsilon = F_{\epsilon-1}(\eta_{\epsilon-1})$ et $x_{F_{\epsilon-1}} \in H^i(F_\epsilon/F_{\epsilon-1})$, on obtient par (R3) que $x \in H^i(F_{\epsilon-1}/F)$. Si $\epsilon = 1$, alors $x = 0$, sinon on déduit par le lemme 4.2(2) que $x = 0$.

(2) Supposons $n = i = m$. Puisque $x_{F_{\epsilon-1}} \in H^n(F_\epsilon/F_{\epsilon-1})$ et $\eta_{\epsilon-1} \in GP_n F_{\epsilon-1}$, on déduit par (R1) que $x_{F_{\epsilon-1}} \in \{0, e^n(\eta)_{F_{\epsilon-1}}\}$ (car $e^n(\eta)_{F_{\epsilon-1}} = e^n(\eta_{\epsilon-1})$). Soit $y \in H^n F$ une classe définie comme suit:

$$y = \begin{cases} x & \text{si } x_{F_{\epsilon-1}} = 0 \\ x + e^n(\eta) & \text{si } x_{F_{\epsilon-1}} = e^n(\eta)_{F_{\epsilon-1}}. \end{cases}$$

On a $y \in H^n(F_{\epsilon-1}/F)$. Si $\epsilon = 1$, alors $y = 0$, sinon le lemme 4.2(2) implique que $y = 0$.

(3) Supposons $n = i > m$. Comme dans le cas (1)(i) $x \in H^n(F(\pi)/F) = \{0, e^n(\pi)\}$. Si de plus la conjecture 2.9 est vraie alors π_{F_ϵ} est anisotrope et donc par le Hauptsatz $\pi_{F_\epsilon} \notin I^{n+1}F$. Par (R4) on a $e^n(\pi)_{F_\epsilon} \neq 0$ et donc $x = 0$.

Proposition 4.3. *On garde les mêmes notations que dans* $(*)$. *On suppose que* $n = m$ *et que* F *est de caractéristique* 0 *lorsque* $n \geq 5$. *Soit* ψ *une forme quadratique telle que* $\psi_{F_\epsilon} \in I^{n+1}F_\epsilon$. *Alors,* $\psi \in I^n F$.

Démonstration. Par hypothèse $e^i(\psi_{F_\epsilon}) = 0$ pour $i \leq n$. On affirme que si $\varphi \in I^j F$ pour un certain $j \in \{1, \cdots, n-1\}$, alors $\psi \in I^{j+1}F$. En effet, puisque $e^j(\psi)_{F_\epsilon} = 0$ on obtient par le théorème 4.1 que $e^j(\psi) = 0$ et par (R4) $\psi \in I^{j+1}F$. Comme $\psi \in IF$, on déduit par itération que $\psi \in I^n F$.

5. Démonstrations

5.1. Démonstration de la proposition 2.7.

Lemme 5.1. *On garde les mêmes notations et hypothèses que dans la proposition 2.7. Pour* K/F *une extension, on a:*
(1) Si $n > m$: $\varphi_K \sim 0 \Longleftrightarrow \pi_K \sim \tau_K \sim 0$ $(\Longleftrightarrow \zeta_K \sim \sigma_K \sim 0)$.
(2) Si $n = m$: φ_K *est voisine* $\Longleftrightarrow \pi_K \cong \tau_K$ $(\Longleftrightarrow \zeta_K \cong \sigma_K)$.

Démonstration. (1) C'est une simple conséquence de l'hypothèse $n > m$ et du fait qu'une forme de Pfister isotrope est hyperbolique.
(2) Dans ce cas $\dim \varphi = 2^{n+1}$. Il est clair que $\pi_K \cong \tau_K$ implique que $\varphi_K \in GP_{n+1}K$. Réciproquement, si $\varphi_K \in GP_{n+1}K$ alors $\varphi_{K(\pi)} \sim 0$ et donc $\tau_{K(\pi)} \sim 0$. On conclut par le théorème de la sous-forme et la multiplicativité d'une forme de Pfister.
Démonstration de la proposition. (1) Supposons $n > m$. Pour $i \in \{0, \cdots, \epsilon\}$ (resp. $j \in \{0, \cdots, \nu\}$) soit r_i (resp. s_j) tel que $2^{r_i} = i_W(\eta_{F_i})$ (resp. $2^{s_j} = i_W(\delta_{G_j})$). Puisque $2^{r_i} = i_W(\eta_{F_i})$, on obtient par la proposition 2.1 que les formes π_{F_i} et τ_{F_i} sont divisibles par une forme de $P_{r_i}F_i$. Par le lemme 5.1, on déduit que pour $i \in \{0, \cdots, \epsilon\}$ on a $i_W(\eta_{F_i}) = i_W(\delta_{F_i})$, et donc 2^{r_i} appartient à la suite des indices de déploiement de δ. De même pour $j \in \{0, \cdots, \nu\}$ l'entier 2^{s_j} appartient à la suite des indices de déploiement de η. Remarquons aussi que $2^{r_i}, 2^{s_j} \leq 2^m$ pour tout $(i,j) \in \{0, \cdots, \epsilon\} \times \{0, \cdots, \nu\}$. Ainsi, on déduit qu'on a nécessairement $\nu = \epsilon$ et par [21, Remark 5.5] on a que F_i est F-équivalent à G_i pour tout $i \in \{0, \cdots, \epsilon\}$.
(2) Supposons $n = m$. Par le lemme 5.1 on a $\zeta_{F_\epsilon} \cong \sigma_{F_\epsilon}$ et $\pi_{G_\nu} \cong \tau_{G_\nu}$. Ainsi, $i_W(\eta_{G_\nu}) = i_W(\delta_{F_\epsilon}) = 2^m$. Par [21, Remark 5.5] on a que F_ϵ est F-équivalent à G_ν.

5.2. Démonstration de la proposition 2.12.

Supposons que φ_{F_ϵ} soit isotrope. Par la proposition 2.2(i) $(\pi_0)_{F_\epsilon} \sim 0$.
(i) Supposons $n > m$. Par la conjecture 2.10, on a $\pi_0 \cong \pi \perp r\pi$ pour un certain $r \in D_F(-\tau)$. Par simplification, on a $a\pi \cong br\pi$. Ainsi, $\varphi \cong br\pi \perp b\tau \cong br\pi \perp -br\tau$. Une contradiction car φ est anisotrope.

(ii) Supposons $n = m$. Par la conjecture 2.10 il existe $\alpha \in F^*$ tel que $\pi_0 \perp -(\eta \perp \alpha\eta) \in I^{n+2}F$. Ainsi, $ab\pi \perp \tau \perp -\alpha\eta \in I^{n+2}F$. Par le Hauptsatz $ab\pi \perp \tau = b\varphi$ est isotrope, une contradiction.

5.3. Démonstration de la proposition 2.13. (1) Si $i_W(\pi \perp -\tau) = 2^m$, alors par la proposition 2.2(iii) $F_\epsilon = F$ et la proposition est évidente.
(2) Si $i_W(\pi \perp -\tau) = 2^{m-1}$. Soient $\rho \in P_{m-1}F$, $\mu = \langle 1\rangle \perp \mu' \in P_{n-m+1}F$ et $d \in F^*$ tels que $\pi \cong \rho \otimes \mu$ et $\tau \cong \rho \otimes \langle\langle d\rangle\rangle$. On a $\eta_{\text{an}} = \rho \otimes (\mu' \perp \langle d\rangle)$ et $F_\epsilon = F(\eta_{\text{an}})$. Soit $\delta \in P_{n+1}F \cap W(F_\epsilon/F)$.
(i) Si $n > m$, alors par la proposition 2.2(iv) $\delta_{F(\pi)} \sim 0$. Ainsi, $\delta \cong \pi \perp \alpha\pi$ pour un certain $\alpha \in F^*$. Par le théorème de la sous-forme, on a $\pi \perp \alpha\pi \cong s\rho \otimes (\mu' \perp \langle d\rangle) \perp \xi$ pour $s \in F^*$ et ξ une forme de dimension 2^n. Soit $e \in D_F(\rho \otimes \mu') \subset D_F(\pi)$. Puisque $e, es \in D_F(\pi \perp \alpha\pi)$, on peut supposer, par multiplicativité, que $s = 1$. Par simplification on a $\tau \perp \alpha\pi \sim \xi$. Par comparaison des dimensions, on a $\tau \perp \alpha\pi$ isotrope. Ainsi, il existe $r \in D_F(-\tau) \cap D_F(\alpha\pi)$. On a alors $\delta \cong \pi \perp r\pi$ avec $r \in D_F(-\tau)$.
(ii) Si $n = m$, alors $\eta_{\text{an}} \in GP_mF$. Ainsi, il existe $x, y \in F^*$ tels que $\delta \cong \langle x, y\rangle \otimes \eta_{\text{an}}$. On a bien $\delta \perp -(\eta \perp xy\eta) \in I^{n+2}F$.

