Documenta Mathematica

Journal der Deutschen Mathematiker-Vereinigung

Band 1

1996

ISSN 1431-0635 Print ISSN 1431-0643 Internet

Documenta Mathematica, Journal der Deutschen Mathematiker-Vereinigung, veröffentlicht Forschungsarbeiten aus allen mathematischen Gebieten und wird in traditioneller Weise referiert.

Documenta Mathematica erscheint am World Wide Web unter der Adresse:

`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, Journal der Deutschen Mathematiker-Vereinigung, 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 the address:

`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 Documenta Mathematica (Print)
ISSN 1431-0643 Documenta Mathematica (Internet)

Anschrift des technischen geschäftsführenden Herausgebers:
Ulf Rehmann, Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, D-33501 Bielefeld

DOCUMENTA MATHEMATICA

JOURNAL DER DEUTSCHEN MATHEMATIKER-VEREINIGUNG

BAND 1, 1996

Étale Cohomology of Rigid Analytic Spaces

Johan de Jong[1] and Marius van der Put

Received: October 16, 1995

Communicated by Peter Schneider

Abstract. The paper serves as an introduction to étale cohomology of rigid analytic spaces. A number of basic results are proved, e.g. concerning cohomological dimension, base change, invariance for change of base fields, the homotopy axiom and comparison for étale cohomology of algebraic varieties. The methods are those of classical rigid analytic geometry and along the way a number of known results on rigid cohomology are re-established.

Key Phrases: "étale cohomology", "rigid analytic spaces", "rigid cohomology", "overconvergent sheaves"

1991 Mathematics Subject Classification: Primary 26E30, secondary 14F20.

1 Introduction

The origin of this paper lies in the questions on étale cohomology for rigid analytic spaces posed in [S-S]. In that paper an étale site and a corresponding cohomology theory for analytic varieties are defined. We prove here that the axioms for an 'abstract cohomology' (as stated in [S-S]) hold for this cohomology theory. In addition, we prove a (quasi-compact) base change theorem for rigid étale cohomology and a comparison theorem comparing rigid and algebraic étale cohomology of algebraic varieties.

The main tools in this paper are analytic (resp. étale) points and rigid (resp. étale) overconvergent sheaves. The rigid overconvergent sheaves on affinoids were first introduced in [P82] and were called constructible in that paper. They were further studied in [S93] and were called conservative there. The term 'overconvergent', also used by P. Berthelot in recent work, seemed more appropriate this time.

In Section 2 we (re)introduce some basic notations concerning analytic points and rigid overconvergent sheaves, which are needed later on. We (re)prove a number of folklore results, most importantly: 1) Rigid cohomology agrees with Čech cohomology on quasi-compact spaces. 2) The cohomological dimension of a paracompact space

[1] The research of Dr. A.J. de Jong has been made possible by a fellowship of the Royal Netherlands Academy of Arts and Sciences.

is at most its dimension. 3) A base change theorem for rigid spaces which is more general than the results of [P82] or [S93].

The rest of the paper deals with étale sites and étale cohomology. Étale points and étale overconvergent sheaves are introduced. A key point is the introduction of special étale morphisms of affinoids $U \to X$, analogous to rational subdomains in the rigid case. Included in the paper is the proof by R. Huber that any étale morphism of affinoids is special étale. This simplifies the original exposition somewhat. A structure theorem for étale morphisms (3.1.2) allows us to give a proof of the étale base change theorem following closely the proof in the rigid case. We calculate the cohomology groups of one dimensional spaces in Section 4. This allows us to prove the basic results mentioned at the beginning of this introduction (Sections 5, 6 and 7).

We have tried to be complete in the proofs of various statements. We hope that this paper may serve as an introduction to rigid and étale cohomology of rigid analytic spaces.

Berkovich, in the paper [B93], develops an étale cohomology theory for analytic spaces. The category of analytic spaces used there was introduced in [B90] and extended in [B93]. It is different from the category of rigid analytic spaces. For this reason we have not borrowed from his work. However, we have to mention that the approach taken here, in some sense, does not differ from his (although in this paper we have to deal with non-overconvergent sheaves also, which do not correspond to sheaves on the Berkovich analytic spaces). For example, Lemma 2.1.1, which controls the étale stalk functors, is more or less equivalent to Theorems 2.1.5 & 2.3.3 of [B93]. Furthermore, using the equality of Berkovich cohomology with ours in the case of paracompact varieties (see [Hu, Section 8.3]), all our results on cohomology of overconvergent sheaves are in principle deducible from the references [B93, B94a, B94b, B94c].

Étale cohomology theories for rigid analytic spaces were developed by O. Gabber (unpublished) and K. Fujiwara, who proved Deligne's conjecture using his theory. As mentioned above R. Huber constructed an étale cohomology theory for his adic spaces, this specializes to give a theory for rigid analytic spaces also.

We thank P. Schneider for sending his informal notes [S91] to the authors for consultation.

1.1 Notations and conventions

- Unless stated otherwise k will be a complete non Archimedean valued field.
- As general reference for the basic facts and definitions concerning rigid analytic varieties we take [BGR].
- All rigid analytic varieties occurring in this work will be quasi-separated analytic varieties. This means that the diagonal morphism $X \to X \times X$ is quasi-compact, or equivalently that the intersection of any two affinoid subvarieties of X is a finite union of affinoid subvarieties of X. It is clear that fibre products of such are still quasi-separated.
- We work frequently with sites and associated topoi as in [SGA 4]. We recall that a morphism of sites $f : S_1 \to S_2$ is a continuous functor $u : S_2 \to S_1$ (remark that u goes in the opposite direction!), which induces a morphism of

associated topoi $S_1^\sim \to S_2^\sim$ (see [SGA 4, IV 4.9]). We remark that if S_2 allows finite projective limits then it suffices that u is continuous and preserves fibred products.

- A sheaf $\mathcal{F}$ on a site S is said to be flabby if for any object U in S we have $\mathrm{H}^q(U, \mathcal{F}) = 0$ for all $q > 0$. It is said to be flasque if for any morphism $U \to V$ the restriction map $\mathcal{F}(V) \to \mathcal{F}(U)$ is surjective. A flasque sheaf is flabby since Čech cohomology may be used to determine whether a sheaf is flabby ([M80, III 2.12]).

2 Analytic points and rigid overconvergent sheaves

In this section we will review the base change theorem for rigid analytic spaces (see [P82, S93]). We will introduce our basic notations and reprove the statements of [P82] (whose proofs are perhaps somewhat sketchy). We try to avoid using results from [B90] except for the basic fact that the space $\mathcal{M}(X)$ (see below) is Hausdorff and compact (this is not hard to prove). Finally, we prove a slightly stronger version of the base change theorem, namely that it holds for arbitrary sheaves.

2.1 Sites, sheaves and analytic points on affinoids

Let X be an affinoid space over some complete non Archimedean valued field k. On X we consider the *special Grothendieck topology* given by the collection of finite unions of open affinoid subspaces and the admissible coverings. (See [FP, GP], this is a G-topology slightly stronger than the weak G-topology of X in [BGR, 9.1.4].) We will write X_{rigid} for the following site:

1. The objects are the admissible open subsets of X. We choose here as admissible opens the finite unions of open affinoid subsets. These will also be called the special subsets of X.

2. A morphism between to objects is an inclusion between the admissible subsets.

3. For an object U the elements of $\mathrm{Cov}(U)$ are those set-theoretical coverings of U by admissible opens which can be refined to finite coverings.

We use the special G-topology rather than the strong G-topology since it behaves better with respect to base change and change of base field. We remark that this gives the same category of sheaves.

It is sometimes easier to work with a subcategory X_{rigid}^{rat} of X_{rigid}. The objects of X_{rigid}^{rat} are the rational subsets of X. A rational subset of X is a set of the form

$$\{x \in X | \, |f_1(x)| \geq |f_i(x)| \text{ for all } i \text{ with } 1 \leq i \leq n\}$$

where $f_1, \ldots, f_n$ are elements of $O(X)$ generating the unit ideal. We note that a small change of the $f_1, \ldots, f_n$ does not affect the subset above. It is known that every open affinoid subset of X is a finite union of rational subsets ([GG]). A *rational covering* of a rational $U \subset X$ is a covering of the form $U = \cup_{i=1}^m U_i$ given by elements $f_1, \ldots f_m \in O(U)$ generating the unit ideal such that the U_i are the rational subsets (of U and also of X) $U_i := \{x \in U | \, |f_i(x)| \geq |f_j(x)| \text{ for all } j\}$. This defines for every object the collection of coverings. The morphism of sites $X_{rigid} \to X_{rigid}^{rat}$ (given by the inclusion functor $X_{rigid}^{rat} \to X_{rigid}$, see our conventions) defines an isomorphism of associated topoi, this follows from the fact that any special subset of X is a finite union of rational subsets and any finite affinoid covering of an affinoid variety can be refined to a rational covering (see for example [BGR, 8.2.2/2]).

It is well known that the set of ordinary points of X is too small to "separate" the sheaves on X_{rigid}. For this purpose one introduces new points, called *analytic points.* (See [P82, S93]). We will adopt here the terminology of [S93].

An *analytic point* a of X is a semi-norm $|\ |_a : O(X) \to \mathbf{R}_{\geq 0}$ on the affinoid algebra $O(X)$ of X satisfying:

1. $|f+g|_a \leq \max(|f|_a, |g|_a)$ for all $f, g \in O(X)$.
2. $|fg|_a = |f|_a|g|_a$ for all $f, g \in O(X)$.
3. For $\lambda \in k$ the value $|\lambda|_a$ is the absolute value of λ.
4. $|\ |_a : O(X) \to \mathbf{R}_{\geq 0}$ is continuous with respect to the norm topology on $O(X)$.

The *filter of the analytic point* a consist of the affinoid subdomains U of X for which there exists a rational covering given by $f_1, ..., f_n$ and an i such that $U \supset U_i$ and $|f_i|_a \geq |f_j|_a$ for all j. This is equivalent with the property that $|\,|_a$ extends to a $|\,|_a : O(U) \to \mathbf{R}_{\geq 0}$, i.e., that a is also an analytic point of U. We write $a \in U$ to denote that U belongs to the filter of a. We will also need the concept of a *wide neighborhood* of an analytic point a of X (see [S93, p. 131]). An element U of the filter of a is a wide neighborhood of a if there exists an affinoid generating system $f_1, \ldots, f_n$ of $O(U)$ over $O(X)$ such that $|f_i|_a < 1$ for all i.

Let $\mathcal{M}(X)$ denote the set of analytic points of X. We give $\mathcal{M}(X)$ the coarsest topology such that for every $g \in O(X)$ the map $\mathcal{M}(Z) \to \mathbf{R}$ given by $a \mapsto |g|_a$ is continuous. For an analytic point a a fundamental system of neighborhoods is given by the subsets $\mathcal{M}(U)$ where U runs through the (affinoid) wide neighborhoods of a. The space $\mathcal{M}(Z)$ is Hausdorff and compact for this topology. These results are not hard to prove, they follow from 1.2.2 and 1.3.3 of [P82], but see [B90, §1], [S93, §1] for more details. We will repeatedly make use of the following corollary of the above: *Suppose that $\{X_i\}_{i\in I}$ are affinoid subdomains of X such that for any analytic point a of X some X_i is a wide neighborhood of a, then the covering $X = \bigcup X_i$ is admissible, i.e., finitely many of the X_i cover X.*

The stalk of a sheaf S on X_{rigid} at an analytic point a is defined as $S_a = \lim_{\to} S(U)$ where the direct limit is taken over all U in the filter of a. The modified stalk of S at a is $S_a^{mod} = \lim_{\to} S(U)$ where the limit is over the wide open neighborhoods of a in X.

For every U in the filter of a the semi-norm $|\ |_a$ extends to a semi-norm on $O(U)$. Hence we get a semi-norm $|\ |_a$ on O_a the stalk of $O = O_X$ at a. A fundamental fact that we will use is (see [P82, 1.3.1]) that for $f \in O(X)$:

$$|f|_a = \inf\{||f||_U\}$$

where U runs through the filter of a. In fact it suffices to consider only wide open neighborhoods of a (use that for $U \subset X$ rational we have $||f||_U = \inf_{r>1} ||f||_{U(r)}$

where $U(r)$ is defined as in 2.3 below). It follows from these considerations that the ideal m_a of elements $f \in O_a$ satisfying $|f|_a = 0$ is the unique maximal ideal of O_a (and similar for O_a^{mod}). The field O_a/m_a will be denoted by k_a. The semi-norm $|\ |_a$ induces a valuation on k_a. This valuation extends the valuation of the subfield k of k_a. In general the field k_a is not complete and its completion is denoted by F_a. (The same constructions give k_a^{mod} and F_a^{mod}.)