5.4. Démonstration de la proposition 2.16. (1) Supposons $n > m$:
(i) Supposons que la conjecture 2.10 soit vraie. Soit $x \in H^{n+1}(F_\epsilon/F)$. Puisque $F_\epsilon(\pi)/F(\pi)$ est transcendante pure, on déduit que $x \in H^{n+1}(F(\pi)/F)$. Par (R1) il existe $\alpha \in F^*$ tel que $x = e^{n+1}(\pi \perp \alpha\pi)$. Puisque $x_{F_\epsilon} = 0$, on a $e^{n+1}((\pi \perp \alpha\pi)_{F_\epsilon}) = 0$. Par (R4) on a $(\pi \perp \alpha\pi)_{F_\epsilon} \in I^{n+2}F_\epsilon$. Par le Hauptsatz, on a $\pi \perp \alpha\pi \in W(F_\epsilon/F)$. D'après la conjecture 2.10, on déduit que $\pi \perp \alpha\pi \cong \pi \perp -r\pi$ avec $r \in D_F(\tau)$. Ainsi, $x = e^n(\pi)(r)$ avec $r \in D_F(\tau)$.
(ii) Supposons que la conjecture 2.15 soit vraie. Soit $\rho \in P_{n+1}F \cap W(F_\epsilon/F)$. Alors, $e^{n+1}(\rho) \in H^{n+1}(F_\epsilon/F)$. Par la conjecture 2.15 on a $e^{n+1}(\rho) = e^n(\pi)(r)$ avec $r \in D_F(\tau)$. Par (R4) on a $\rho \perp -(\pi \perp -r\pi) \in I^{n+2}F$. Puisque $\dim(\rho \perp -(\pi \perp -r\pi))_{\text{an}} < 2^{n+2}$, le résultat se déduit par le Hauptsatz.
(2) Supposons $n = m$:
Supposons que la conjecture 2.15 soit vraie. Soit $\delta \in P_{n+1}F \cap W(F_\epsilon/F)$. Alors, $e^{n+1}(\delta) \in H^{n+1}(F_\epsilon/F)$. Par la conjecture 2.15 il existe $s \in F^*$ tel que $e^{n+1}(\delta) = e^n(\eta)(-s)$. Par (R4) on obtient $\delta \perp -(\eta \perp s\eta) \in I^{n+2}F$.

5.5. Démonstration du théorème 2.19. (1) C'est une conséquence de [8, Theorem 1].

Puisque la conjecture 2.10 implique la conjecture 2.9, on déduit qu'on a par hypothèse $(\pi_0)_{F_\epsilon}$ anisotrope.
(ii) $\Longrightarrow$ (i) Si $\psi \in P_{n+1}F$ et φ contient une voisine de ψ, alors on sait que $\varphi_{F(\psi)}$ est isotrope. Si $\varphi \perp \alpha\psi \in I^{n+2}F$, alors par le Hauptsatz on a $(\varphi \perp \alpha\psi)_{F(\psi)} \sim 0$ et donc $\varphi_{F(\psi)}$ est isotrope.
(i) $\Longrightarrow$ (ii) Soit ψ de dimension 2^{n+1} tel que $\varphi_{F(\psi)}$ soit isotrope. En particulier, $(\pi_0)_{F_\epsilon(\psi)}$ est isotrope et donc hyperbolique. On peut supposer que $1 \in D_F(\psi)$.

Ainsi,

$$\psi_{F_\epsilon} \cong (\pi_0)_{F_\epsilon} \tag{3}$$

• Si $n > m$: Par la proposition 2.2(iv) et l'équation (3) $\psi \in W(F(\pi)/F)$. Ainsi, $\psi \cong \pi \perp \alpha\pi$ pour un certain $\alpha \in F^*$. Par l'équation (3), on a $\alpha\pi \perp -ab\pi \in W(F_\epsilon/F)$. Par la conjecture 2.10, $\pi \perp -ab\alpha\pi \cong \pi \perp r\pi$ pour un certain $r \in D_F(-\tau)$. Par la simplification de Witt, on a $ab\alpha\pi \cong -r\pi$. Ainsi, $\varphi \cong -br(\alpha\pi \perp \tau)$. Puisque $\psi \cong \pi \perp \alpha\pi$, on voit que $-br(\alpha\pi \perp \langle 1\rangle)$ est une voisine de ψ contenue dans φ. Donc l'assertion (2) est prouvée.
• Si $n = m$: Puisque $\psi_{F_\epsilon} \in I^{n+1}F_\epsilon$ on déduit par la proposition 4.3 que $\psi \in I^nF$. Ainsi, $e^n(\psi) \in H^n(F_\epsilon/F)$. Par le théorème 4.1, on a $e^n(\psi) \in \{0, e^n(\eta)\}$.
(i) Si $e^n(\psi) = e^n(\eta)$, alors $\psi \perp -\eta \in I^{n+1}F$ par (R4). On a $e^{n+1}(\psi \perp -\eta)_{F_\epsilon} = e^{n+1}(\psi_{F_\epsilon})$ (car $\eta_{F_\epsilon} \sim 0$). Par l'équation (3) on a $e^{n+1}(\psi \perp -\eta)_{F_\epsilon} = e^{n+1}(\pi_0)_{F_\epsilon}$. Par la conjecture 2.15 on a $e^{n+1}(\psi \perp -\eta) + e^{n+1}(\pi_0) = e^{n+1}(\eta \perp r\eta)$ pour un certain $r \in F^*$. Par (R4) on a $\psi \perp -\eta \perp \pi_0 \perp \eta \perp r\eta \in I^{n+2}F$. Après simplification et puisque $\pi_0 \perp r\pi_0 \in I^{n+2}F$, on obtient $\psi \perp -rb\varphi \in I^{n+2}F$.
(ii) Si $e^n(\psi) = 0$, alors $\psi \in I^{n+1}F$ et donc $\psi \in P_{n+1}F$. Par l'équation (3) $e^{n+1}(\psi \perp -\pi_0) \in H^{n+1}(F_\epsilon/F)$. Par la conjecture 2.15 il existe $s \in F^*$ tel que $e^{n+1}(\psi \perp -\pi_0) = e^{n+1}(\eta \perp s\eta)$. Par (R4) on a $\psi \perp -\pi_0 \perp \eta \perp s\eta \in I^{n+2}F$. Après simplification, on a $\psi \perp -b\varphi \perp s\eta \in I^{n+2}F$. Puisque η est isotrope, on a $\dim(-b\varphi \perp s\eta)_{\mathrm{an}} < 2^{n+2}$. Par le Hauptsatz on a $(-b\varphi \perp s\eta)_{F(\psi)} \sim 0$. Par conséquent, $-b\varphi \perp s\eta \sim u\psi$ pour un certain $u \in F^*$. On a $\dim \eta_{\mathrm{an}} < 2^{n+1}$. Ainsi, $i_W(-b\varphi \perp -u\psi) > 2^n$, et donc les formes $-b\varphi$ et $u\psi$ contiennent en commun une sous-forme μ de dimension $> 2^n$ c'est-à-dire φ contient une voisine de ψ. Donc l'assertion (3) est prouvée.

5.6. Démonstration du corollaire 2.21. Dans ce cas, φ_{F_ϵ} est une voisine anisotrope de $(\pi_0)_{F_\epsilon}$.
(2) $\Longrightarrow$ (1) Evident.
(1) $\Longrightarrow$ (2) Puisque φ n'est pas voisine, on a $i_W(\eta) = 1$ et donc η_{an} est de dimension 8. Par la proposition 2.2(ii) on a $F_\epsilon = F(\eta_{\mathrm{an}})$. Supposons que $\varphi_{F(\psi)}$ soit isotrope et $1 \in D_F(\psi)$. Alors, ψ_{F_ϵ} est une sous-forme de $(\pi_0)_{F_\epsilon}$. On écrit $(\pi_0)_{F_\epsilon} \cong \psi_{F_\epsilon} \perp \xi'$ avec ξ' est une F_ϵ-forme quadratique.
(i) Si $\dim \psi \geq 13$, alors $\dim \xi' \leq 3$. D'après [16, Theorem 2] il existe δ_1 une F-forme telle que $\xi' \cong (\delta_1)_{F_\epsilon}$. On pose $\psi_1 = \psi \perp \delta_1$. On a $(\pi_0)_{F_\epsilon} \cong (\psi_1)_{F_\epsilon}$. D'après les propositions 2.13 et 2.12, on obtient $\varphi_{F(\psi_1)}$ isotrope.
(ii) Si $\dim \psi = 12$, alors $\dim \xi' = 4$. On a $\eta_{\mathrm{an}} \notin I^2F$. D'après [16, Theorem 6] il existe δ_2 une F-forme telle que $\xi' \cong (\delta_2)_{F_\epsilon}$. On pose $\psi_2 = \psi \perp \delta_2$. Comme dans le cas (i) $\varphi_{F(\psi_2)}$ est isotrope.
(iii) Si $\dim \psi = 11$ et ind $C_0(\psi) \leq 2$. Comme $c(\psi)_{F_\epsilon} = c(\xi')$ et $\dim \xi' = 5$, on déduit que $\xi' \cong \langle d'\rangle \perp \tau$ pour $\tau \in GP_2F_\epsilon$ et $d' = d_\pm \xi' = -d_\pm \psi$ [21, Page 10]. D'après [16, Theorem 6] il existe $\delta_3 \in GP_2F$ telle que $\tau \cong (\delta_3)_{F_\epsilon}$. On pose $\psi_3 = \psi \perp \langle d'\rangle \perp \delta_3$. Comme dans le cas (i) $\varphi_{F(\psi_3)}$ est isotrope.

Dans chacun des trois cas précédents, on obtient par le théorème 2.19 que $\psi_i \in GP_{n+1}F$ et φ contient une voisine de ψ_i pour $i = 1, 2, 3$. Par conséquent, ψ est voisine d'une 4-forme de Pfister dont φ contient une voisine.

5.7. DÉMONSTRATION DU THÉORÈME 2.24. On suppose $1 \in D_F(\psi)$. Puisque $i_W(\eta) = 2$, on a $\dim \eta_{\mathrm{an}} = 8$, et par la proposition 2.2(ii) $F_\epsilon = F(\eta_{\mathrm{an}})$. Dans ce cas la conjecture 2.10 est vraie (proposition 2.13). Ainsi, φ_{F_ϵ} est une voisine anisotrope de $(\pi_0)_{F_\epsilon}$. Posons $d = d_\pm\psi$.
Supposons que $\varphi_{F(\psi)}$ soit isotrope. Alors, $(\pi_0)_{F_\epsilon(\psi)} \sim 0$, et par conséquent

$$(\pi_0)_{F_\epsilon} \cong \psi_{F_\epsilon} \perp \xi \tag{4}$$

pour ξ une F_ϵ-forme quadratique.
(1) On suppose que $\dim \psi \geq 13$. Dans ce cas on reprend la même méthode que celle utilisée dans la démonstration du corollaire 2.21 pour montrer qu'il existe $\psi' \in GP_4F$ tel que $\psi \subset \psi'$ et φ contient une voisine de ψ'.
(2) On suppose que $\dim \psi = 12$. On a $c(d\psi)_{F_\epsilon} = c(\xi)$ et $c(\eta_{\mathrm{an}}) = c(\tau)$.
(i) Si $d \neq 1$, alors d'aprés [16, Théorème 2], il existe ξ' une F-forme de dimension 4 telle que $\xi = (\xi')_{F_\epsilon}$. Puisque $(\pi_0)_{F_\epsilon} \cong (\psi \perp \xi')_{F_\epsilon}$, on déduit par la proposition 2.13 que π_0 est isotrope sur $F(\psi \perp \xi')$. Par le théorème 2.19 on a $\psi \perp \xi' \in GP_4F$ et φ contient une voisine de $\psi \perp \xi'$.
(ii) Si $d = 1$ et $c(\psi) \neq c(\tau)$. Alors, $c(\psi)_{F_\epsilon} \neq c(\tau)_{F_\epsilon}$ car $H^2(F_\epsilon/F) = \{0\}$ (théorème 4.1). Ainsi, $c(\xi) \neq c(\tau)_{F_\epsilon}$. De nouveau par [16, Théorème 2] il existe ξ'' une F-forme de dimension 4 telle que $\xi \cong \xi''_{F_\epsilon}$. Comme dans le cas (i) on a $\psi \perp \xi'' \in GP_4F$ et φ contient une voisine de $\psi \perp \xi''$.
(iii) Si $d = 1$ et $c(\psi) = c(\tau)$, alors ξ est semblable à τ. Ainsi, $(\pi_0)_{F_\epsilon(\tau)} \sim 0$ et donc $\psi_{F_\epsilon(\tau)} \sim 0$. Comme $F_\epsilon(\pi)/F(\pi)$ est transcendante pure, on obtient $\psi_{F(\pi)(\tau)} \sim 0$.
• Si $\psi_{F(\tau)} \sim 0$, alors ψ est divisible par τ et donc voisine d'une 4-forme de Pfister ρ. Par conséquent, $\varphi_{F(\rho)}$ est isotrope et par le théorème 2.19 φ contient une voisine de ρ.
• Si $\psi_{F(\tau)} \not\sim 0$. Alors $\psi_{F(\tau)} \sim \lambda\pi$ pour un certain $\lambda \in F(\tau)^*$. Par l'excellence de $F(\tau)/F$ ([2], [34]) il existe π_1 une F-forme quadratique de dimension 8 telle que $\psi_{F(\tau)} \sim (\pi_1)_{F(\tau)}$. Puisque $c(\pi_1)_{F(\tau)} = 0$, on peut supposer que $\pi_1 \in GP_3F$ [4, 2.10]. Puisque $\psi \perp -\pi_1 \perp \tau \in I^3F$, on obtient que $e^3(\psi \perp -\pi_1 \perp \tau) \in H^3(F(\tau)/F) = c(\tau) \cdot H^1F$ [1]. Ainsi, il existe $c \in F^*$ tel que $e^3(\psi \perp -\pi_1 \perp \tau) = c(\tau) \cdot (c) = e^3(\tau \perp -c\tau)$. Par conséquent,

$$\psi \perp -\pi_1 \perp c\tau \in I^4F \tag{5}$$

Soit $\alpha \in D_F(\pi_1)$. Puisque $\lambda\pi \cong (\pi_1)_{F(\tau)}$, on a $\pi_{F(\tau)} \cong (\alpha\pi_1)_{F(\tau)}$. Ainsi, $(\pi \perp -\alpha\pi_1)_{F(\tau)} \sim 0$. Comme $\dim(\pi \perp -\alpha\pi_1)_{\mathrm{an}} \leq 14$ et $c(\pi \perp -\alpha\pi_1) = 0$, on obtient

$$\pi \perp -\alpha\pi_1 \sim \rho' \otimes \tau \tag{6}$$

pour ρ' une forme quadratique de dimension paire ≤ 2. Des équations (5) et (6) et modulo I^4F, on obtient

$$\psi \perp -\pi \perp (\rho' \perp \langle c\rangle) \otimes \tau \in I^4F \tag{7}$$

Soit $e \in F^*$ tel que $(\rho' \perp \langle c\rangle) \otimes \tau \perp e\tau \in I^4F$. De l'équation (7) on a $\psi \perp -\pi \perp -e\tau \in I^4F$. Avec l'équation (4) on a

$$(ab\pi)_{F_\epsilon} \perp -\xi \perp (-e\tau)_{F_\epsilon} \in I^4F_\epsilon \tag{8}$$

En particulier, $(ab\pi)_{F_\epsilon} \perp -\xi \perp (-e\tau)_{F_\epsilon} \in GP_4F_\epsilon$. Ainsi, $\pi_{F_\epsilon} \cong \tau_{F_\epsilon} \perp e\xi$. Par conséquent, $(e\eta)_{F_\epsilon} \sim \xi$ et par l'équation (4) $(\pi_0)_{F_\epsilon} \sim (\psi \perp e\eta)_{F_\epsilon}$. Comme $\mu := \pi_0 \perp -\psi \perp -e\eta \in I^3F$ (car d'invariant de Clifford trivial), on a $e^3(\mu)_{F_\epsilon} = 0$. Puisque η_{an} n'est pas voisine, on a $H^3(F_\epsilon/F) = \{0\}$ [1]. Par conséquent $\mu \in I^4F$ et donc $e^4(\mu) \in H^4(F_\epsilon/F)$. D'après les propositions 2.13 et 2.16, il existe $r \in D_F(\tau)$ tel que $e^4(\mu) = e^3(\pi) \cdot (r)$. Par (R4) $ab\pi \perp -\psi \perp -e\eta \perp r\pi \in I^5F$. D'où le résultat.
Réciproquement, supposons qu'il existe $x \in F^*$, $y \in D_F(\tau)$ tels que $ab\pi \perp -\psi \perp x\eta \perp y\pi \in I^5F$. Alors, $\pi_0 \perp -\psi \perp x\eta \perp -(\pi \perp -y\pi) \in I^5F$. Par la proposition 2.11(1) on a $\pi \perp -y\pi \in W(F_\epsilon/F)$. Ainsi, $(\pi_0 \perp -\psi \perp x\eta)_{F_\epsilon} \in I^5F_\epsilon$. Puisque $\dim(\eta_{F_\epsilon})_{\mathrm{an}} = 4$, on obtient $\nu := (\pi_0 \perp -\psi)_{F_\epsilon} \perp ((x\eta)_{F_\epsilon})_{\mathrm{an}} \in GP_5F_\epsilon$. Par le Hauptsatz, $\nu_{F_\epsilon(\psi)} \sim 0$ c'est-à-dire $(\pi_0)_{F_\epsilon(\psi)} \sim \psi_{F_\epsilon(\psi)} \perp -((x\eta)_{F_\epsilon(\psi)})_{\mathrm{an}}$. Ainsi, $(\pi_0)_{F_\epsilon(\psi)}$ est isotrope. On applique successivement les propositions 2.13 et 2.12 pour déduire que $\varphi_{F(\psi)}$ est isotrope.
(3) On suppose que $\dim\psi = 11$:
Puisque $\operatorname{ind} C_0(\psi) \leq 2$, on a aussi $\operatorname{ind} C_0(\xi) \leq 2$ et donc $\xi = \langle d\rangle \perp \xi'$ pour une certaine $\xi' \in GP_2F_\epsilon$ [21, Page 10] (c'est-à-dire ξ est voisine). Par conséquent, $(\pi_0)_{F_\epsilon(\delta)} \sim 0$ où $\delta = \psi \perp \langle d\rangle$. On applique successivement les propositions 2.13 et 2.12 pour déduire que $\varphi_{F(\delta)}$ est isotrope.

5.8. Démonstration de la proposition 3.5.

Lemme 5.2. *On garde les mêmes notations que dans* $(**)$. *Soit* $\eta_1 = \langle 1, -d\rangle \otimes \pi_1' \perp \langle -dd'\rangle$ *et* $\eta' = \eta_1 \perp \langle d'\rangle$. *Alors, on a:*
(1) $W(F(\eta)/F) \subset W(F(\eta_1)/F)$.
(2) $c(\eta') = c(\tau)$, $\eta' \in I^2F$ *et les corps* $F(\eta_1)$ *et* $F(\eta')$ *sont* F*-équivalents. En particulier,* $W(F(\eta)/F) \subset W(F(\eta')/F)$.
(3) La forme η_1 *n'est pas une voisine.*

Démonstration. (1) Puisque $\eta_1 \subset -bd'\eta$, on a que $\eta_{F(\eta_1)}$ est isotrope et donc il existe une F-place de $F(\eta)$ vers $F(\eta_1)$ [20, Theorem 3.3]. Par conséquent, $W(F(\eta)/F) \subset W(F(\eta_1)/F)$.
(2) On vérifie facilement que $c(\eta') = c(\tau)$. Il est clair que $\eta' = \langle 1, -d\rangle \otimes (\pi_1' \perp \langle d'\rangle)$. Ainsi, $\eta' \in I^2F$ et $i_W(\eta'_{F(\eta')}) \geq 2$. Par le lemme 3.1 η_1 est isotrope sur $F(\eta')$. Puisque η' est isotrope sur $F(\eta_1)$, les corps $F(\eta')$ et $F(\eta_1)$ sont F-équivalents et donc $W(F(\eta_1)/F) = W(F(\eta')/F)$. Par l'assertion (1) on obtient $W(F(\eta)/F) \subset W(F(\eta')/F)$.
(3) Si η_1 était voisine, on aurait $\eta' \in GP_nF$ (car $\dim\eta' = 2^n$) et donc $c(\eta') = c(\tau) = 0$, une contradiction.
Le théorème suivant jouera un rôle important dans la démonstration.