Let $\phi : O(X) \to F_a$ denote the continuous homomorphism of k-algebras obtained above from $|\ |_a$. Then one sees that $|f|_a = |\phi(f)|$. This remark shows that our definition of analytic point coincides with the equivalence classes of analytic points as defined in [S93]. Every ordinary point of X is also an analytic point (with $F_a = k_a$ a finite extension of k). The following lemma will be useful in our study of the étale site of X.

Lemma 2.1.1 *Notations are as above.*

1. *O_a and O_a^{mod} are Henselian local rings.*
2. *k_a and k_a^{mod} are Henselian valued fields.*
3. *F_a is finite over a complete subfield K which has a dense subfield $k(t_1, ..., t_d)$ with $d \leq$ the dimension of X.*
4. *The homomorphism $O_a^{mod} \to O_a$ is local, flat and induces an isomorphism $F_a^{mod} \cong F_a$.*

Proof. Let $O_a \subset A$ be a finite free extension of rings. We claim the following: the ring $A \hat{\otimes} F_a$ has a nontrivial idempotent if and only if A has one. (We also claim a similar result for O_a^{mod}.)

This immediately implies (1) (see [R70, I Proposition 5]). Statement (2) means that the valuation ring of k_a (resp. k_a^{mod}) is an Henselian ring. Our claim implies that a finite separable ring extension $k_a \subset k'$ contains a copy of k_a if and only if the tensor product $k' \otimes F_a$ contains a copy of F_a (use a lift $O_a \to A$ of the finite extension $k_a \to k'$). This gives that any scheme étale over the valuation ring of k_a has a k_a-valued point if and only if it has a F_a-valued point. This assertion combined with the fact that the valuation ring of F_a is Henselian implies that k_a is a Henselian valued field (use the criterium of [R70, Proposition 3 page 76]).

To prove our claim, note that the ring extension $O_a \subset A$ comes from a finite free ring extension $O(U) \subset A_U$ for some U in the filter of a. Clearly, A_U is an affinoid algebra and hence determines a finite flat morphism $\phi : V = \mathrm{Spm}(A_U) \to U$. The fact that $A \hat{\otimes} F_a \cong A_U \hat{\otimes} F_a$ has a nontrivial idempotent is equivalent to the fact that $\phi^{-1}(a) = b_1, \dots, b_s$ has at least two elements. Let us take disjoint wide neighbourhoods V_i of the b_i in V. There exists a smaller U' in the filter of a such that $\phi^{-1}(U')$ is contained in $\cup V_i$ (see Lemma 3.1.6 below; the reader may check that this lemma is not used before that lemma). Therefore the algebra $A_{U'} = A_U \otimes O(U')$ decomposes and hence so does A. The proof for O_a^{mod} is the same.

(3) After dividing $O(X)$ by a prime ideal we may suppose that $|\ |_a$ is a norm on $O(X)$. The field of quotients of $O(X)$ is a dense subfield of F_a. The algebra $O(X)$ is finite over some $A := k\langle T_1, \dots, T_d\rangle$ with d equal to the dimension of X. Let $K \subset F_a$ denote the completion of the field of quotients of A with respect to $|\ |_a$. The field F_a is finite over K and K has $k(T_1, \dots, T_d)$ as dense subfield with respect to $|\ |_a$.

(4) It is clear that the homomorphism $O_a^{mod} \to O_a$ is local and flat. Suppose that $\wp$ is the kernel of the seminorm $|\ |_a$ on $O(X)$. It is clear that the fraction field of $O(X)/\wp$ is dense in both F_a and F_a^{mod}. The result follows. $\square$

Remark 2.1.2 It follows from this lemma and its proof that there are equivalences between the following categories: the category of finite separable extensions of F_a, of finite separable extensions of k_a, of finite separable extensions of k_a^{mod}, of finite étale extensions of local rings $O_a \subset A$, and of finite étale extensions of local rings $O_a^{mod} \subset A$. Furthermore, any such extension comes from a finite étale (see paragraph 4) morphism $V \to U$ where U is a wide neighbourhood of a.

It is clear that the above constructions are functorial in the following sense. If $f : Y \to X$ is a morphism of affinoids over k, then we get a morphism of sites $Y_{rigid} \to X_{rigid}$ (resp. $Y_{rigid}^{rat} \to X_{rigid}^{rat}$). Indeed, if $U \subset X$ is an affinoid subdomain (resp. rational subset) then so is $f^{-1}(U) \subset Y$. Hence a functor $X_{rigid} \to Y_{rigid}, U \mapsto f^{-1}(U)$, it is easy to see that this is continuous and compatible with fibre products (i.e., intersections). The associated adjoint functors on sheaves are denoted f^*, f_* as usual.

The morphism f also induces a continuous map: $\mathcal{M}(Y) \to \mathcal{M}(X)$. The seminorm $O(Y) \to \mathbf{R}_{\geq 0}$ is mapped to the composition $O(X) \to O(Y) \to \mathbf{R}_{\geq 0}$. We remark that if f identifies Y with an affinoid subdomain of X then 1) $Y_{rigid} \cong X_{rigid}/Y$ and 2) the analytic points of Y are identified with those analytic points a of X such that Y is in the filter of a, i.e., $a \in Y$.

2.2 Sites, sheaves and analytic points for general X

To the analytic variety X we associate the site X_{rigid} by exactly the same definition as for affinoid X's. The objects are the finite unions of affinoid open subvarieties and the coverings are coverings which can be refined to finite coverings. (Since X is quasi-separated, the intersection of two affinoid open subvarieties is an object of the category X_{rigid}, so that X_{rigid} is indeed a site.) We remark that the the associated topos $X_{rigid}^{\sim}$ is again naturally isomorphic to the category of sheaves on X (as defined in [BGR, 9.2]). A morphism $f : Y \to X$ induces a morphism of topoi $f_{rigid} : Y_{rigid}^{\sim} \to X_{rigid}^{\sim}$ but not in general a morphism of sites $Y_{rigid} \to X_{rigid}$. Indeed, this morphism of sites exists if and only if f is quasi-compact.

The space X has some admissible covering $\{X_i\}$ by affinoids subsets. The analytic points of X are just the analytic points of the X_i, subject to the usual equivalence relation. (For a more precise definition see [S93, §2].) We remark that our $f : Y \to X$ induces a map on analytic points.

Finally, suppose $f : Y \to X$ is an open immersion (in the sense of [BGR, p. 354]). It is easy to prove (using the above) that: 1) f induces an injection between the sets of analytic points and 2) f induces an isomorphism $Y_{rigid}^{\sim} \to X_{rigid}^{\sim}/Y$ (where Y denotes the sheaf $V \mapsto \mathrm{Mor}_X(V, Y)$ on X_{rigid}). However, it is not true that any f satisfying 1) and 2) is an open immersion.

2.3 Overconvergent sheaves on affinoids

Let X be an affinoid variety over k. The collection of analytic points of X is still not large enough to "separate" the Abelian sheaves on X_{rigid}. We can introduce a larger collection of points as in [P82] to remedy this fact. However, this larger collection of points seems not to be of much use for questions like base change theorems et cetera. We choose to work with a restricted collection of sheaves, namely the overconvergent sheaves on X_{rigid}.

Suppose that $V \subset U$ are special subsets of X. We will say that V is inner in U (w.r.t. X), or that U is a *wide neighborhood* of V in X, if for any analytic point a of V there is an affinoid wide neighborhood U_a of a in X with $U_a \subset U$. Notation: $V \subset\subset_X U$. It is proved in [S93, §1 Proposition 23] that this agrees with the notion V is relatively compact in U over X (see [BGR, 9.6.2]) if V and U are affinoid subdomains of X: $V \subset\subset_X U \Leftrightarrow$ there is an affinoid generating system $f_1, \ldots, f_r$ of $O(U)$ over $O(X)$ such that

$$V \subset \{x \in U; |f_1(x)| < 1, \ldots, |f_r(x)| < 1\}.$$

Suppose $V \subset X$ is rational in X given by the inequalities $|g_0| \geq |g_1|, ..., |g_m|$. For $r > 1$ and $r \in \sqrt{|k^*|}$ we define the rational set $V(r)$ by the inequalities $r|g_0| \geq |g_1|, ..., |g_m|$. It is easy to see that $V \subset\subset_X V(r)$. (The notation $V(r)$ will be used even if no explicit system $g_0, \ldots, g_m$ defining V and $V(r)$ is indicated.)

LEMMA 2.3.1 *With notations as above.*

1. *The $V(r)$ form a co-final system of (special) wide neighborhoods in X of the rational set V.*

2. *If $V_1, \ldots, V_n$ are rational in X then*

$$V_1 \cap \ldots \cap V_m \subset\subset_X V_1(r) \cap \ldots \cap V_m(r)$$

($r > 1$ and $r \in \sqrt{|k^*|}$) and this forms a co-final system of wide neighborhoods of $V_1 \cap \ldots \cap V_m$. Similarly for $V_1 \cup \ldots \cup V_m \subset\subset_X V_1(r) \cup \ldots \cup V_m(r)$.

Proof. Suppose that $V \subset\subset_X U$ (with U a special subset of X). We claim the covering $X = U \cup (X \setminus V)$ is admissible. This is proved in [P92, Lemma 1.1], but let us indicate another proof: For any analytic point a of X, $a \notin V$ choose an affinoid wide neighborhood W_a of a with $W_a \cap V = \emptyset$ (just define W_a by suitable inequalities). For an analytic point $a \in V$ we choose the affinoid wide neighborhood W_a of a in X which is contained in U. Since $\mathcal{M}(X)$ is compact the covering $X = \bigcup W_a$ is admissible (see 3.1), hence so is $X = U \cup (X \setminus V)$. This proves our claim. In particular there is a special $W \subset X \setminus V$ such that $X = U \cup W$.

Next, put $W_i = \{w \in W; \ |g_i(x)| \geq |g_j(x)| \ j = 0, \ldots, n\}$ for $i = 1, \ldots, n$. Of course $W = \bigcup W_i$ since $W \cap V = \emptyset$. On W_i the function g_i is invertible hence we can put

$$\epsilon_i = ||g_0/g_i||_{W_i} \quad \text{and} \quad \epsilon = \max_i \epsilon_i.$$

By the maximum modulus principle on W_i and since $W \cap V = \emptyset$ we get $\epsilon_i < 1$ and $\epsilon < 1$. It is now clear that for any $r \in \sqrt{|k^*|}$, $\epsilon^{-1} > r > 1$ we have $V(r) \cap W = \emptyset$ and hence $V(r) \subset U$.

We prove 2) only in the case $m = 2$. Suppose that V_1 is given by the inequalities $|g_0| \geq |g_1|, \ldots, |g_n|$ and that V_2 is given by the inequalities $|f_0| \geq |f_1|, \ldots, |f_{n'}|$. The intersection $V_1(r) \cap V_2(r)$ is given by the inequalities $r^2|g_0 f_0| \geq |g_i f_j|$, $i = 0, \ldots, n$, $j = 0, \ldots, n'$. The result follows. The statement for unions is trivial from 1). □

At this point we are able to define the rigid overconvergent sheaves on our affinoid variety X. A (pre)sheaf S (on X_{rigid}) is called *(rigid) overconvergent* if for every admissible open $V \subset X$ we have

$$S(V) = \varinjlim_{V \subset\subset_X U} S(U).$$

It follows from the lemma above that if S is a sheaf then S is overconvergent if and only if $S(V) = \lim S(V(r))$ for any rational $V \subset X$. These sheaves were called the constructible sheaves in [P82]; they agree with the conservative sheaves of [S93] by [S93, §1 Lemma 25]. In [S93, §1] it is shown that these overconvergent sheaves correspond to sheaves on the topological space $\mathcal{M}(X)$.

LEMMA 2.3.2 *(Properties of overconvergent sheaves.) In this lemma all (pre)sheaves are (pre)sheaves of Abelian groups on the affinoid variety X.*

1. *The sheaf associated to a overconvergent presheaf is overconvergent.*
2. *For any overconvergent sheaf S the presheaves $U \mapsto H^i(U, S)$ are overconvergent.*
3. *The category of overconvergent sheaves is an exact subcategory of the category of all sheaves.*
4. *If $f : Y \to X$ is a morphism of affinoids then f^* and f_* preserve overconvergent sheaves. The same holds for $R^q f_*$.*
5. *If $X = \bigcup X_i$ is written as the finite union of affinoid subdomains then a sheaf S on X is overconvergent if and only if the restriction of S to any of the X_i is overconvergent.*
6. *A overconvergent sheaf S is zero if and only if all of its stalks S_a at analytic points of X are zero.*

Proof. Let S be a overconvergent presheaf. Suppose $V \subset X$ is the union of rational subsets $V_1, \ldots, V_m$ of X. Denote by $\mathcal{V} = \{V_i\}$ the covering of V and by $\mathcal{V}(r) = \{V_i(r)\}$ the covering of $V(r) := \bigcup_i V_i(r)$. It is immediate from Lemma 2.3.1 that

$$\mathcal{C}^{\cdot}(\mathcal{V}, S) = \varinjlim{}_{r>1} \mathcal{C}^{\cdot}(\mathcal{V}(r), S).$$

(These symbols denote Čech complexes.) It is therefore clear that the map

$$\varinjlim_{V \subset\subset_X U} \check{H}^p(U, S) \longrightarrow \check{H}^p(V, S)$$

is surjective.

Let us prove that it is also injective. Take a special $U \subset X$ with $V \subset\subset_X U$, an admissible covering $\mathcal{U} = \{U_i\}$ of U, a co-cycle $\xi \in \mathcal{C}^p(\mathcal{U}, S)$ whose Čech cohomology class maps to zero in $\check{H}^p(V, S)$. This means there is a covering $\mathcal{V} = \{V_j\}$ of V which refines $\mathcal{U} \cap V$, i.e., there is a function α such that $V_j \subset U_{\alpha(j)} \cap V$, and a chain $\eta \in \mathcal{C}^{p-1}(\mathcal{V}, S)$ with $\alpha(\xi) - \mathrm{d}\eta = 0 \in \mathcal{C}^p(\mathcal{V}, S)$. Here $\alpha(\xi)$ is the image of ξ under the map $\mathcal{C}^p(\mathcal{U}, S) \to \mathcal{C}^p(\mathcal{V}, S)$ determined by α. By refining $\mathcal{U}$ and $\mathcal{V}$ we may assume that $\mathcal{U}$ and $\mathcal{V}$ are finite and that all U_i and V_j are rational subdomains of X.

By the above, the co-cycle ξ lifts to a co-cycle $\xi' \in \mathcal{C}^p(\mathcal{U}(r), S)$ for some $r > 1$. Lemma 2.3.1 implies that there exists an $r^* > 1$ such that $V_j(r^*) \subset U_{\alpha(j)}(r) \forall j$. For an even smaller r^*, we may also assume η lifts to a chain $\eta' \in \mathcal{C}^{p-1}(\mathcal{V}(r^*), S)$. The co-cycle $\alpha(\xi') - \mathrm{d}\eta' \in \mathcal{C}^p(\mathcal{V}(r^*), S)$ maps to zero as a chain in $\mathcal{C}^p(\mathcal{V}, S)$, thus it is already zero in some $\mathcal{C}^p(\mathcal{V}(r^{**}), S)$, $r^* > r^{**} > 1$. We conclude that the cohomology class of ξ in $\check{H}^p(V(r^{**}), S)$ is zero, which was what we wanted to show.

The isomorphism of Čech cohomologies above proves that the presheaf $\check{\mathcal{H}}^0(S)$ is overconvergent if S is overconvergent. Hence also the sheaf associated to S is overconvergent. It proves (2) since Čech cohomology agrees with usual cohomology for any special $U \subset X$. (See [P82, 1.4.4] or our Proposition 2.5.4.)

The third statement of our lemma means that the kernels and co-kernels of overconvergent sheaves are overconvergent and that if a short exact sequence of sheaves $0 \to S_1 \to S_2 \to S_3 \to 0$ is given, S_1 and S_3 are overconvergent then so is S_2. These statements follow easily from (1) and (2).

(4) If $V \subset X$ is a rational subset, then $f^{-1}(V)$ is a rational subdomain of Y and we have: $f^{-1}\big(V(r)\big) = \big(f^{-1}(V)\big)(r)$. Thus it is clear from Lemma 2.3.1 that for special $V \subset\subset_X U$ in X we have $f^{-1}(V) \subset\subset_Y f^{-1}(U)$ and that these $f^{-1}(U)$ form a co-final system of wide neighborhoods of $f^{-1}(V)$.

Take an overconvergent sheaf S on Y. The sheaf $R^q f_* S$ is the sheaf associated to the presheaf $U \mapsto H^q(f^{-1}(U), S)$. It is immediate from the remarks above and (2) that this presheaf is overconvergent.

If S is a sheaf on X then $f^* S$ is the sheaf associated to the presheaf P defined as follows on $V \in X_{rigid}$:

$$P(V) = \varinjlim_{U \in X_{rigid}, f^{-1}(U) \supset V} S(U)$$

Suppose S is overconvergent. If $t \in P(V)$, i.e., t comes from $s \in S(U)$ for some $U \subset X$ as in the limit, then s comes from $s' \in S(U')$ for some $U' \in X_{rigid}$ with $U \subset\subset_X U'$. By the above we see that $V \subset\subset_Y f^{-1}(U')$. We conclude that the map

$$\varinjlim_{V \subset\subset_Y V'} P(V') \to P(V)$$

is surjective. Let us prove that it is injective: Suppose $t' \in P(V')$ comes from some $s' \in S(U')$ with $f^{-1}(U') \supset V'$ and maps to zero in some $S(U)$ with $U \subset U'$ and $f^{-1}(U) \supset V$. There exists a wide neighborhood U'' of U' and $s'' \in S(U'')$ mapping to s'. Since S is overconvergent there is a special U''' with $U''' \subset U''$, $U''' \supset\supset_X U$ such that s'' maps to zero in U'''. It is clear that $V'' := V' \cap f^{-1}(U''')$ is a wide neighborhood of V in Y such that t' maps to zero in $P(V'')$. We have proved that P, hence $f^* S$, is overconvergent.

(5) This follows from (3) and (4) since any sheaf S on X fits into an exact sequence

$$0 \longrightarrow S \longrightarrow \bigoplus_i S|_{X_i} \longrightarrow \bigoplus_{i,j} S|_{X_i \cap X_j}.$$

Here $S|_{X_i} := j_* j^* S$ where $j : X_i \to X$ is the inclusion.

(6) Take a section $s \in \Gamma(X, S)$. By assumption any analytic point a in X has an affinoid wide neighborhood $V_a \subset X$ such that $s|_{V_a} = 0$. By compactness of $\mathcal{M}(X)$ we get that the covering $X = \bigcup V_a$ is admissible, hence $s = 0$. The same proof gives that $\Gamma(V, S) = 0$ for arbitrary special $V \subset X$. □

2.4 Overconvergent sheaves on general X

Let X be an arbitrary analytic variety over k. We will say that a sheaf S on X is *overconvergent* if for any affinoid open subvariety $V \subset X$, the restriction $S|_V$ is overconvergent on V. Suppose $X = \bigcup X_i$ is an admissible affinoid covering. It follows from Lemma 2.3.2 that S is overconvergent if and only if $S|_{X_i}$ is overconvergent for all i.

Suppose $f : Y \to X$ is a morphism of rigid varieties. It is clear from Lemma 2.3.2 that f^* preserves overconvergent sheaves. This is not true in general for f_* or $R^q f_*$. But it is true if f is quasi-compact.

Proposition 2.4.1 *If $f : Y \to X$ is a quasi-compact morphism then f_* and $R^q f_*$ preserve overconvergent presheaves.*

Proof. Take an overconvergent sheaf S on Y. The question is local on X, hence we may assume X affinoid. Thus Y is quasi-compact and hence by Lemma 2.5.3 we can find a finite admissible affinoid covering $Y = \bigcup Y_i$ such that all intersections $Y_{i_0 \ldots i_q} := Y_{i_0} \cap \ldots \cap Y_{i_q}$ are affinoid. At this point we use the spectral sequence (deduced from the Cartan-Leray spectral sequence [SGA 4, V 3.3]) $\{E_n^{pq}\}$ abutting to $R^n f_* S$ and with E_2-term:

$$E_2^{pq} = \bigoplus_{i_0 \ldots i_q} R^p \big(f|_{Y_{i_0 \ldots i_q}}\big)_* S|_{Y_{i_0 \ldots i_q}}$$

By Lemma 2.3.2 all its terms are overconvergent sheaves. Hence by the same lemma we see that $R^n f_* S$ is overconvergent too. □

2.5 Cohomology and Čech cohomology

In this subsection we prove that cohomology agrees with Čech cohomology on quasi-compact varieties. Further we prove that the cohomological dimension of such an analytic variety is at most its dimension.

Lemma 2.5.1 *Let X be an affinoid variety, $V \subset X$ special and a an analytic point of X. There exists a wide neighborhood $W = W_a$ of a such that $W \cap V$ is a finite union of Weierstrass domains, each defined by invertible functions.*

Proof. Since V is a finite union of rational subsets of X we may assume that V is rational itself. Say it is defined by the inequalities $|g_0| \geq |g_1|, \ldots, |g_n|$, where the g_i generate the unit ideal of $O(X)$. If $a \notin V$, then we can find a wide neighborhood W of a disjoint with V. If $a \in V$ then $|g_0|_a \geq |g_i|_a$ and since the g_i generate the unit ideal we get $|g_0|_a > 0$. Thus we may replace X by a wide neighborhood of a, so that g_0 becomes invertible. In this situation V is defined by $1 \geq |f_i|$ with $f_i = g_i/g_0$, i.e., V is a Weierstrass domain in X. For those i such that $\epsilon_i := |f_i|_a < 1$, we may replace

X by the wide neighborhood of a defined by $|f_i| \leq 1/2(1+\epsilon_1)$ and drop f_i. At this point $V \subset X$ is defined as $1 \geq |f_i|$ with $|f_i|_a = 1$ for all i. Hence the subset $|f_i| \geq |\pi|$, $\pi \in k, 0 < |\pi| < 1$ defines a wide neighborhood of a such that f_i is invertible on it. □

Lemma 2.5.2 *Suppose X is affinoid, $V \subset X$ special. There exists a finite covering $X = \bigcup X_i$ by affinoids of X such that $X_i \cap V$ is affinoid for all i.*

Proof. By compactness of $\mathcal{M}(X)$ and the lemma above we may assume $V \subset X$ is a finite union of Weierstrass domains, each given by invertible functions. Say $V = \bigcup_{i=1}^n V_i$ and V_i is defined by $1 \geq |f_1^i|, \ldots, |f_{n_i}^i|$ and each f_j^i invertible.

Consider combinatorial data of the form $A = (i, (j_1, \ldots, \widehat{j_i}, \ldots, j_n))$ where $i \in \{1, \ldots, n\}$ and $j_l \in \{1, \ldots, n_l\}$ for each $l \neq i, l \in \{1, \ldots, n\}$. We put

$$V_A = \left\{x \in X; |f_j^i(x)| \leq |f_{j_l}^l(x)|,\, l = 1, \ldots, \hat{i}, \ldots, n,\, j = 1, \ldots, n_i\right\}$$

Remark that $X = \bigcup_A V_A$ since for any $x \in X$ there is some $i \in \{1, \ldots, n\}$ such that $\max_j |f_j^i(x)| \leq \max_j |f_j^l(x)|$ for all $l \neq i$. On the other hand, if $A = (i, (j_1, \ldots, \widehat{j_i}, \ldots, j_n))$ as above then

$$V_A \cap V \subset V_i,$$

and hence $V_A \cap V = V_A \cap V_i$ is affinoid. This is immediate from the definitions. □

We remark that in proving the lemmata above we proved something slightly stronger: Suppose we had started with an admissible affinoid covering $V = \bigcup V_i$. This we can refine to a finite covering $V = \bigcup V_i$ with $V_i \subset X$ rational. The proof of Lemma 2.5.1 shows that we can cover X by finitely many affinoids X_j such that each $X_j \cap V_i$ is a Weierstrass domain in X_j defined by invertible functions. The proof of Lemma 2.5.2 shows that we can cover each X_j by finitely many $X_{j,A}$'s such that $X_{j,A} \cap (V \cap X_j) = X_{j,A} \cap V$ is contained in some V_i. Thus we have proved the first statement of the following lemma in the case that X is affinoid.