THÉORÈME 5.3. *(Fitzgerald* [5, Proposition 1.4]*)*
Soient φ une forme quadratique voisine d'une n-forme de Pfister ρ et $\varphi' = \varphi \perp \langle y \rangle$ pour $y \in F^*$. Supposons que φ' ne soit pas voisine de ρ. Soit $\varphi'' \in W(F(\varphi')/F)$. Alors, $\varphi'' \cong \pi_1 \perp \cdots \perp \pi_s$ pour un certain entier $s \geq 1$ et $\pi_i \in GP_{n+1}F \cap W(F(\varphi')/F)$ pour tout $i \in \{1, \cdots, s\}$.

Démonstration de la proposition. (1) On a $-1 \in S(\varphi)$. En effet, il est clair que $\tau \perp -\pi$ est isotrope. Puisque $i_W(\tau \perp -\pi) = 2$, on déduit que $(\tau \perp -\pi)_{\mathrm{an}} = \langle\langle d \rangle\rangle \otimes (-\pi_1' \perp \langle -d' \rangle)$ et donc $-d' \in D_F((\tau \perp -\pi)_{\mathrm{an}})$. Ainsi, $d' \langle 1, b \rangle \perp (\tau \perp -\pi)_{\mathrm{an}}$ est isotrope.
(2) Puisque $\tau \perp \alpha\pi$ est isotrope, alors $(\tau \perp \alpha\pi)_{\mathrm{an}}$ est semblable à $(\tau \perp -\pi)_{\mathrm{an}}$ qui est de dimension 2^n. Puisque $d' \langle 1, b \rangle \perp (\tau \perp \alpha\pi)_{\mathrm{an}}$ est isotrope, le résultat s'en déduit.
(3) Soit $\rho \in P_{n+1}F \cap W(F(\eta)/F)$ anisotrope. Puisque η est isotrope sur $F(\pi)$ (lemme 3.3), on déduit que $\rho_{F(\pi)} \sim 0$. Alors,

$$\rho \cong \pi \perp \alpha\pi$$

pour $\alpha \in F^*$. Par le théorème de la sous-forme on a

$$\rho \cong -bd'\eta \perp \xi$$

pour ξ une forme de dimension 2^n. La simplification de Witt dans la relation $\rho \cong \pi \perp \alpha\pi \cong -bd'\eta \perp \xi$ implique

$$\langle 1, -d \rangle \perp \alpha\pi \cong \langle -bd', -dd' \rangle \perp \xi \tag{9}$$

Par le lemme 5.2(2) on a $\rho_{F(\eta')} \sim 0$, et par le théorème de la sous-forme on a $\rho \cong \eta' \perp \xi'$ pour ξ' une forme quadratique de dimension 2^n (η' comme dans le lemme 5.2). La simplification de Witt dans la relation $\rho \cong -bd'\eta \perp \xi \cong \eta' \perp \xi'$ implique

$$\xi \sim d' \langle 1, b \rangle \perp \xi' \tag{10}$$

Des équations (9) et (10) on obtient

$$\tau \perp \alpha\pi \sim \xi'.$$

On substitue dans l'équation (10) pour obtenir $\xi \sim d' \langle 1, b \rangle \perp (\tau \perp \alpha\pi)_{\mathrm{an}}$. Par comparaison des dimensions on a $d' \langle 1, b \rangle \perp (\tau \perp \alpha\pi)_{\mathrm{an}}$ isotrope. D'où le résultat.
Réciproquement, soit $\alpha \in S(\varphi)$ et $\rho = \pi \perp \alpha\pi \in P_{n+1}F$. On vérifie facilement les relations

$$-bd'\eta \sim \langle 1, -d \rangle \otimes \pi_1' \perp \langle -bd', -dd' \rangle \tag{11}$$

$$\pi \perp -\tau \sim \langle 1, -d \rangle \otimes \pi_1' \perp d' \langle 1, -d \rangle \tag{12}$$

Des relations (11) et (12) on a

$$-bd'\eta \sim (\pi \perp -\tau)_{\mathrm{an}} \perp -d' \langle 1, b \rangle \tag{13}$$

Puisque $\rho \sim (\pi \perp -\tau)_{\mathrm{an}} \perp (\tau \perp \alpha\pi)_{\mathrm{an}}$, on a par l'équation (13) que

$$\rho \sim -bd'\eta \perp (d' \langle 1, b\rangle \perp (\tau \perp \alpha\pi)_{\mathrm{an}})_{\mathrm{an}}.$$

Puisque $\dim \eta = 2^n$ et $\dim(d' \langle 1, b\rangle \perp (\tau \perp \alpha\pi)_{\mathrm{an}})_{\mathrm{an}} \leq 2^n$, on déduit que $\rho_{F(\eta)}$ est isotrope.

5.9. Démonstration du théorème 3.6.

Proposition 5.4. *La forme $\varphi_{F(\eta)}$ est une voisine anisotrope de $(\pi_0)_{F(\eta)}$.*

Démonstration. Comme dans la démonstration de l'assertion (3) du lemme 3.3 on a $\varphi_{F(\eta)} \subset (\pi_0)_{F(\eta)}$. Si $(\pi_0)_{F(\eta)}$ est isotrope, alors on obtient par le théorème de la sous-forme

$$\pi \perp -bd'\pi \cong x\eta \perp \xi$$

pour ξ une forme quadratique de dimension 2^n et $x \in F^*$. Soit $\beta \in D_F(\pi'_1)$. Alors, $-bd'\beta \in D_F(\eta) \cap D_F(\pi \perp -bd'\pi)$. Comme $-xbd'\beta \in D_F(x\eta) \subset D_F(\pi \perp -bd'\pi)$, on peut supposer, par multiplicativité, que $x = 1$. Par simplification, on obtient que $\varphi \sim \xi$. Une contradiction, car φ est anisotrope.
Démonstration du théorème. On suppose que $1 \in D_F(\psi)$.
(1) $\Longrightarrow$ (2) Puisque $\varphi_{F(\psi)}$ est isotrope, on déduit que $\varphi_{F(\eta)(\psi)}$ est isotrope et donc $(\pi_0)_{F(\eta)(\psi)} \sim 0$. Comme $(\pi_0)_{F(\eta)}$ est anisotrope (proposition 5.4), on a par le théorème de la sous-forme $(\pi_0)_{F(\eta)} \cong \psi_{F(\eta)}$. Ainsi, $\pi_0 \perp -\psi \in W(F(\eta)/F)$. On a $\dim(\pi_0 \perp -\psi)_{\mathrm{an}} \leq 2^{n+2} - 2 < 2^{n+2}$. Soit η_1 comme dans le lemme 5.2. Puisque $W(F(\eta)/F) \subset W(F(\eta_1)/F)$ et η_1 n'est pas voisine (lemme 5.2), il existe par le théorème 5.3 une forme $\rho \in GP_{n+1}F \cap W(F(\eta_1)/F$ telle que

$$(\pi_0 \perp -\psi)_{\mathrm{an}} \sim \rho \tag{14}$$

Comme $\pi_0, \rho \in P_{n+1}F$, on a $\psi \in I^{n+1}F$ et donc $\psi \in P_{n+1}F$. On a aussi $\rho_{F(\eta)} \sim 0$ (car $(\pi_0 \perp -\psi)_{F(\eta)} \sim 0$). Par la proposition 3.5, il existe $s \in S(\varphi)$, $u \in F^*$ tels que $\rho \cong u(\pi \perp s\pi)$. De l'équation (14) on a $\pi_0 \perp -\rho$ isotrope. Soit $v \in D_F(\pi_0) \cap D_F(\rho)$. Alors, $\pi_0 \perp -\rho \sim v(\pi_0 \perp -(\pi \perp s\pi)) \sim \psi$. Ainsi, $\psi \cong -vs(\pi \perp sbd'\pi)$.
(2) $\Longrightarrow$ (1) Soit $s \in S(\varphi)$ tel que ψ soit semblable à $\rho := \pi \perp sbd'\pi$. On a $\pi_0 \perp -\rho \sim -bd'(\pi \perp s\pi)$. D'après la proposition 3.5, on déduit que $(\pi_0 \perp -\rho)_{F(\eta)} \sim 0$. Ainsi, $(\pi_0)_{F(\eta)} \cong \rho_{F(\eta)}$. Par la proposition 5.4, on déduit que $\varphi_{F(\rho)}$ est isotrope et donc $\varphi_{F(\psi)}$ l'est aussi.

5.10. Démonstration du théorème 3.7. Par la proposition 5.4, on a que $\varphi_{F(\psi)}$ est isotrope si et seulement si $\psi_{F(\eta)}$ est semblable à une sous-forme de $(\pi_0)_{F(\eta)}$. La forme $\eta \notin I^2F$ car sinon par un simple calcul on aurait φ isotrope. Puisque η est de dimension 8 on peut utiliser les mêmes techniques de descente que dans la démonstration du corollaire 2.21, et on finit la preuve en utilisant le théorème 3.6.