Lemma 2.5.3 *Let X be a quasi-compact variety over k.*

1. *Given an admissible covering $\mathcal{V} : V = \bigcup V_i$ of the special subset V of X, there exists a finite affinoid covering $\mathcal{U} : X = \bigcup X_j$ such that the covering $\mathcal{U} \cap V$ refines $\mathcal{V}$. In addition we may assume $X_j \cap V_i$ affinoid for all j.*

2. *There exists a finite affinoid covering $X = \bigcup X_j$ such that $X_i \cap X_j$ is affinoid for all i, j.*

Proof. (1) This assertion follows immediately from the case X affinoid (proved above) by writing X as the finite admissible union of affinoids (use that X is quasi-separated by our conventions).

(2) Take first an arbitrary finite affinoid covering $X = \bigcup X_i$. By (1) we can find finite affinoid coverings $\mathcal{U}_{ij} : X_i = \bigcup_k X_{ijk}$ such that $X_{ijk} \cap (X_i \cap X_j)$ is affinoid for all k. Next we take a finite affinoid covering $\mathcal{U}_i : X_i = \bigcup_l X_{il}$ refining $\mathcal{U}_{ij}$ for all j. It is clear that $X_{il} \cap X_{jm} = X_{il} \cap (X_i \cap X_j) \cap X_{jm}$ is affinoid (all intersections are taken in X). Thus the covering $X = \bigcup X_{il}$ works. □

Proposition 2.5.4 *Suppose X is a quasi-compact (and quasi-separated) variety. Čech cohomology agrees with cohomology on X.*

Proof. The Leray spectral sequence relating Čech cohomology with cohomology [SGA 4, V 3.4] shows that it suffices to prove: $\check{H}^p(X, S) = 0$ if S is a presheaf whose associated sheaf is zero. Suppose $\mathcal{V}$ is some finite admissible covering of X and $\xi = \prod \xi_{i_o \ldots i_p} \in \mathcal{C}^p(\mathcal{V}, S) = \prod_{i_o \ldots i_p} S(V_{i_o \ldots i_p})$. We can find a covering $\mathcal{V}_{i_o \ldots i_p}$ of $V_{i_o \ldots i_p}$ such that $\xi_{i_o \ldots i_p}$ restricts to zero on each member of $\mathcal{V}_{i_o \ldots i_p}$. By Lemma 2.5.3 we can find a covering $\mathcal{U}_i : V_i = \bigcup U_{ij}$ of V_i such that $\mathcal{U}_i \cap V_{i_o \ldots i_p}$ (some $i_l = i$) refines $\mathcal{V}_{i_o \ldots i_p}$ for all choices of the i_l. Put $\mathcal{U} = \bigcup \mathcal{U}_i$, it is an admissible covering of X and the map

$$\alpha : \mathcal{C}^*(\mathcal{V}, S) \longrightarrow \mathcal{C}^*(\mathcal{U}, S)$$

is defined using $U_{ij} \subset V_i$. It is clear that the chain ξ maps to zero under α. □

Remark 2.5.5 By Lemma 2.5.3 this is a special case of [P82, 1.4.4]. The argument in the proof of [P82, 1.4.5] together with Lemma 2.5.3 shows that Čech cohomology agrees with cohomology on any (quasi-separated, see conventions) X which is of countable type (see Definition 2.5.6 below).

We introduce some convenient topological notions for the Grothendieck topology on our analytic varieties X.

Definition 2.5.6 *Let X be an analytic variety over k.*

1. *We say that X is of countable type if there exists a countable admissible affinoid covering of X.*

2. *Suppose that $X = \bigcup X_i$ is an admissible affinoid covering of X. We say that the covering is locally finite if each X_i meets finitely many X_j.*

3. *The variety X will be called paracompact if there exists an admissible locally finite affinoid covering.*

Lemma 2.5.7 *A paracompact space X is the admissible disjoint union of paracompact varieties of countable type. A connected paracompact variety X can be written as the admissible union $X = \bigcup_{n \in \mathbb{N}} X_n$, with X_n quasi-compact and $X_i \cap X_j = \emptyset$ when $|i - j| \geq 2$.*

Proof. Since any rigid analytic space is the admissible disjoint union of its connected components, it suffices to prove the second statement. Therefore we assume that X is connected and has a locally finite affinoid admissible covering $X = \bigcup X_\alpha$. Let us choose a fixed index α_0. For any α we define the distance $d(\alpha)$ of α to α_0 to be the minimal length d of a sequence of indices $\alpha_0, \alpha_1, \ldots, \alpha_d = \alpha$ such that $X_{\alpha_i} \cap X_{\alpha_{i+1}} \neq \emptyset$ for all $i = 0, \ldots, d-1$. Since X is connected all distances are finite. We put $X_n = \bigcup_{d(\alpha)=n} X_\alpha$. Since the covering was locally finite the spaces X_n are quasi-compact. The last condition of the lemma follows immediately from our definition of distance. □

In the proof of the next proposition we need the relation of rigid analytic geometry with formal geometry (see [R70] and [BL]). We recall that if $\mathfrak{X}$ is a formal scheme of finite type and flat over $\mathrm{Spf}(k^\circ)$ then there is canonically associated a quasi-compact rigid analytic variety $X = \mathfrak{X}^{rig}$. If $\mathfrak{U} \subset \mathfrak{X}$ is a formal open subscheme then $\mathfrak{U}^{rig} \subset \mathfrak{X}^{rig}$ is an open subvariety. If $\mathfrak{X} = \bigcup \mathfrak{U}_i$ then $\mathfrak{X}^{rig} = \bigcup \mathfrak{U}_i^{rig}$ is an admissible covering (see [BL, §4]). Thus we get a morphism of sites $X_{rigid} = \mathfrak{X}^{rig}_{rigid} \to \mathfrak{X}_{Zar}$.

It is also possible to perform the construction $\mathfrak{X} \mapsto \mathfrak{X}^{rig}$ for formal schemes $\mathfrak{X}$ which are only locally of finite type over $\mathrm{Spf}(k^\circ)$. It is not true that any rigid variety X comes from such a formal scheme. A counterexample can be constructed by gluing a countable number of closed discs to a fixed closed disc along mutually disjoint closed sub-discs. (This is also an example of a variety of countable type which is not paracompact.) It can be proved using the lemma above and [BL] that any paracompact X comes from a (paracompact) formal scheme $\mathfrak{X}$.

Proposition 2.5.8 *(See [P82, 1.4.13]). If X is a quasi-compact rigid analytic variety of dimension d then $H^p(X, S) = 0$ for all $p > d$ and all sheaves S on X.*

Proof. Let us choose a formal scheme $\mathfrak{X}$ with $\mathfrak{X}_{rig} \cong X$ (see [R70] or [BL, Theorem 4.1]). Let us denote by $\{\mathfrak{X}_\alpha\}$ the directed system of admissible blowing ups of $\mathfrak{X}$. These all satisfy $\mathfrak{X}_\alpha^{rig} \cong X$. Hence we get the morphism of sites $\pi_\alpha : X_{rigid} \to \mathfrak{X}_{\alpha,Zar}$. Let us write $S_\alpha := \pi_{\alpha,*}S$. There is a map $H^p_\alpha := H^p(\mathfrak{X}_{\alpha,Zar}, S_\alpha) \to H^p(X, S)$ deduced from the map $\pi^*\pi_{\alpha,*}S \to S$. It is proved in [BL, 4.4] that any finite covering of X comes from a covering of some $\mathfrak{X}_\alpha$. Therefore, by our result that Čech cohomology agrees with cohomology on X, we see that any cohomology class in $H^p(X, S)$ comes from some H^p_α. At this point we just remark that the underlying Zariski topological space associated to $\mathfrak{X}_\alpha$ is the underlying topological space of a scheme of finite type over the field $\bar{k}$ of dimension at most n. The result follows. □

Remark 2.5.9 If we allow in X_{rigid} only finite coverings then it is true that $\lim \mathfrak{X}_{\alpha,Zar} \cong X_{rigid}$ as sites (see letter of Deligne to Raynaud of 23 august 1992). In this way it becomes clear that in fact $\lim H^p_\alpha = H^p(X, S)$. This follows from the following general fact: Suppose the site $\mathcal{S}$ is the direct limit of a directed system of sites $\mathcal{S}_\alpha$. Then for any sheaf $\mathcal{F}$ on $\mathcal{S}$ there is a canonical isomorphism

$$\varinjlim_\alpha H^q(\mathcal{S}_\alpha, \mathcal{F}|_{\mathcal{S}_\alpha}) \cong H^q(\mathcal{S}, \mathcal{F}).$$

This isomorphism is in fact easy to prove by induction on q, using the Cartan-Leray spectral sequence and the fact that any cohomology class can be killed by some covering.

Corollary 2.5.10 *If X is paracompact and of dimension $\le d$ then cohomology of sheaves on X is zero in degrees $\ge d+1$.*

Proof. It suffices to do the case where X is connected. Choose a covering $X = \bigcup X_n$ as in Lemma 2.5.7. Put $V_1 = \bigcup_{n \text{ odd}} X_n$ and $V_2 = \bigcup_{n \text{ even}} X_n$. The spaces V_1, V_2 and $V_1 \cap V_2$ are admissible disjoint unions of quasi-compact varieties. Note that for any sheaf S on X the maps $H^d(X_n, S) \oplus H^d(X_{n+1}, S) \to H^d(X_n \cap X_{n+1}, S)$ is surjective, otherwise the sheaf S on $X_n \cup X_{n+1}$ would have a nontrivial $d+1^{th}$-cohomology group, a contradiction with the proposition. With these remarks the result of the corollary follows from a consideration of the Cartan-Leray spectral sequence associated to the covering $X = V_1 \cup V_2$. $\square$

Remark 2.5.11 Any separated variety of dimension 1 is paracompact. See [LP]. Similarly, the analytic space associated to a scheme of finite type over $\mathrm{Spec}(k)$ is paracompact.

2.6 General morphisms

Consider an extension of complete valued fields $k \subset K$. In [BGR, 9.3.6] there is constructed a base change functor $X \mapsto X\hat{\otimes}K$ of analytic varieties over k to analytic varieties over K. If X is affinoid then $X\hat{\otimes}K$ is affinoid with algebra $O(X)\hat{\otimes}_k K$. In general, if $X = \bigcup X_i$ is an admissible affinoid covering then $X\hat{\otimes}K$ is defined as the gluing of the $X_i\hat{\otimes}K$. If V is an affinoid open subvariety of X then so is $V\hat{\otimes}K \subset X\hat{\otimes}K$. In this way (use [BGR, 9.3.6/1& 2]) we see that there is a morphism of sites

$$\varphi = \varphi_{K/k} : (X\hat{\otimes}K)_{rigid} \to X_{rigid}.$$

Lemma 2.6.1 *The functors φ^*, φ_* and $R^q\varphi_*$ preserve overconvergent sheaves.*

Proof. There is a trivial reduction to the case that X is affinoid. Let V be a rational subdomain of X. It is clear that $V(r)\hat{\otimes}K = \big(V\hat{\otimes}K\big)(r)$ for $r > 1, r \in \sqrt{|k^*|}$ (see [BGR, 9.3.6/1]). These form a co-final system of wide neighborhoods of $V\hat{\otimes}K$ since $\sqrt{|k^*|}$ is dense in $\mathbf{R}_{\geq 0}$. Thus it is clear from Lemma 2.3.1 that for special $V \subset\subset_X U$ in X we have $V\hat{\otimes}K \subset\subset_{X\hat{\otimes}K} U\hat{\otimes}K$ and that these $U\hat{\otimes}K$ form a co-final system of wide neighborhoods of $V\hat{\otimes}K$. The rest of the proof is exactly the same as the proof of Lemma 2.3.2 part 4. $\square$

Let $k \subset K$ denote an extension of complete valued fields. Let X (resp. Y) denote an arbitrary analytic variety over the field k (resp. K). The most convenient way to define a *general morphism* $f : Y \to X$ is to say that f is a morphism of the K-analytic spaces $Y \to X\hat{\otimes}K$. If both X and Y are affinoid then this is simply a continuous k-algebra homomorphism $O(X) \to O(Y)$, since any such factors as $O(X) \to O(X)\hat{\otimes}_k K \to O(Y)$. By the above, a general morphism $f : Y \to X$ gives rise to a morphism of topoi $f_{rigid} : Y^{\sim}_{rigid} \to X^{\sim}_{rigid}$. The pullback functor, written f^*, preserves overconvergent sheaves. We say that the morphism f is *quasi-compact* if $Y \to X\hat{\otimes}K$ is quasi-compact. In this case f induces a morphism of sites $Y_{rigid} \to X_{rigid}$ and $R^q f_*$ preserves overconvergent sheaves for all q. (Use the lemma above and Proposition 2.4.1.)

If, in addition, we are given a morphism $Z \to X$ of analytic varieties over k, then we can form the fibre product:

$$Y \times_X Z := Y \times_{X\hat{\otimes}K} X\hat{\otimes}K$$

It is an analytic variety over K which satisfies a certain universal property regarding general morphisms; we leave it to the reader to describe this property explicitly.

2.7 Base change

The aim of the base change theorem is to compare $H^q(Y_a, S|_{Y_a})$ with $\left(R^q f_* S\right)_a$ for sheaves S on Y. Here Y_a is the fibre of a morphism f over the analytic point a. Let us first define this fibre.

Consider a morphism $f : Y \to X$ of analytic varieties over k and let an analytic point a of X be given. The *fibre* Y_a of f over a is defined as the fibre product of the general morphism $\mathrm{Spm}(F_a) \to X$ with f. It can also be defined as the fibre of $f\hat{\otimes}F_a : Y\hat{\otimes}F_a \to X\hat{\otimes}F_a$ over the usual point $a \in X\hat{\otimes}F_a$. There results a general morphism $\alpha : Y_a \to Y$. We remark that α is quasi-compact; the morphism of sites $(Y_a)_{rigid} \to Y_{rigid}$ comes from the functor $V \mapsto V_a$ on special subsets of Y. For a sheaf S on Y we write $S|Y_a$ instead of $\alpha^*(S)$. Finally, we remark that if both X and Y are affinoid then Y_a is affinoid with algebra $O(Y)\hat{\otimes}_{O(X)}F_a$.

Lemma 2.7.1 *(Key lemma for the rigid case.) Let a morphism $f : Y \to X$ of affinoid spaces over k be given together with an analytic point a of X. Write $\alpha : Y_a \to X$ for the resulting general morphism.*

1. *For every admissible open $V \subset Y_a$ (i.e., $V \in (Y_a)_{rigid}$) there is an admissible open $W \subset Y$ such that $V = W_a$.*
2. *Suppose W, Z are admissible open in Y and $W_a \subset Z_a$. There is a U in the filter of a such that $W \cap f^{-1}(U) \subset Z$.*

Proof. (1) We may assume that V is a rational subset of Y_a. Thus V is given by inequalities $|g_1| \geq |g_1|, ..., |g_m|$ with elements $g_1, ..., g_m \in O(Y_a) = O(Y)\hat{\otimes}_{O(X)}F_a$ generating the unit ideal. Say that $f_1g_1 + \ldots + f_mg_m = 1$. We may suppose that the g_i come from elements $g_i \in O(Y) \otimes_{O(X)} k_a$. So there is some U in the filter of a and elements $G_i \in O(Y)\hat{\otimes}_{O(X)}O(U) = O(f^{-1}(U))$ mapping to the g_i. If we take $F_i \in O(Y)\hat{\otimes}_{O(X)}O(U) = O(f^{-1}(U))$ mapping to elements close to the f_i then we see that $F_1G_1 + \ldots + F_mG_m = 1 + \delta$ where δ maps to an element of $O(Y_a) = O(Y)\hat{\otimes}_{O(X)}F_a$ with small norm, say with spectral norm < 1. By Lemma 2.7.2 this implies that δ gets spectral norm < 1 in $O(Y)\hat{\otimes}_{O(X)}O(U) = O(f^{-1}(U))$ for some smaller U in the filter of a. Hence we see that $G_1, \ldots, G_m$ generate the unit ideal in $O(f^{-1}(U))$. Thus $W \subset f^{-1}U$ given by the inequalities $|G_1| \geq |G_1|, ..., |G_m|$ works.

(2) We may assume that W is a rational subdomain of Y. Next we write Z as a finite union $Z = \bigcup Z_i$ of rational subdomains Z_i of Y. The finite covering $W_a = \bigcup_i W_a \cap (Z_i)_a$ can be refined by a rational covering $W_a = \bigcup_j V_j$ given by a number of elements $g_1, \ldots, g_m$ in $O(W_a)$ generating the unit ideal. Arguing as above, we may suppose that the g_i come from $G_i \in O(W)$ generating the unit ideal, after replacing X by some U in the filter of a. The rational subsets W_j of W defined by $|G_j| \geq |G_1|, ..., |G_m|$ cover W and each $(W_j)_a$ is contained in some $(Z_i)_a$. If we solve the problem for all the pairs (W_j, Z_i) with $(W_j)_a \subset (Z_i)_a$ then we solve the problem for (W, Z). Thus we have reduced to the case that both W and Z are rational subdomains of Y.

At this point we replace Z by $Z \cap W$, then we are in the situation that $Z \subset W$ is a rational subdomain, $Z_a = W_a$ and we want to show that there is some U such that $W \cap f^{-1}(U) = Z \cap f^{-1}(U) \subset Z$. Suppose that Z is given by inequalities $|h_0| \geq |h_1|, ..., |h_n|$ where $h_0, ..., h_n$ generate the unit ideal in $O(W)$. In particular, h_0

is an invertible function on Z, hence on $Z_a = W_a$. Arguing as in (1), we may shrink X and assume that h_0 is invertible on W. Dividing by h_0 we see that we may suppose that Z is given by the inequalities $|h_1| \leq 1, \ldots, |h_m| \leq 1$. The h_i have norms ≤ 1 on W_a. Hence, by Lemma 2.7.2, we can find a U in the filter of a such that the h_i have norm ≤ 1 on $W \cap f^{-1}(U)$, i.e., such that $W \cap f^{-1}(U) = Z \cap f^{-1}(U)$. □

Lemma 2.7.2 *Let $f : Y \to X$ be a morphism of afinoid spaces over k, let a be an analytic point of X. Let $g \in O(Y)$ whose image $\alpha(g) \in O(Y_a)$ has spectral norm ≤ 1 (resp. < 1). There is a U in the filter of a such that the spectral norm of g on $f^{-1}(U)$ is ≤ 1 (resp. < 1).*

Proof. Let us write $O(Y) = O(X)\langle T_1, ..., T_n\rangle/(G_1, ..., G_m)$. With obvious notations we have $O(Y_a) = F_a\langle T_1, ..., T_n\rangle/(G_1(a), ..., G_m(a))$. If the spectral norm of $\alpha(g)$ is ≤ 1 it follows that $\alpha(g)$ is integral over the ring $F_a\langle T_1, ...T_n\rangle^o$. Let such an equation be

$$\alpha(g)^e + c_{e-1}\alpha(g)^{e-1} + ... + c_0 = 0$$

Write $c_i = \sum c_{i,\beta}T^\beta$ with all $c_{i,\beta} \in F_a$ satisfying $|c_{i,\beta}|_a \leq 1$.

Choose some $\pi \in k$ with $0 < |\pi| < 1$. For the $c_{i,\beta}$ with $|c_{i,\beta}|_a \geq |\pi|$ (there are only finitely many of these!) we take a suitable U in the filter of a and elements $C_{i,\beta} \in O(U)$ with images $\alpha(C_{i,\beta}) \in F_a$ such that $|\alpha(C_{i,\beta}) - c_{i,\beta}|_a < |\pi|$. (This is possible, the image of O_a is dense in F_a.) It follows that $|\alpha(C_{i,\beta})|_a \leq 1$. Thus the inequalities $|C_{i,\beta}| \leq 1$ define a smaller U in the filter of a where the elements $C_{i,\beta}$ have spectral norm ≤ 1. For convenience we replace X by U and Y by $f^{-1}U$. The $C_{i,\beta} \in O(X)$ are elements with spectral norm ≤ 1. We consider the expression

$$R := g^e + \gamma(\sum C_{e-1,\beta}T^\beta)g^{e-1} + ... + \gamma(\sum C_{0,\beta}T^\beta)$$

where γ denotes the map $O(X)\langle T_1, ..., T_n\rangle \to O(Y)$. This element $R \in O(Y)$ has an image $\alpha(R) \in O(Y_a)$ with spectral norm $< |\pi|$. If we can find a U in the filter of a such that the spectral norm of R on $f^{-1}U$ is < 1 then we replace again X by U and Y by $f^{-1}U$. After this is done the spectral norm of R on Y is < 1 and the spectral norms of the $\gamma(\sum C_{i,\beta}T^\beta)$ are ≤ 1. It follows at once that the spectral norm of g on Y is ≤ 1.

In this way we have reduced the case ≤ 1 of the lemma to the case < 1. Let us therefore assume that the spectral norm of $\alpha(g)$ is < 1. For some $N \geq 1$ the element $\alpha(g^N) \in O(Y_a)$ has a pre-image $g_1 \in F_a\langle T_1, ..., T_n\rangle$ with norm < 1. Take also a $g_2 \in O(X)\langle T_1, ..., T_n\rangle$ with image $g^N \in O(Y)$. Then $\alpha(g_2) - g_1 \in F_a\langle T_1, ..., T_n\rangle$ lies in the ideal generated by the $\{G_1(a), ..., G_m(a)\}$ and we can write

$$\alpha(g_2) - g_1 = \sum_i G_i(a)(\sum_\beta a_{i,\beta}T^\beta)$$

where the coefficients $a_{i,\beta} \in F_a$ have limit 0. For the $a_{i,\beta}$ with $|a_{i,\beta}| \geq |\pi|$ we choose a U in the filter of a and elements $A_{i,\beta} \in O(U)$ such that the difference of the image of $A_{i,\beta}$ and $a_{i,\beta}$ in F_a has absolute value $< |\pi|$. We may suppose again that $U = X$. We suppose that π is chosen such that all coefficients of $\pi G_i(a)$ have norm < 1 (in F_a). After changing g_2 into $g_2 - \sum_i G_i(\sum_\beta A_{i,\beta}T^\beta) \in O(X)\langle T_1, ..., T_n\rangle$ we have the situation that $\alpha(g_2) - g_1 \in F_a\langle T_1, ..., T_n\rangle$ and $\alpha(g_2) \in F_a\langle T_1, ..., T_n\rangle$ are power series with coefficients having absolute values < 1. For the finitely many coefficients in $O(X)$ of g_2 with absolute value ≥ 1 we can find a U in the filter of a such that their spectral norms on $f^{-1}U$ are < 1. After shrinking X to U and Y to $f^{-1}U$ we arrive at the situation where all the coefficients of g_2 are < 1. Hence the spectral norm of g^N on Y is < 1 and so the spectral norm of g on Y is < 1. □

Lemma 2.7.3 *In the situation of the Key Lemma.*

1. *The functor $S \mapsto S|_{Y_a}$ preserves flasque sheaves.*
2. *For any sheaf S on Y the sheaf $\alpha^*_{rigid}(S) = S|_{Y_a}$ can be described as follows:*

$$\Gamma(W_a, S|_{Y_a}) = \varinjlim_{a \in U \subset X} S\big(W \cap f^{-1}(U)\big).$$

Here W is any special subset of Y.

Proof. (1) This is clear from the first assertion of our Lemma 2.7.1. (2) From Lemma 2.7.1 it follows that for any S we have

$$\Gamma(W_a, S|_{Y_a}) = \varinjlim_{Z \subset Y, Z_a = W_a} S\big(Z\big).$$

The limit is over all admissible open $Z \subset Y$ such that $Z_a = W_a$. From Lemma 2.7.1 part 2 it follows that the $Z = W \cap f^{-1}(U)$ are co-final in this system. □

Finally, we come to the base change theorem. To give a natural statement recall that a δ-functor between to Abelian categories $\mathcal{A}$ and $\mathcal{B}$ is a sequence of functors $\{T_n\}_{n \geq 0}$ (equipped with certain boundary operators) such that any short exact sequence in $\mathcal{A}$ is transformed into a long exact sequence in $\mathcal{B}$. (For a more precise definition see for example [H77]).