Bibliographie

[1] *J. Kr. Arason*, Cohomologische Invarianten quadratischer Formen, J. Alg. 36 (1975), 448–491.

[2] *J. Kr. Arason*, Excellence of $F(\varphi)/F$ for 2-fold Pfister forms, Appendice de [4].

[3] *R. Elman, T. Y. Lam*, Pfister forms and K-theory of fields, J. Alg. 23 (1972), 181–213.

[4] *R. Elman, T. Y. Lam, A. Wadsworth*, Amenable fields and Pfister extensions, Conf. Quadratic forms (1976) (G. Orzech, ed.). Queen's papers on Pure and Appl. Math. 46. Queen's Univ. Kingston, Ont., 445–492 (1977).

[5] *R. W. Fitzgerald*, Witt kernels of function fields extensions, Pacific J. Math. 109 (1983), 89–106.

[6] *D. W. Hoffmann*, Isotropy of 5-dimensional quadratic forms over the function field of a quadric, Proceedings of 1992 Santa Barbara Summer Research Institute (eds. B. Jacob, A. Rosenberg). Proc. Symp. Pure Math. 58.2 (1995), 217–225.

[7] *D. W. Hoffmann*, On 6-dimensional quadratic forms isotropic over the function field of a quadric, Comm. Alg. 22 (1994), 1999–2014.

[8] *D. W. Hoffmann*, Isotropy of quadratic forms over the function field of a quadric, Math. Z. 220 (1995), 461–476.

[9] *D. W. Hoffmann*, Twisted Pfister forms, Doc. Math. J. DMV 1 (1996), 67–102.

[10] *J. Hurrelbrink, U. Rehmann*, Splitting patterns of excellent quadratic forms, J. reine angew. Math. 444 (1993), 183–192.

[11] *J. Hurrelbrink, U. Rehmann*, Splitting patterns and linear combinations of two Pfister forms, J. reine angew. Math. 495 (1998), 163–174.

[12] *O. T. Izhboldin, N. A. Karpenko*, Isotropy of virtual Albert forms over function fields of quadrics, Math. Nachr. 206 (1999), 111-122.

[13] *O. T. Izhboldin, N. A. Karpenko*, Isotropy of 6-dimensional quadratic forms over function fields of quadrics, J. Algebra 209 (1998), 65-93.

[14] *W. Jacob, M. Rost*, Degree four cohomological invariants for quadratic forms, Invent. Math. 96 (1989), 551–570.

[15] *B. Kahn*, Les formes quadratiques de hauteur et de degré 2, Indag. Mathem. 7 (1), 47–66 (1996).

[16] *B. Kahn*, A descent problem for quadratic forms, Duke Math. J. 80.1, 139–159 (1995).

[17] *B. Kahn, M. Rost et R. Sujatha,* Unramified cohomology of quadrics, I, Amer. J. Math. 120 (1998), 841–891.

[18] *B. Kahn, R. Sujatha,* Motivic cohomology and unramified cohomology of quadrics, J. Eur. Math. Soc. 2 (2000), 145–177.

[19] *B. Kahn, R. Sujatha,* Unramified cohomology of quadrics, II, Duke Math. J. 106.3 (2001), 449–484.

[20] *M. Knebusch*, Generic splitting of quadratic forms I, Proc. London. Math. Soc. 33 (1976) 65–93.

[21] *M. Knebusch*, Generic splitting of quadratic forms II, Proc. London. Math. Soc. 34 (1977) 1–31.

[22] *A. Laghribi*, Isotropie de certaines formes quadratiques de dimensions 7 et 8 sur le corps des fonctions d'une quadrique, Duke Math. J. 85.2 (1996), 397–410.

[23] *A. Laghribi*, Formes quadratiques en 8 variables dont l'algèbre de Clifford est d'indice 8, K-Theory J. 12.4 (1997), 371–383.

[24] *A. Laghribi*, Formes quadratiques de dimension 6, Math. Nach. 204 (1999), 125–135.

[25] *T. Y. Lam*, The algebraic theory of quadratic forms, (2^e édition) Benjamin, New York, 1980.

[26] *T. Y. Lam*, Fields of u-invariant 6 after A. S. Merkurjev, Israel Math. Conf. Proc. Vol. 1: Ring Theory 1989 (in honor of S. A. Amitsur), (L. Rowen, ed.), Weismann Science Press, Jerusalem, 1989, pp. 12–31.

[27] *D. Leep*, Function fields results, notes manuscrites prises par T. Y. Lam, 1989.

[28] *A. S. Merkurjev*, L'homomorphisme de résidu normique de degré 2 (en russe), Dokl. Akad. Nauk SSSR 261 (1981), 542–547. Traduction anglaise : Soviet Math. Doklady 24 (1981), 546–551.

[29] *A. S. Merkurjev, A. A. Suslin*, L'homomorphisme de résidu normique de degré 3 (en russe), Izv. Akad. Nauk SSSR 54 (1990), 339–356. Traduction anglaise: Math. USSR Izv. 36 (1991), 349–367.

[30] F. Morel, *Suite spectrale d'Adams et invariants cohomologiques des formes quadratiques,* C. R. Acad. Sci. Paris 328 (1999), 963–968.

[31] *D. Orlov, A. Vishik, V. Voevodsky:* An exact sequence for Milnor's K-theory with applications to quadratic forms, http://www.math.uiuc.edu/K-theory/454/.

[32] *E. Peyre*, Products of Severi–Brauer varieties and Galois cohomology, Proceedings of Symposia in Pure Mathematics, AMS Summer Research Institute, Santa Barbara, 1992. Volume 58.2, 369–401.

[33] *M. Rost*, Hilbert's theorem 90 for K_3^M for degree-two extensions, prépublication, Regensburg, 1986.

[34] *M. Rost*, On quadratic forms isotropic over the function field of a conic, Math. Ann. 288 (1990), 511–513.

[35] *D. Shapiro*, Similarities, quadratic forms, and Clifford algebras, PhD Thesis, California University, Berkeley, 1974.

[36] *W. Scharlau*, Quadratic and Hermitian forms, Springer, Berlin, 1985.

[37] *M. Szyjewski*, The fifth invariant of quadratic forms, (in Russian), Algebra i Analiz. 2 (1990), 213–234. English translation: St. Petersburg Math. J. 2 (1991), 179–198.

[38] *V. Voevodsky*, The Milnor conjecture, http://www.math.uiuc.edu/K-theory/0170/.

[39] *A. R. Wadsworth,* Noetherian pairs and function fields of quadratic forms, PhD Thesis, Chicago University, 1972.

Ahmed Laghribi
Faculté Jean Perrin
Rue Jean Souvraz - SP18
62307 Lens Cedex
France
laghribi@agel.ucl.ac.be

A Weak Hasse Principle for Central Simple Algebras with an Involution

David W. Lewis, Claus Scheiderer, Thomas Unger

Received: May 28, 2001

Communicated by Ulf Rehmann

Abstract. The notions of totally indefinite and weakly isotropic algebras with involution are introduced and a proof is given of the fact that a field satisfies the Effective Diagonalization Property (ED) if and only if it satisfies the following weak Hasse principle: every totally indefinite central simple algebra with involution of the first kind over the given field is weakly isotropic. This generalizes a known result from quadratic form theory.

2000 Mathematics Subject Classification: 16K20, 11E39, 12J15
Keywords and Phrases: Real fields, central simple algebras, involutions, weak Hasse principles, hermitian squares

1. Introduction

Let F be a field of characteristic different from two and let (A, σ) be a central simple algebra over F with involution of the first kind (i.e. $\sigma|_F = \mathbf{1}_F$). Recall that σ is called *orthogonal* (resp. *symplectic*) if σ is adjoint to a symmetric (resp. skew-symmetric) bilinear form, after scalar extension to a splitting field of A.

The connection between orthogonal involutions and quadratic forms has been a motivation for extending quadratic form theoretic concepts and theorems to the realm of algebras with involution (of any kind). For example, the classical invariants (discriminant, Clifford algebra, signature) of quadratic forms have been defined for algebras with involution (see [7]) and classification theorems à la Elman and Lam [5] have been obtained by Lewis and Tignol [14]. Some more examples include: a Cassels-Pfister theorem [19], an orthogonal sum for Morita-equivalent algebras with involution [2] and analogues of the Witt ring [12, 3].

In this paper we will examine the extension to central simple algebras with an involution of the first kind of the following weak Hasse principle for weak isotropy:

(WH): *Every totally indefinite quadratic form over F is weakly isotropic* and prove an analogue of the following theorem due to Prestel [15] and Elman et al. [4]:

Theorem 1.1. *F satisfies* (WH) $\iff$ *F satisfies the Strong Approximation Property.*

In particular, we will show that every totally indefinite central simple F-algebra with involution of the first kind is weakly isotropic if and only if F satisfies the Effective Diagonalization Property. Our result can also be re-interpreted to give a partial generalization of a theorem of Lewis [11] on sums of squares representing zero in a central simple algebra.
We mention that there is a refined (and more difficult) version of Theorem 1.1, which holds for arbitrary base fields, due to Bröcker [1, 3.9] and Prestel [16, p. 93]. It says that if ϕ is a totally indefinite quadratic form over a field F, and if for every valuation with real residue class field, at least one residue class form of ϕ is weakly isotropic, then ϕ is weakly isotropic. This statement can also be found in [18, 3.7.12]. Its converse is easily seen to be true.
All involutions on central simple algebras considered in this paper are of the first kind and all forms (quadratic, hermitian, ...) are assumed to be nonsingular. Standard references are [8] and [18] for the theory of quadratic forms, [7] for central simple algebras with an involution and [16] for real fields.

2. Weakly isotropic and totally indefinite algebras

In this section we will generalize the notions of totally indefinite and weakly isotropic quadratic forms to the setting of central simple algebras (A, σ) with an involution of the first kind over a field F of characteristic $\neq 2$. We denote the space of orderings of F by X_F and an arbitrary ordering of F by P.

Definition 2.1. Let (A, σ) be a central simple F-algebra with involution of the first kind. A right ideal I in A is called *isotropic* (with respect to the involution σ) if for all x and y in I we have that $\sigma(x)y = 0$. The algebra with involution (A, σ) is called *isotropic* if A contains a nonzero isotropic right ideal, or equivalently, if there exists an idempotent $e \neq 0$ in A such that $\sigma(e)e = 0$ (see [7, 6.A]). We also say that (A, σ) is *anisotropic* if for $x \in A$, $\sigma(x)x = 0$ implies $x = 0$.

Recall that a quadratic form q over F is weakly isotropic if there exists an $n \in \mathbb{N}$ such that $n \times q$ is isotropic.

Definition 2.2. The algebra with involution (A, σ) is called *weakly isotropic* if there exist nonzero $x_1, \dots, x_n \in A$ such that $\sigma(x_1)x_1 + \cdots + \sigma(x_n)x_n = 0$ and *strongly anisotropic* otherwise.

Remark 2.3. In [21] an n-fold orthogonal sum $\boxplus^n(A, \sigma)$ is defined and it is shown there that $\boxplus^n(A, \sigma) \cong (M_n(F), t) \otimes_F (A, \sigma)$, where t denotes the transpose involution. This is on the one hand in accordance with Dejaiffe's [2] construction of an orthogonal sum of two Morita-equivalent algebras with involution and on the other hand what one would expect since $n \times q = \underbrace{\langle 1, \dots, 1\rangle}_{n\times} \otimes q$ and $\boxplus^n(A,\sigma)$ reduces to $n \times q$ in the split case when σ is adjoint to a quadratic form q. In analogy with the quadratic form case, one could define (A,σ) to be weakly isotropic by requiring that $\boxplus^n(A,\sigma)$ is isotropic for some positive integer n and it is easy to see that this condition is equivalent with the one given in Definition 2.2.

Let (D,ϑ) be a central division algebra over F with involution of the first kind and (V,h) an ε-hermitian form over (D,ϑ), $\varepsilon = \pm 1$. Recall [7, 4.A] that the adjoint involution σ_h of h on $\mathrm{End}_D(V)$ is implicitly defined by

$$h\bigl(x, f(y)\bigr) = h\bigl(\sigma_h(f)(x), y\bigr) \quad \text{for } x, y \in V \text{ and } f \in \mathrm{End}_D(V)$$

and that σ_h is also of the first kind.
Just as for quadratic forms, we say that the ε-hermitian form h is weakly isotropic if there exists a positive integer n such that $n \times h$ is isotropic.

Lemma 2.4. *Let (D,ϑ), (V,h) and σ_h be as above. Then $(\mathrm{End}_D(V),\sigma_h)$ is weakly isotropic if and only if h is weakly isotropic. More precisely, there exist $f_1,\dots,f_n \in \mathrm{End}_D(V)$ such that $\sigma_h(f_1)f_1 + \cdots + \sigma_h(f_n)f_n = 0$ if and only if there exist $x_1,\dots,x_n \in V$ such that $h(x_1,x_1) + \cdots + h(x_n,x_n) = 0$.*

Proof. The lemma is folklore, and we only give the argument since we couldn't find a suitable reference. It suffices to show that $(\mathrm{End}_D(V),\sigma_h)$ is isotropic if and only if h is isotropic.
If σ_h is isotropic, there is $0 \neq f \in \mathrm{End}_D(V)$ with $\sigma_h(f)f = 0$. Choose $v \in V$ with $f(v) \neq 0$. Then

$$0 = h\bigl(\sigma_h(f)(f(v)),\, v\bigr) = h\bigl(f(v),\, f(v)\bigr)$$

shows that h is isotropic. Conversely, if $h(v,v) = 0$ for some $v \in V$, then $\sigma_h(f)f = 0$ for any $f \in \mathrm{End}_D(V)$ with $f(V) \subset vD$. ∎

Corollary 2.5. *Let $(\mathrm{End}_F(V),\sigma_q)$ be a split algebra with involution, adjoint to a quadratic form q on V. Then there exist $f_1,\dots,f_n \in \mathrm{End}_F(V)$ such that $\sigma_q(f_1)f_1 + \cdots + \sigma_q(f_n)f_n = 0$ if and only if there exist $x_1,\dots,x_n \in V$ such that $q(x_1) + \cdots + q(x_n) = 0$.*

Now suppose that F is a real field and that P is an ordering of F. In [13] Lewis and Tignol defined the signature of an algebra (A,σ) with involution of the first kind as

$$\mathrm{sig}_P\,\sigma = \sqrt{\mathrm{sig}_P\,T_\sigma},$$

where T_σ is the involution trace form, defined by $T_\sigma(x) := \mathrm{Trd}_A(\sigma(x)x), \forall x \in A$. If (A,σ) is split with orthogonal involution, $(A,\sigma) \cong (\mathrm{End}_F(V),\sigma_q)$, then Lewis and Tignol showed that $\mathrm{sig}_P\,\sigma_q = |\,\mathrm{sig}_P\,q|$.
Recall that a quadratic form q over F is called totally indefinite if it is indefinite for each ordering P of F, i.e. $|\,\mathrm{sig}_P\,q| < \dim q$ for each P.

Definition 2.6. The algebra with involution (A, σ) is called *indefinite* for the ordering P of F if $\operatorname{sig}_P \sigma < \deg A$ and *totally indefinite* if it is indefinite for each ordering P of F.

3. The weak Hasse principle

We now have all the ingredients ready to generalize (WH) to:

(WHA): *Every totally indefinite algebra with involution of the first kind over F is weakly isotropic.*

In [20, Ch. 5] Unger showed that (WHA) holds for fields with a unique ordering, algebraic number fields and $\mathbb{R}(t)$. These fields are some of the standard examples of SAP fields, as described below.

Definition 3.1. The field F satisfies the *Strong Approximation Property* (or is SAP, for short) if the following equivalent conditions hold:

(i) Every clopen subset of X_F has the form $\{P \in X_F \mid a >_P 0\}$ for some $a \in F^\times$.
(ii) For all $a, b \in F^\times$ the quadratic form $\langle 1, a, b, -ab\rangle$ is weakly isotropic.
(iii) Every quadratic form q such that a power of q is weakly isotropic, is itself weakly isotropic.
(iv) For any two disjoint closed subsets X, Y of X_F, there exists an $a \in F^\times$ such that $a >_P 0, \forall P \in X$ and $a <_P 0, \forall P \in Y$.
(v) For every (Krull) valuation $v : F^\times \twoheadrightarrow \Gamma$ with value group Γ and real residue class field $\overline{F}_v$, either (a) or (b) holds:
  (a) $\Gamma = 2\Gamma$;
  (b) $|\Gamma/2\Gamma| = 2$ and $\overline{F}_v$ has a unique ordering.

Condition (iv) is the original definition of SAP fields, due to Knebusch et al. [6, Thm. 12]. The equivalence (i) $\iff$ (iv) is given in [6, Thm. 12, Cor. 13]. Prestel [15, (2.2), (3.1)] showed (ii) $\iff$ (iii) $\iff$ (v) $\iff$ F is a Pasch field, while the equivalence F is SAP $\iff$ F is Pasch can be found in [4, Thm. C]. The notion of a Pasch field was first introduced by Prestel; for a definition we refer the reader to [15]. Additional references for SAP fields are the monographs by Lam [9] and Prestel [16].

Example 3.2. Here are some examples of SAP fields:

(1) Fields with only one ordering.
(2) Algebraic number fields.
(3) Fields of transcendence degree ≤ 1 over a real-closed field, e.g. $\mathbb{R}(t)$.
(4) $F((t))$ if F has at most one ordering.

The following fields are not SAP:

(5) The rational function field $\mathbb{Q}(x)$.
(6) The rational function field $F(x, y)$, where F is any real field.

Based on the results in [20, Ch. 5] it was tempting to think that (WHA) would hold for all SAP fields. The Strong Approximation Property is definitely required, for if F is not SAP, we can construct a counterexample as follows:

There exist $a, b \in F^\times$ such that $q := \langle 1, a, b, -ab\rangle$ is strongly anisotropic. Hence the algebra $(A, \sigma) = (\mathrm{End}_F(F^4), \sigma_q)$ is strongly anisotropic by Corollary 2.5. However, the form $T_\sigma = q \otimes q$ is equal to

$$\begin{aligned} q \otimes q &= q \perp aq \perp bq \perp -abq \\ &= \langle 1, a, b, -ab, a, 1, ab, -b, b, ab, 1, -a, -ab, -b, -a, 1\rangle \\ &= 6 \times \langle 1, -1\rangle \perp \langle 1, 1, 1, 1\rangle, \end{aligned}$$

so T_σ is in fact isotropic and hence totally indefinite. Therefore the orthogonal involution σ is totally indefinite.

For quadratic forms, this argument was of course already known in the 1970's, as testified by Theorem 1.1. We merely presented it from the point of view of algebras with involution.

A counterexample in the symplectic case can be constructed by tensoring the previous algebra with the quaternion division algebra $(-1, -1)_F$ equipped with the canonical (symplectic) involution, which is strongly anisotropic.

As it turns out, a property stronger than SAP is needed, the Effective Diagonalization Property, first defined by Ware [22], which we will describe now.