Theorem 2.7.4 *(Base change for rigid spaces.) Let $f : Y \to X$ be a quasi-compact morphism of rigid analytic varieties over k. Take any analytic point a of X and denote by Y_a the fibre of f over a. The functors $S \mapsto H^n(Y_a, S|_{Y_a})$ (resp. $S \mapsto \big(R^n f_* S\big)_a$) on the category of Abelian sheaves on Y_{rigid} form a δ-functor. These δ-functors are isomorphic: $\big(R^n f_* S\big)_a \cong H^n(Y_a, S|_{Y_a})$ for any Abelian sheaf S on Y.*

Proof. The functor $S \mapsto S|_{Y_a}$ is exact and so is the functor $\mathcal{F} \mapsto \mathcal{F}_a$ on sheaves on X. From this follows immediately that the functors under consideration form δ-functors.

Let us define the canonical morphisms:

$$(*) \qquad \big(R^n f_* S\big)_a \longrightarrow H^n(Y_a, S|_{Y_a})$$

Since $S|_{Y_a} = \alpha^*(S)$ there are canonical homomorphisms $H^n(X, S) \to H^n(Y_a, S|_{Y_a})$ and these form a transformation of δ-functors. For any open subvariety $U \subset X$, with $a \in U$, we have $f^{-1}(U)_a = Y_a$ and $\big(S|_{f^{-1}(U)}\big)|_{Y_a} = S|_{Y_a}$. Hence the same argument gives

$$H^n(f^{-1}(U), S) \longrightarrow H^n(Y_a, S|_{Y_a}).$$

Since $\big(R^n f_* S\big)_a = \lim H^n(f^{-1}(U), S)$ (the limit is taken over U as above) we get the desired map of δ-functors.

To prove that (*) is an isomorphism we may assume that X is affinoid. Let us first do the case that Y is affinoid. The result for $n = 0$ is Lemma 2.7.3 with $W = Y$. For a flasque sheaf on Y both sides of (*) are zero for $n \geq 1$ (use 2.7.3), hence the standard argument gives the result for general n. (Inject S into a flasque sheaf and argue by induction on n.)

There are two ways to get the result for general quasi-compact Y.
(1) Choose a finite affinoid covering $Y = \bigcup Y_i$ such that all $Y_{i_0 \ldots i_q} = Y_{i_0} \cap \ldots \cap Y_{i_q}$ are affinoid (see 2.5.3). The maps (*) for Y and all $Y_{i_0 \ldots i_q}$ induce a morphism of spectral sequences $\{{}_1E_n^{pq}\} \to \{{}_2E_n^{pq}\}$ abutting to the maps $\left(R^{p+q} f_* S\right)_a \longrightarrow H^{p+q}(Y_a, S|_{Y_a})$ and with as E_2-terms the maps (*):

$$\bigoplus_{i_0 \ldots i_q} \left(R^p (f|_{Y_{i_0 \ldots i_q}})_* S\right)_a \longrightarrow \bigoplus_{i_0 \ldots i_q} H^p\big((Y_{i_0 \ldots i_q})_a, S|_{Y_{i_0 \ldots i_q}}\big)$$

These maps are isomorphisms by the above hence we get the result.
(2) Here we just remark that the Key Lemma holds for $f : Y \to X$ with X affinoid and Y quasi-compact. This follows immediately from the Key Lemma as it stands now. The base change theorem now follows from the same argument as for the case Y affinoid. □

Remark 2.7.5 (1) The result is of course most useful for overconvergent sheaves S since in that case the sheaves $R^n f_* S$ are overconvergent too and hence "determined" by their stalks at analytic points.

(2) The proof given above is the one of [P82]. In [S93] the translation of rigid overconvergent sheaves on Z to sheaves on $\mathcal{M}(Z)$ is used to translate the statement into the topological base change theorem for the continuous map $\mathcal{M}(f) : \mathcal{M}(Y) \to \mathcal{M}(X)$.

(3) One aim of this paper is to develop a theory of étale points and étale overconvergent sheaves such that the base change theorem and related theorems are valid.

3 Étale points and étale overconvergent sheaves

A morphism $f : Y \to X$ of analytic spaces over k is called *étale* if for every $y \in Y$ the induced homomorphism of the local rings $O_{X,f(y)} \to O_{Y,y}$ is flat and un-ramified. The term un-ramified means that $O_{Y,y}/mO_{Y,y}$ is a (finite) separable field extension of the field $O_{X,f(y)}/m$ where m denotes the maximal ideal of $O_{X,f(y)}$.

This notion of étale morphism is somewhat complicated. First of all the image of an étale morphism is in general not an admissible open subset. For affinoids Y, X however, it has been shown in [M81] that $f(Y)$ is a finite union of affinoid subdomains of X. We will give a proof of this fact below (see Proposition 3.1.7).

We define the étale site in 3.2 (see [S-S]) and we compare the étale topology with the rigid topology. We define étale points and étale stalks in 3.3. In order to be able to work with étale overconvergent sheaves we construct étale wide neighborhoods in the affinoid case. The proof of the étale base change theorem is then similar to the proof in the rigid case.

3.1 Étale morphisms of affinoids

Let an extension of rings $A \to B$ be given. Let $d : B \to \Omega^f_{B/A}$ denote the universal *finite* differential module of B over A. By definition $\Omega^f_{B/A}$ is a finitely generated B-module and every derivation of B/A into a finitely generated B-module factors uniquely over $d : B \to \Omega^f_{B/A}$. This module exists in many cases where the usual universal differential module of B over A is not a finitely generated module. For affinoid algebras A, B over the same field k one can give B a presentation $B = A\langle T_1, ..., T_n\rangle/(G_1, ..., G_m)$. In this case $\Omega^f_{B/A}$ exists and is the quotient of the free B-module generated by $dT_1, ..., dT_n$ by its submodule generated by $dG_1, ..., dG_m$. Let m be a maximal ideal of B and n the corresponding maximal ideal of A. The completions of the local rings are denoted by $\hat{B}_m$ and $\hat{A}_n$. One can show that $\Omega^f_{\hat{B}_m/\hat{A}_n}$ coincides with $\Omega^f_{B/A} \otimes \hat{B}_m$. In the situation $Y = Spm(B)$ and $X = Spm(A)$ and y, x corresponding to m, n one has that $\hat{B}_m, \hat{A}_n$ are the completions of the local rings $O_{Y,y}$ and $O_{X,f(y)}$. The map between the last two local rings is un-ramified if and only if the map between the completed rings is un-ramified. The last statement is equivalent with $\Omega^f_{\hat{B}_m/\hat{A}_n} = 0$. From this and the fact that flatness is a local property one finds the following:

Observation 3.1.1 *A morphism of affinoid spaces $Y \to X$ is étale if and only if $O(X) \to O(Y)$ is flat and $\Omega^f_{O(Y)/O(X)} = 0$. Further $\Omega^f_{O(Y)/O(X)} = 0$ if and only if the $n \times n$ minors of the matrix $(\frac{\partial G_k}{\partial T_l})$ generate the unit ideal in $O(Y)$.*

A *special étale morphism* of affinoid spaces $f : Y \to X$ is a morphism such that $O(Y)$ has a presentation

$$O(Y) = O(X)\langle T_1, ..., T_n\rangle/(G_1, ..., G_n)$$

such that the functional determinant $\Delta := \det(\partial G_i/\partial T_j)$ is an invertible function on Y. The morphism $Y \to X$ is indeed étale. We need only prove flatness. Let us check this in a point $y \in Y$ where $T_i = 0$. The completion of the local ring $O_{Y,y}$ is isomorphic to:

$$\widehat{O_{X,f(y)}}[[T_1, \dots, T_n]]/(G_1, \dots, G_n)$$

Our assumption on Δ gives that $(G_1, \dots, G_n) = (T_1, \dots, T_n)$ in this ring. Hence the map $O_{X,f(y)} \to O_{Y,y}$ is flat and un-ramified since it induces an isomorphism on completions (this is not true for general points $y \in Y$!).

We now present the proof by Huber of the fact that any étale morphism of affinoids is special étale.

Let $Y \to X$ be an étale morphism of affinoids. Choose a surjection $O(X)\langle T_1, \dots, T_n\rangle \to O(Y)$ with kernel I. Since the module of differentials of $O(Y)$ over $O(X)$ is zero, there is an isomorphism

$$I/I^2 \longrightarrow O(Y)dT_1 \oplus \dots \oplus O(Y)dT_n.$$

Thus we may choose $G_1, \dots, G_n \in I$ whose classes mod I^2 are a basis of I/I^2. It follows that $\mathrm{Sp}\big(O(X)\langle T_1, \dots, T_n\rangle/(G_1, \dots, G_n)\big)$ is equal to $Y \amalg Z$ for some affinoid

Z. Let us choose an element $G \in O(X)\rangle T_1, \ldots, T_n\rangle$ which is 1 on Y and 0 on Z. It follows that $O(Y)$ has the presentation

$$O(Y) = O(X)\langle T, T_1, \ldots, T_n\rangle/(TG - 1, G_1, \ldots, G_n).$$

It follows immediately that Y is special étale over X.

OBSERVATION 3.1.2 *[Hu, 1.7.1] Any étale morphism $Y \to X$ of affinoids is a special étale morphism.*

Let $f : Y \to X$ be an étale morphism of affinoids and choose a representation $O(Y) = O(X)\langle T_1, \ldots, T_n\rangle/(G_1, \ldots, G_n)$ such that $\{\Delta, G_1, \ldots, G_n\}$ generate the unit ideal of the algebra $O(X)\langle T_1, \ldots, T_n\rangle$. If $Z \to X$ is a general morphism, where Z is an affinoid variety over K, then the fibre product $Z \times_X Y$ is given by the affinoid algebra:

$$O(Z)\hat{\otimes}_{O(X)}O(Y) = O(Z)\langle T_1, ..., T_n\rangle/(G_1, ..., G_m)$$

Thus it is clear that $Z \times_X Y \to Z$ is again special étale.

OBSERVATION 3.1.3 *Étale morphisms of affinoids are preserved by general base change. It follows that arbitrary étale morphisms $f : Y \to X$ are preserved by general base change. In particular, if b is an analytic point with image a in X then F_b is a finite separable extension of F_a.*

The following proposition shows that any étale morphism of affinoids can locally be embedded in a finite étale morphism.

PROPOSITION 3.1.4 *Let $f : Y \to X$ be an étale morphism of affinoids. There exists a finite affinoid covering $X = \bigcup X_j$, finite étale morphisms $g_j : Z_j \to X_j$ and open immersions $h_j : f^{-1}(X_j) \to Z_j$ such that $f|_{f^{-1}(X_j)} = g_j \circ h_j$.*

Proof. Let us take an analytic point a of X. By compactness of $\mathcal{M}(X)$ we need only to find a wide neighborhood of a in X over which f can be factored as in the proposition. By Lemma 3.1.6 we may assume that $f^{-1}(\{a\}) = \{b\}$ for some analytic point b of Y. By the above the field extension $F_a \subset F_b$ is finite separable. Hence we can find a wide U in the filter of a and a finite étale morphism $\phi : V \to U$ such that $\phi^{-1}(\{a\})$ consists of one analytic point v with $F_v \cong F_b$ as F_a extensions. See Remark 2.1.2. Let us consider the fibre product $Y \times_X V$ and its projections. The projection to the first factor is a finite étale map onto $f^{-1}(U)$, the projection to the second factor is étale to V and there is an analytic point $c =$"$b \times_a v$" with $F_v = F_c = F_b$. Thus by the lemma below a wide neighborhood W of b is isomorphic to a wide neighborhood of c which is mapped isomorphically to an affinoid subdomain of a wide neighborhood of v. Lemma 3.1.6 shows that replacing U by a smaller wide U we get that $f^{-1}(U) \subset W$ is isomorphic to an affinoid open subdomain of V. □

LEMMA 3.1.5 *Let $f : Y \to X$ be an étale morphism of affinoids and b an analytic point of Y. Put $a = f(b)$. If $F_a \cong F_b$ then there exists a wide U in the filter of a such that $f^{-1}(U) = V \amalg W$ where V is a wide neighborhood of b in Y and the morphism $V \to U$ is an open immersion. If f is finite then $V \to U$ is an isomorphism.*

Proof. In the case that f is finite we may assume that $f^{-1}(\{a\}) = \{b\}$ by replacing X by U as in Lemma 3.1.6 and Y by a connected component of $f^{-1}(U)$. Now $O(Y)$ is a finite locally free $O(X)$-module which we may assume to have constant rank by replacing X by one of its connected components. Our assumptions imply that $F_a \cong O(Y)\hat{\otimes}_{O(X)} F_a \cong O(Y) \otimes_{O(X)} F_a$ hence this rank must be one. This proves the finite case.