Definition 3.3. A quadratic form $\langle a_1, \dots, a_n\rangle$ is *effectively diagonalizable* if it is isometric to a form $\langle b_1, \dots, b_n\rangle$ satisfying $b_i \in P \implies b_{i+1} \in P$ for all $1 \le i < n$ and all $P \in X_F$. The field F satisfies the *Effective Diagonalization Property* (or is ED, for short) if every quadratic form over F is effectively diagonalizable.

The class of ED fields is a proper subclass of the class of SAP fields.

Example 3.4. The field $\mathbb{Q}((t))$ is SAP, but not ED.

Prestel and Ware [17] proved the following characterization theorem:

Theorem 3.5. *F is* ED *if and only if for every (Krull) valuation $v : F^\times \twoheadrightarrow \Gamma$ with value group Γ and real residue class field $\overline{F}_v$, we have $|\Gamma/2\Gamma| \le 2$ and $\overline{F}_v$ is euclidean in case $|\Gamma/2\Gamma| = 2$.*

(Recall that a field is *euclidean* if it is uniquely ordered and every positive element is a square.) They also showed:

Theorem 3.6. *If F is* ED *then every 2-extension of F is also* ED*. (In particular, the pythagorean closure of F is* ED*.)*

(Recall that an extension K of F is called a 2-*extension* of F if K is contained in the quadratic closure of F.)

Remark 3.7. The ED property also played an important role in the classification theorems of Lewis and Tignol [14].

Our generalization of Theorem 1.1 reads:

Theorem 3.8. *F is* ED $\iff$ *F satisfies* (WHA).

The proof will follow from the results below (Theorems 3.11 and 3.12).

Lemma 3.9. *Let (A,σ) be a central simple algebra with involution of the first kind over F. Let $d \in F$ be a sum of squares, and let $K = F(\sqrt{d})$. Suppose that $(A \otimes_F K, \sigma_K)$ is weakly isotropic. Then (A,σ) is weakly isotropic.*

Proof. We may assume that $(A \otimes_F K, \sigma_K)$ is isotropic. Hence there exist $x, y \in A$, not both zero such that

$$\big(\sigma(x) + \sigma(y)\sqrt{d}\big)\big(x + y\sqrt{d}\big) = 0.$$

Separating, this implies $\sigma(x)x + d\sigma(y)y = 0$. Suppose $d = d_1^2 + \cdots + d_r^2$ with $d_i \in F$, then

$$\sigma(x)x + \sigma(d_1 y)(d_1 y) + \cdots + \sigma(d_r y)(d_r y) = 0,$$

i.e. (A,σ) is weakly isotropic. ∎

Lemma 3.10. *Let F be a pythagorean* SAP *field and A a central simple algebra of exponent 2 over F. Then A is Brauer-equivalent to a quaternion division algebra $(-1, f)_F$ for some $f \in F^\times$.*

Proof. By a well-known theorem of Merkurjev, A is Brauer-equivalent to a tensor product of finitely many quaternion division algebras over F. Without loss of generality, we may assume that A is Brauer-equivalent to $(a,b)_F \otimes_F (a',b')_F$ for certain $a, a', b, b' \in F^\times$, and that $(a,b)_F$ and $(a',b')_F$ do not split. Since $(a,b)_F$ is a division algebra, its norm form $\langle 1, -a, -b, ab\rangle$ is anisotropic. Hence $\langle a, b, -ab\rangle$ is anisotropic. Since F is SAP, the quadratic form $\langle 1, a, b, -ab\rangle$ is weakly isotropic, and thus isotropic, since F is pythagorean. Hence $\langle 1, a, b, -ab\rangle \simeq \langle 1, -1, c, d\rangle$ for certain $c, d \in F^\times$. Comparing determinants, we get $\langle 1, a, b, -ab\rangle \simeq \langle 1, -1, c, c\rangle$, which implies $\langle a, b, -ab\rangle \simeq \langle -1, c, c\rangle$, and thus $(a,b)_F \cong (-1, c)_F$.
Similarly, $(a',b')_F \cong (-1, c')_F$ for some $c \in F^\times$, and so A is Brauer-equivalent to $(-1, cc')_F$. ∎

Theorem 3.11. *Assume that F is* ED*, then F satisfies* (WHA).

Proof. Let (A,σ) be totally indefinite. We will show that (A,σ) is weakly isotropic.
If A is split, the theorem is true by Theorem 1.1 (when σ is orthogonal) or trivial (when σ is symplectic).
If the degree of A is odd, then A is split and σ is orthogonal. So we are done in this case. Hence we may assume that A is not split and $\deg A = n = 2m$ is even.
Since F is ED, its pythagorean closure is ED. By Lemma 3.9 we may replace F by its pythagorean closure. (The pythagorean closure F_{pyth} is in general an infinite extension of F but, for any given algebra A, we only need to pass to a finite extension of F, sitting inside F_{pyth}. Then we apply Lemma 3.9 finitely many times.) So we assume from now on that F is a pythagorean ED field.

By Lemma 3.10, A is Brauer-equivalent to a quaternion division algebra $D := (-1, f)_F$ for some $f \in F^\times$. So now we have

$$(A, \sigma) \cong (\mathrm{End}_D(D^m), \sigma_h) \cong (M_m(D), \sigma_h),$$

where σ_h is the adjoint involution of a form $h : D^m \times D^m \longrightarrow D$, which is hermitian or skew-hermitian with respect to quaternion conjugation $^-$ on D, according to whether σ is symplectic or orthogonal.

Suppose first that σ is symplectic, so that h is hermitian. By [18, 7.6.3] there exists a basis $\{e_1, \dots, e_m\}$ of D^m over D which is orthogonal with respect to h. Let $\lambda_i = h(e_i, e_i)$ for $i = 1, \dots, m$. Lewis and Tignol [13, Cor. 2] showed that $\lambda_i \in F$ for all $i = 1, \dots, m$ and that

$$T_\sigma = \langle 2 \rangle \otimes N \otimes \Lambda \otimes \Lambda$$

where N is the norm form of D and $\Lambda = \langle \lambda_1, \dots, \lambda_m \rangle$. By assumption T_σ is totally indefinite, and hence weakly isotropic. Then $N \otimes \Lambda \otimes \Lambda$ is weakly isotropic and so $N \otimes N \otimes \Lambda \otimes \Lambda$ is weakly isotropic. Since F is SAP, this implies (by Definition 3.1(iii)) that $N \otimes \Lambda$ is weakly isotropic. Since $h(x, x) = \lambda_1 N(x_1) + \cdots + \lambda_m N(x_m)$ for $x = (x_1, \dots, x_m) \in D^m$, this implies that the hermitian form h is weakly isotropic over D and hence that (A, σ) is weakly isotropic.

Suppose next that σ is orthogonal, so that h is skew-hermitian. Put $K = F(\sqrt{f})$ (note that f is not a square, since D is a division algebra). Over K, the algebra A splits. Since (A, σ) is totally indefinite, it is clear that $(A, \sigma) \otimes_F K$ is also totally indefinite. Being a 2-extension of the ED field F, the field K is SAP and, by Theorem 1.1, it follows that $(A, \sigma) \otimes_F K$ is weakly isotropic, since $A \otimes_F K$ is split. This implies that the skew-hermitian form h becomes weakly isotropic over K (i.e. as a form over $D_K \cong M_2(K)$). ¿From this we will now deduce that that the form h itself is weakly isotropic, i.e. that the algebra (A, σ) is weakly isotropic.

Replacing h by $N \times h$ for $N \gg 0$ if necessary, there are $x, y \in D^m$, not both zero, such that $h_K(x + y\sqrt{f}, x + y\sqrt{f}) = 0$. This implies $h(x, x) + f h(y, y) = 0$. If $h(y, y) = 0$, it follows that h is (weakly) isotropic and we are done. Otherwise, $u := h(y, y) \in D$ is a non-zero pure quaternion (since h is skew-hermitian) and h has a diagonalization

$$h \simeq \langle -fu, \dots \rangle.$$

Now let $d := u^2 = -\mathrm{Nrd}(u) \in F^\times$. Then $d < 0$ on $\{P \in X_F | f <_P 0\}$ and therefore $D_{F_P} \cong (d, f)_{F_P}$ for all orderings $P \in X_F$ (here F_P denotes the real closure of F with respect to P). Hence $D \cong (d, f)_F$ by Pfister's local-global principle (note that the Witt ring of F is torsion free, since F is pythagorean; see [18, 2.4.10–11]), and there exists a pure quaternion $v \in D$ with

$$v^2 = -\mathrm{Nrd}(v) = f \quad \text{and} \quad uv + vu = 0.$$

Thus $\overline{v} = \mathrm{Nrd}(v) v^{-1} = -f v^{-1}$, and so

$$\overline{v} u v = -f v^{-1} u v = -f(-u) = fu.$$

Therefore $h(yv, yv) = \overline{v}uv = fu$ and h also has a diagonalization

$$h \simeq \langle fu, \ldots \rangle.$$

This implies that $h \perp h \simeq \langle -fu, fu, \ldots \rangle$, which is isotropic. In other words, h is weakly isotropic and hence σ is weakly isotropic. We are done. ∎

Theorem 3.12. *For any non-*ED *field F, there is an algebra (A, σ) with involution of the first kind (and of either type) over F which is strongly anisotropic but totally indefinite.*