The general case. We may replace X by a U as in the lemma below and Y by one of its connected components, hence we may assume that $f^{-1}(\{a\}) = \{b\}$. Let us consider the fibre product $Y \times_X Y$. Since f is étale, the diagonal $\triangle(Y)$ is a union of connected components of $Y \times_X Y$: $\triangle$ is a closed immersion and it is étale (look at local rings!), hence by the finite case above it is also open immersion. Put $Z = Y \times_X Y \setminus \triangle(Y)$, it is an affinoid variety. By assumption, b is not in the image of $\mathrm{pr}_1|_Z : Z \to Y$. Hence we can find a wide neighborhood V of b in Y such that $\mathrm{pr}_1^{-1}(V) \cap Z = \emptyset$. (Use that the spaces $\mathcal{M}(Z)$ and $\mathcal{M}(Y)$ are Hausdorff and compact.) Next we replace X by a wide neighborhood U of a such that $f^{-1}(U) \subset V$ (see lemma below) and Y by $f^{-1}(U)$. We see that $Y \times_X Y \cong Y$. Thus $Y \to X$ is an open immersion ([BGR, 7.3.3], look at complete local rings in ordinary points of Y) and we have won. □

Lemma 3.1.6 *Suppose $f : Y \to X$ is a morphism of affinoids and a is an analytic point of X.*

1. *If $Y_a = \bigcup Y_i$ is the decomposition of the fibre of f into connected components, then there is a wide U in the filter of a such that $f^{-1}(U) = \coprod V_i$ with $(V_i)_a = Y_i$.*
2. *If $Y_a = \{b_1, \ldots, b_s\}$ and we are given wide neighborhoods $W_i \subset Y$ of b_i then we may choose U such that $f^{-1}(U) \subset \bigcup W_i$.*

Proof. The first assertion is a direct consequence of the base change theorem combined with the fact that $f_*\mathbb{Z}$ is overconvergent. For 2) take neighborhoods W_i' of b_i in Y such that $W_i' \subset\subset_Y W_i$. By our Key Lemma we can find a neighborhood U' of a such that $f^{-1}(U') \subset \bigcup W_i'$. For some $U \subset X$ with $U' \subset\subset_X U$ we get $f^{-1}(U) \subset \bigcup W_i$. (Since $\bigcup W_i' \subset\subset_Y \bigcup W_i$, compare with proof of Lemma 2.3.2 part 4.) □

Proposition 3.1.7 *Let $f : Y \to X$ be an étale morphism with Y quasi-compact.*

1. *The image $f(Y)$ of f is a special subset of X, i.e., it is a finite union of open affinoid subvarieties of X.*
2. *An analytic point a of X comes from an analytic point of $f(Y)$ if and only if there exists an analytic point of Y mapping to a.*
3. *The formation of the image of f commutes with general base change: if $X' \to X$ is a general morphism then $f(Y \times_X X') = f(Y) \times_X X'$.*

Proof. We remark that the last assertion follows from the other two.

Let us take an admissible affinoid covering $X = \bigcup X_j$. The admissible covering $Y = \bigcup f^{-1}(X_j)$ has a finite affinoid refinement $Y = \bigcup_{i=1}^n Y_i$. It suffices to prove the proposition for the maps $Y_i \to X_{\alpha(i)}$. Thus we may assume that both X and Y are affinoid. At this point let us prove the assertion on analytic points assuming proven the result on the image. Take an analytic point a of X. If a is not an analytic point of $f(Y)$ then there exists a neighborhood $U \subset X$ of a such that $U \cap f(Y) = \emptyset$. Hence $f^{-1}(U) = \emptyset$ and so $Y_a = \emptyset$. On the other hand, if $a = f(b)$ for some analytic point b of Y then for any U in the filter of a, $f^{-1}(U) \neq \emptyset$. Hence $U \cap f(Y) \neq \emptyset$, hence a is an analytic point of $f(Y)$.

Let us prove the first assertion. Using our preceding proposition we may assume that f factors as $Y \to Z \to X$ where $Y \to Z$ is an open immersion and $Z \to X$ is finite étale. We may also assume that Y is a rational subdomain of Z. We have a morphism

$$\varphi : Z \longrightarrow (\mathbb{P}_n)^{an}$$

with $Y = \varphi^{-1}(R)$ where $R = \{(x_0, \ldots, x_n); |x_0| \geq |x_i|\}$.

Suppose the degree of $Z \to X$ is constant and equal to d. Consider the d-fold fibre product

$$Z^d := Z \times_X Z \times \ldots \times_X Z$$

The diagonals $\triangle_{ij} = \{(z_1, \ldots, z_d) \in Z^d | z_i = z_j\}$ are unions of connected components of Z^d since $Z \to X$ is étale. We put

$$W := Z^d \setminus \bigcup_{i,j} \triangle_{ij}$$

It is an affinoid variety, finite étale over X endowed with an action of S_d (the symmetric group on d letters). The quotient of W under this action is X in the sense that $\Gamma(W, O_W)^{S_d} = \Gamma(X, O_X)$. (Since $Z \to X$ is finite we are doing just algebraic geometry here.) There is a S_d-equivariant map

$$\varphi \times \ldots \times \varphi : W \longrightarrow \mathbb{P}_n^{an} \times \ldots \times \mathbb{P}_n^{an}$$

which descends to a morphism

$$S_d(\varphi) : X \to \big((\mathbb{P}_n)^d / S_d\big)^{an}.$$

It is clear that $f(Y) = S_d(\varphi)^{-1}\big(R(d)/S_d\big)$ with

$$R(d) = \bigcup_i \mathbb{P}_n^{an} \times \ldots \times R \times \ldots \mathbb{P}_n^{an}.$$

There is an obvious formal scheme $(\mathbb{P}^d_{n,k^\circ})^\wedge$ giving rise to $(\mathbb{P}^d_n)^{an}$ and $R(d)$ corresponds to a S_d-stable formal open subscheme of it, namely:

$$\mathcal{U} := \bigcup_i (\mathbb{P}_{n,k^\circ})^\wedge \times \ldots \times (\mathbb{A}_{n,k^\circ})^\wedge \times \ldots \times (\mathbb{P}_{n,k^\circ})^\wedge$$

It follows that $R(d)/S_d$ corresponds to the formal open subscheme $\mathcal{U}/S_d$ of $(\mathbb{P}^d_{n,k^\circ})^\wedge$. Thus $R(d)/S_d$ is a special subset of $\big((\mathbb{P}_n)^d/S_d\big)^{an}$ and hence so is $S_d(\varphi)^{-1}$ of it. $\square$

3.2 The étale site

Let X be an analytic variety over k. In this subsection we recall the definition of the étale site of X (see [S-S, p. 58]). We give a criterium for a presheaf to be a sheaf and we give some examples of étale sheaves. Finally, we prove Hilbert 90 in our situation and we prove that étale cohomology of coherent O-modules agrees with rigid cohomology.

The underlying category of the site $X_{\acute{e}tale}$ will be the category of étale morphisms f of analytic varieties $f : Y \to X$. A morphism of f into f' is a morphism $g : Y \to Y'$ such that $f' \circ g = f$; the morphism g is automatically étale.

We say that a family of étale morphisms $\{g_i : Z_i \to Y\}_{i \in I}$ is an *étale covering* if it has the following property:

> For any (some) choice of admissible affinoid coverings $Z_i = \bigcup_j Z_{i,j}$ we have $Y = \bigcup_{i,j} g_i(Z_{i,j})$ and this is an admissible covering in the G-topology of Y.

This makes sense since the subsets $g_i(Z_{i,j})$ are admissible (special) subsets (see Proposition 3.1.7). We remark that the property is local on Y in the following sense: if $Y = \bigcup Y_l$ is an admissible affinoid covering then $\{g_i : Z_i \to Y\}$ is an étale covering if and only if $\{g_i : g_i^{-1}(Y_l) \to Y_l\}$ is an étale covering for all l. This is so since both assertions are equivalent to the following assertion:

> For each l there are finitely many (i_α, j_α), $\alpha = 1, \dots, n$ such that $Y_l \subset \bigcup_{\alpha=1}^n g_{i_\alpha}(Z_{i_\alpha, j_\alpha})$.

From this it also immediately follows that if $\{Z_i \to Y\}$ is an étale covering and $\{X_{i,j} \to Z_i\}$ are étale coverings then $\{X_{i,j} \to Y\}$ is an étale covering.

Lemma 3.2.1 *Suppose $\{Y_i \to X\}$ is an étale covering and $Z \to X$ is a general morphism. The fibre product $\{Z \times_X Y_i \to Z\}$ is an étale covering.*

Proof. This follows immediately from the definition, the remarks above and Proposition 3.1.7. □

It follows from the above that the category $X_{\acute{e}tale}$, equipped with the family of étale coverings as defined above is a site. It is also clear from the lemma that any (general) morphism $f : Z \to X$ defines a morphism of sites $Z_{\acute{e}tale} \to X_{\acute{e}tale}$ (given by the functor $(Y \to X) \mapsto (Z \times_X Y \to Z)$). The functors on étale sheaves will be denoted by f_* and f^* as usual.

For any object $Y \to X$ of $X_{\acute{e}tale}$ we get a morphism of sites

$$r_{Y/X} : X_{\acute{e}tale} \longrightarrow Y_{rigid},$$

comparing rigid and étale topologies. It is defined by the inclusion of categories $Y_{rigid} \subset X_{\acute{e}tale}$, if S is a sheaf on $X_{\acute{e}tale}$ then $\Gamma(V, (r_{Y/X})_* S) = \Gamma(V, S)$. Sometimes we will use the notation $S|_{Y_{rigid}}$ in stead of $(r_{Y/X})_* S$; we will also use this notation for presheaves S on $X_{\acute{e}tale}$. If $Y = X$ the morphism $r_{X/X}$ will be denoted $r : X_{\acute{e}tale} \to X_{rigid}$. If a is an analytic point of X then we put $S_a := (S|_{X_{rigid}})_a = r_*(S)_a$.

Proposition 3.2.2 *The presheaf S on $X_{\acute{e}tale}$ is a sheaf if and only if the following two conditions hold:*

1. *For any Y in $X_{étale}$ the presheaf $S|_{Y_{rigid}}$ is a sheaf.*

2. *For any surjective finite étale morphism $Y' \to Y$ of affinoids in $X_{étale}$ the sequence $\emptyset \to S(Y) \to S(Y') \rightrightarrows S(Y' \times_Y Y')$ is exact.*

Proof. Suppose S satisfies 1) and 2). We claim that S also satisfies 2) for any finite étale morphism $Y' \to Y$ in $X_{étale}$ with Y quasi-compact. Just cover Y by affinoids as in Lemma 2.5.3 and use 1) to show that it suffices to know 2) for all the resulting affinoid finite étale coverings.

Let us take a morphism $\varphi : Z \to U$ in $X_{étale}$ such that

1. φ is surjective,

2. Z and U are quasi-compact,

3. φ factors as $Z \to V \to U$ with $V \to U$ finite étale and $Z \to V$ an open immersion.

We claim that for any such φ the sequence

$$\emptyset \to S(U) \to S(Z) \rightrightarrows S(Z \times_U Z)$$

is exact. We prove this by induction on the degree of the morphism $V \to U$. (If it is 1 then φ is an isomorphism and our claim trivial.) Suppose therefore that the degree of $V \to U$ is d and that we have proved our claim in the cases where the corresponding degree is less than d.

Since $V \to U$ is finite étale we have that the diagonal $\triangle(V) \subset V \times_U V$ is a union of connected components of $V \times_U V$. Its complement $W \subset V \times_U V$ is thus a quasi-compact variety and the morphism $pr_2 : W \to V$ is finite étale of degree $< d$. Put $Z' = Z \times_U V \cap W$ and $U' = pr_2(Z')$, both are quasi-compact (see Proposition 3.1.7). The surjective étale morphism $\varphi' = pr_2 : Z' \to U'$ factors through $V' := W \cap pr_2^{-1}(U') \to U'$ which is finite étale of degree $< d$. Furthermore, it is clear that $V = U' \cap Z$.