Proof. The statement is clear if the field is not SAP (there is an involution which is totally indefinite and strongly anisotropic, as explained just after Example 3.2), so we concentrate on the case of a SAP field which is not ED.
Let F be such a field. Then F has a (Krull) valuation v whose value group Γ satisfies $\Gamma/2\Gamma = \mathbb{Z}/2\mathbb{Z}$, and whose residue field $\overline{F}_v$ is real without being euclidean (this follows from Definition 3.1(v) and Theorem 3.5). Let $\pi \in F^\times$ with $v(\pi) \notin 2\Gamma$, and let $a \in F^\times$ be a v-unit whose residue class in $\overline{F}_v$ is a sum of squares but not a square. We choose a such that a is a sum of squares in F, and consider the quaternion (division) algebra $A = (a, \pi)$ over F. Let $1, i, j, k = ij$ be the standard F-basis of A, satisfying $i^2 = a$, $j^2 = \pi$ and $k^2 = -a\pi$. Let h be the (diagonal) skew-hermitian form $h = \langle j, k \rangle$ over $(A, ^-)$ (where $^-$ denotes the standard (symplectic) involution on A), and let σ be the adjoint involution of h on $M_2(A)$. We claim that σ is totally indefinite, but not weakly isotropic. To show this, let $L = F(\sqrt{a})$. We fix the splitting $\phi\colon A_L \xrightarrow{\sim} M_2(L)$ over L given by

$$i \mapsto \begin{pmatrix} \sqrt{a} & 0 \\ 0 & -\sqrt{a} \end{pmatrix} \quad \text{and} \quad j \mapsto \begin{pmatrix} 0 & 1 \\ \pi & 0 \end{pmatrix}.$$

Under ϕ, h corresponds to a similarity class of quadratic forms q (of rank 4) over L. We are going to calculate q.
For $x \in A^\times$ with $x + \overline{x} = 0$, let σ_x be the (orthogonal) involution on A given by $\sigma_x(z) = x^{-1}\overline{z}x$. Under ϕ, $\sigma_x \otimes \mathbf{1}$ corresponds to a similarity class of quadratic forms q_x over L. Writing $J := \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ we have $q_x = J \cdot \phi(x)$. In particular, taking $x = j$ and $x = k$, we find

$$q_j = \langle -\pi, 1 \rangle \quad \text{and} \quad q_k = \langle \pi\sqrt{a}, \sqrt{a} \rangle.$$

Thus

$$q \simeq \langle 1, \sqrt{a}, -\pi, \pi\sqrt{a} \rangle.$$

The form q is totally indefinite. Since the extension L/F is totally real, the involution trace form T_σ (over F) is totally indefinite as well and hence σ is totally indefinite.
On the other hand, the residue forms of q with respect to v are $\langle 1, \sqrt{\overline{a}} \rangle$ and $\langle -1, \sqrt{\overline{a}} \rangle$ (note that $^-$ denotes taking residue classes here). Neither of them is totally indefinite. Hence q is strongly anisotropic. Therefore σ cannot be weakly isotropic.

The symplectic case can be treated again by tensoring our algebra with the quaternion division algebra $(-1,-1)_F$, equipped with quaternion conjugation. ■

Putting everything together now yields a proof of Theorem 3.8.

4. Sums of hermitian squares

In [11], Lewis proved the following theorem, settling a conjecture of Leep et al. [10].

Theorem 4.1. *Let A be a central simple algebra over a field F of characteristic $\neq 2$. Then 0 is a nontrivial sum of squares, i.e. there exist nonzero $x_1, \ldots, x_\ell \in A$ such that $0 = x_1^2 + \cdots + x_\ell^2$, if and only if the trace form T_A is weakly isotropic.*

The natural adaptation of this theorem in the setting of algebras with involution of the first kind is an easy consequence of the work we have done hitherto:

Definition 4.2. Let (A, σ) be a central simple algebra with involution of the first kind over a field F and $x \in A$. Then $\sigma(x)x$ is called a *hermitian square* in A.

Theorem 4.3. *Let (A, σ) be a central simple algebra with involution of the first kind over an* ED*-field F. Then 0 is a nontrivial sum of hermitian squares, i.e. there exist nonzero $x_1, \ldots, x_\ell \in A$ such that $0 = \sigma(x_1)x_1 + \cdots + \sigma(x_\ell)x_\ell$, if and only if the involution trace form T_σ is weakly isotropic.*

Proof. The necessary condition follows trivially (and does not require ED) by simply taking the reduced trace of both sides of $0 = \sum_{i=1}^{\ell} \sigma(x_i)x_i$.
For the sufficient condition, suppose that T_σ is weakly isotropic. Then T_σ, and hence σ, is totally indefinite. Therefore σ is weakly isotropic (since F is ED), i.e. there exist nonzero $x_1, \ldots, x_\ell \in A$ such that $\sigma(x_1)x_1 + \cdots + \sigma(x_\ell)x_\ell = 0$. ■

Remark 4.4. For several special classes of algebras with involution of the first kind, the condition on F can be relaxed and the conclusion of Theorem 4.3 will still hold. This happens for example

(1) when (A, σ) is an algebra of index 2 with symplectic involution over a SAP field F;
(2) when (Q, σ) is a quaternion algebra with involution of the first kind over a field F of characteristic not 2;
(3) when $(A, \sigma) \cong (Q_1, \sigma_1) \otimes_F \cdots \otimes_F (Q_\ell, \sigma_\ell)$ is a multi-quaternion algebra over a field F of characteristic not 2 and each σ_i is an arbitrary involution of the first kind.

For proofs, see [20, Ch. 5].

Finally, we obtain a version of Springer's theorem for strongly anisotropic involutions:

Corollary 4.5. *Let (A, σ) be a central simple algebra with involution of the first kind over an* ED*-field F and let K/F be any finite extension of odd degree. If (A, σ) is strongly anisotropic, then $(A \otimes_F K, \sigma_K)$ is (strongly) anisotropic.*

Proof. Since σ is strongly anisotropic, T_σ is strongly anisotropic by Theorem 4.3. By Springer's theorem (see e.g. [18, 2.5.3]), $(T_\sigma)_K = T_{\sigma_K}$ is strongly anisotropic over K. Hence σ_K is strongly anisotropic by contraposition of the trivial direction of Theorem 4.3. ∎

Acknowledgements

This research was supported by the TMR research network (ERB FMRX CT-97-0107) on "K-theory and algebraic groups".

References

[1] L. Bröcker, Zur Theorie der quadratischen Formen über formal reellen Körpern, *Math. Ann.* 210 (1974), 233–256.

[2] I. Dejaiffe, Somme orthogonale d'algèbres à involution et algèbre de Clifford, *Comm. Algebra* 26 (1998), 1589–1612.

[3] I. Dejaiffe, D. W. Lewis, J. P. Tignol, Witt equivalence of central simple algebras with involution, *Rend. Circ. Mat. Palermo (2)* 49(2) (2000), 325–342.

[4] R. Elman, T.Y. Lam, A. Prestel, On some Hasse principles over formally real fields, *Math. Z.* 134 (1973), 291–301.

[5] R. Elman, T. Y. Lam, Classification theorems for quadratic forms over fields, *Comment. Math. Helv.* 49 (1974), 373–381.

[6] M. Knebusch, A. Rosenberg, R. Ware, Structure of Witt rings, quotients of abelian group rings, and orderings of fields, *Bull. Amer. Math. Soc.* 77 (1971), 205–210.

[7] M.-A. Knus, A.S. Merkurjev, M. Rost, J.-P. Tignol, *The Book of Involutions*, Coll. Pub. 44, Amer. Math. Soc., Providence, RI (1998).

[8] T.Y. Lam, *The Algebraic Theory of Quadratic Forms*, Reading, Mass. (1973).

[9] T.Y. Lam, *Orderings, Valuations and Quadratic Forms*, CBMS Notes 52, Amer. Math. Soc. (1983).

[10] D.B. Leep, D.B. Shapiro, A.R. Wadsworth, Sums of squares in division algebras, *Math. Z.* 190 (1985), 151–162.

[11] D.W. Lewis, Sums of squares in central simple algebras, *Math. Z.* 190 (1985), 497–498.

[12] D.W. Lewis, The Witt semigroup of central simple algebras with involution, *Semigroup Forum* 60 (2000), 80–92.

[13] D.W. Lewis, J.-P. Tignol, On the signature of an involution, *Arch. Math.* 60 (1993), 128–135.

[14] D.W. Lewis, J.-P. Tignol, Classification theorems for central simple algebras with involution, *Manuscripta Math.* 100(3) (1999), 259–276. With an appendix by R. Parimala.

[15] A. Prestel, Quadratische Semi-Ordnungen und quadratische Formen, *Math. Z.* 133 (1973), 319–342.

[16] A. Prestel, *Lectures on Formally Real Fields*, Lecture Notes in Mathematics 1093, Springer-Verlag, Berlin (1984).

[17] A. Prestel, R. Ware, Almost isotropic quadratic forms, *J. London Math. Soc.* (2) 19 (1979), 214–244.

[18] W. Scharlau, *Quadratic and Hermitian Forms*, Grundlehren Math. Wiss. 270, Springer-Verlag, Berlin (1985).

[19] J.-P. Tignol, A Cassels-Pfister theorem for involutions on central simple algebras, *J. Algebra* 181(3) (1996), 857–875.

[20] T. Unger, *Quadratic Forms and Central Simple Algebras with Involution*, Ph.D. Thesis, National University of Ireland, Dublin (2000), unpublished.

[21] T. Unger, Clifford algebra periodicity for central simple algebras with an involution, *Comm. Algebra* 29(3) (2001), 1141–1152.

[22] R. Ware, Hasse principles and the u-invariant over formally real fields, *Nagoya Math. J.* 61 (1976), 117–125.

David W. Lewis
Department of Mathematics
University College Dublin
Belfield, Dublin 4
Ireland
david.lewis@ucd.ie

Claus Scheiderer
Fachbereich Mathematik
Universität Duisburg
47048 Duisburg
Germany
claus@uni-duisburg.de

Thomas Unger
Department of Mathematics
University College Dublin
Belfield, Dublin 4
Ireland
thomas.unger@ucd.ie