We have the following commutative diagram:

$$\begin{array}{ccccccc} \emptyset & \longrightarrow & S(U) & \longrightarrow & S(Z) & \rightrightarrows & S(Z \times_U Z) \\ & & \downarrow & & \downarrow & & \downarrow \\ \emptyset & \longrightarrow & S(U') & \longrightarrow & S(Z') & \rightrightarrows & S(Z' \times_U' Z') \end{array}$$

The diagram shows that any element $s \in S(Z)$ such that $p_1^*(s) = p_2^*(s)$ gives a unique element (by induction) $s' \in S(U')$ such that $s'|_{Z'} = s|_{Z'}$. It is also true that $s'|_{U' \cap Z} = s|_{U' \cap Z}$ (use induction hypothesis for the morphism $(\varphi')^{-1}(U' \cap Z) \to U' \cap Z$). Hence by 1) for the covering $V = Z \cup U'$ we get a unique section $s_V \in S(V)$ with $s_V|_Z = s$ and $s_V|_{U'} = s'$. We want to show that $p_1^*(s_V) = p_2^*(s_V)$ on $V \times_U V$. Remark that $V \times_U V$ has the following admissible special covering

$$V \times_U V = Z \times_U Z \cup U' \times_U Z \cup Z \times_U U' \cup U' \times_U U'.$$

Hence by 1) we need only to prove $p_1^*(s_V) = p_2^*(s_V)$ on each of these. For the most difficult case, namely $U' \times_U U'$, we remark that the morphism $Z' \times_U Z' \to U' \times_U U'$

is the composition $Z' \times_U Z' \to U' \times_U Z' \to U' \times_U U'$ of morphisms to which our induction hypothesis applies. Hence the map $S(U' \times_U U') \to S(Z' \times_U Z')$ is injective. At this point the commutative diagram

$$\begin{array}{ccc} S(Z) & \rightrightarrows & S(Z \times_U Z) \\ \downarrow & & \downarrow \\ S(Z') & \rightrightarrows & S(Z' \times_U' Z') \end{array}$$

gives the desired result.

To prove that the presheaf S is a sheaf we have to show that any étale covering $\{g_i : Z_i \to Y\}$ in $X_{\acute{e}tale}$ gives an exact sequence

$$\emptyset \longrightarrow S(Y) \longrightarrow \prod S(Z_i) \rightrightarrows \prod S(Z_i \times_Y Z_j).$$

By choosing an admissible affinoid covering $Y = \bigcup Y_j$ and using 1) it is easy to reduce to the case Y affinoid. Similarly we may reduce to the case all Z_i affinoid also. Using propositions 3.1.4 and 3.1.7 we may assume that each $Z_i \to Y$ factors as $Z_i \to V_i \to U_i \subset Y$ as above. It is now easy to deduce the result from our claim above. Compare also with [M80, II 1.5]. $\square$

Examples of sheaves on the étale site. It follows easily from the criterium given above that the following presheaves are sheaves. A general object of $X_{\acute{e}tale}$ will be denoted by $f : Y \to X$.

1. The structure sheaf $\mathbb{G}_a$ defined by $Y \mapsto \Gamma(Y, O_Y)$.

2. The sheaf $\mathbb{G}_m$ defined by $Y \mapsto \Gamma(Y, O_Y^*)$.

3. For any real number r we can look at the subsheaf of $\mathbb{G}_a$ given by $Y \mapsto \{f \in \Gamma(Y, O_Y) : |f(y)| \leq r \forall y \in Y\}$. We can also replace the $\leq$-sign by the $<$ sign. If $r \leq 1$ we can define a subsheaf of $\mathbb{G}_m$ by inequalities of the form $|1 - f(y)| \leq r$.

4. Any representable sheaf $Y \mapsto \mathrm{Mor}_X(Y, Z)$ given by some variety Z over X.

5. For any Abelian group A we have the constant sheaf A_X with stalks A defined by: $Y \mapsto$ the set of maps $Y \to A$ constant on connected components of Y. (This is in fact a representable sheaf, namely represented by $\coprod_{a \in A} X$.)

6. If n is prime to the characteristic of k then we define μ_n as the kernel of the homomorphism $\mathbb{G}_m \to \mathbb{G}_m$ given by multiplication by n. If k contains a primitive n^{th} root of unity ζ then $\mu_n \cong \mathbb{Z}/n\mathbb{Z}_X \cdot \zeta$. There is a Kummer exact sequence
$$1 \longrightarrow \mu_n \longrightarrow \mathbb{G}_m \longrightarrow \mathbb{G}_m \longrightarrow 1.$$

7. Suppose that $\mathcal{F}$ is a coherent sheaf of O_X-modules on X. We define a sheaf $W(\mathcal{F})$ of $\mathbb{G}_a$-modules on $X_{\acute{e}tale}$ as follows: $Y \mapsto \Gamma(Y, f^*\mathcal{F})$, here f^* denotes pullback of coherent O-modules: $f^*\mathcal{F} := f^*(\mathcal{F}) \otimes_{f^*O_X} O_Y$. It is clear that $W(O_X) = \mathbb{G}_a$.

Suppose that we are given an étale covering $\{Y_i \to X\}$. We claim that this coverings allows effective descent of coherent O-modules. This means the following: suppose we are given for each i a coherent O_{Y_i}-module $\mathcal{F}_i$ and descent data. This means isomorphisms of coherent sheaves

$$\varphi_{ij} : pr_1^* \mathcal{F}_i \longrightarrow pr_2^* \mathcal{F}_j$$

on $Y_i \times_X Y_j$ satisfying the co-cycle relation $\varphi_{ij} \circ \varphi_{jk} = \varphi_{ik}$ on $Y_i \times_X Y_j \times_X Y_k$. In this situation there exists a unique coherent sheaf of O_X-modules $\mathcal{F}$ giving rise to $\mathcal{F}_i$ on each Y_i and inducing the isomorphisms φ_{ij}. In addition, homomorphisms of systems $(\mathcal{F}_i, \varphi_{ij})$ as above should correspond to homomorphisms of the corresponding sheaves $\mathcal{F}$. The proof of this is gotten by paraphrasing the proof of Proposition 3.2.2 in this case. Indeed, the question we are considering is whether the association $Y \mapsto$ the category of coherent O_X-modules defines a sheaf of categories. It is clear that rigid coverings and finite étale coverings allow effective descent for coherent O-modules and hence the reasoning of the Proposition applies.

COROLLARY 3.2.3 *Any descent datum for coherent sheaves over an étale covering of X is effective.*

COROLLARY 3.2.4 *For any analytic variety X we have the following isomorphisms:*

$$H^1(X, \mathbb{G}_m) \cong H^1(X, O_X^*) \cong Pic(X)$$

Proof. Of course the group $Pic(X)$ is the group of isomorphism classes of line bundles on X. Since $H^1 = \check{H}^1$ any element in $H^1(X, \mathbb{G}_m)$ can be considered as descent data for invertible O-modules. By the above these are effective and hence come from an element of $H^1(X, O_X^*)$. □

PROPOSITION 3.2.5 *Suppose $\mathcal{F}$ is a coherent sheaf of O_X-modules. The natural maps $H^i(X_{rigid}, \mathcal{F}) \to H^i(X, W(\mathcal{F}))$ are isomorphisms.*

Proof. The maps arise from the identification $r_* W(\mathcal{F}) \cong \mathcal{F}$ and the adjunction map $r^* \mathcal{F} = r^* r_* W(\mathcal{F}) \to W(\mathcal{F})$. Hence the result for $i = 0$. We are going to prove the proposition by induction on i. Take n and suppose the proposition is proven for all $X, \mathcal{F}$ and $i \leq n - 1$.

For any $f : Y \to X$ in $X_{étale}$ consider the map

$$H^n(Y, f^* \mathcal{F}) = H^n(Y, (r_{Y/X})_* W(\mathcal{F})) \longrightarrow H^n(Y, W(\mathcal{F})).$$

This map is injective: by induction hypothesis the sheaves $R^i(r_{Y/X})_* W(\mathcal{F})$ on Y_{rigid} are zero for $i = 1, \ldots, n-1$ (they are the sheaves associated to the presheaves $U \mapsto H^i(U, W(\mathcal{F}))$). Thus the spectral sequence $H^j(Y, R^i(r_{Y/X})_* W(\mathcal{F})) \Rightarrow H^{i+j}(Y, W(\mathcal{F}))$ gives the result. Consider the presheaf $\mathcal{H}^n$ on $X_{étale}$ defined by

$$Y \longmapsto \mathcal{H}^n := \mathrm{Coker}\big(H^n(Y, f^* \mathcal{F}) \to H^n(Y, W(\mathcal{F}))\big).$$

The sheaf associated to this presheaf is zero since any cohomology class in $H^n(Y, W(\mathcal{F}))$ can be killed by an étale covering. Therefore, if we show that $\mathcal{H}^n$ is a sheaf then we are done. To do this we use the criterium from Proposition 3.2.2.

Take any admissible covering $\mathcal{U} : Y = \bigcup Y_j$ of some étale $f : Y \to X$. We have the morphism of spectral sequences

$$\begin{array}{ccc} \check{H}^i(\mathcal{U}, \underline{H}^j(f^*\mathcal{F})) & \Rightarrow & H^{i+j}(Y, f^*\mathcal{F}) \\ \downarrow & & \downarrow \\ \check{H}^i(\mathcal{U}, \underline{H}^j(W(\mathcal{F}))) & \Rightarrow & H^{i+j}(Y, W(\mathcal{F})) \end{array}$$

(see for example [M80, III Proposition 2.7]). We leave it to the reader to verify that this and our induction hypothesis immediately imply that

$$0 \longrightarrow \mathcal{H}^n(Y) \longrightarrow \prod \mathcal{H}^n(Y_j) \longrightarrow \prod \mathcal{H}^n(Y_i \cap Y_j)$$

is exact.

Finally, let $Z \to Y$ be a finite étale morphism of affinoids in $X_{étale}$. Put $A = O(Y)$, $B = O(Z)$ and $M = \Gamma(Y, W(\mathcal{F}))$. We use the notation $Z^n = Z \times_X Z \times \ldots \times_X Z$. It is an affinoid variety. Thus we have that $H^i(Z^n_{rigid}, \mathcal{F} \otimes O^n_Z) = 0$ for all i, n. Furthermore, the complex

$$0 \longrightarrow M \longrightarrow M \hat{\otimes}_A B \longrightarrow M \hat{\otimes}_A B \hat{\otimes}_A B \longrightarrow \ldots$$

is exact. (As the ring extension $A \subset B$ is finite we may replace the completed tensor products by usual ones and then the result is classical.) Thus the spectral sequence $\check{H}^i(\mathcal{U}, \underline{H}^j(W(\mathcal{F}))) \Rightarrow H^{i+j}(Y, W(\mathcal{F}))$ for the covering $\mathcal{U} = \{Z \to Y\}$ and induction hypothesis gives that

$$0 \longrightarrow H^n(Y, W(\mathcal{F})) \longrightarrow H^n(Z, W(\mathcal{F})) \longrightarrow H^n(Z \times_Y Z, W(\mathcal{F}))$$

is exact. We have won. □

Corollary 3.2.6 *Suppose the homomorphism $A \to B$ of affinoid algebras defines a surjective étale morphism of affinoids. For any finite A-module M the complex*

$$0 \longrightarrow M \longrightarrow M \hat{\otimes}_A B \longrightarrow M \hat{\otimes}_A B \hat{\otimes}_A B \longrightarrow \ldots$$

is exact.

3.3 Étale points and stalks

Let us define an étale point of the analytic variety X. An *étale point* e above the analytic point a of X is a separable closure $F_a \subset \mathcal{H}_e$ of F_a. We will always denote by F_e the completion of $\mathcal{H}_e$. Note that the field F_e is algebraically closed (see [BGR, 3.4.1/6]). Therefore an étale point e over a also corresponds to an algebraically closed complete extension $F_a \subset F_e$ such that the algebraic closure of F_a lies dense in F_e. The group $\mathrm{Gal}(\mathcal{H}_e/F_a)$ is equal to the group of continuous F_a-isomorphisms $F_e \to F_e$; this pro-finite group will be denoted $\mathcal{G}_e$.

An *étale neighborhood* of e is a triple (Y, b, ϕ), where Y is a variety étale over X, the analytic point b of Y maps to a and $\phi : F_b \to F_e$ is an F_a-embedding. A morphism $(Y, b, \phi) \to (Y', b', \phi')$ is a morphism $g : Y \to Y'$ over X such that $g(b) = b'$ and $\phi' = \phi \circ g^*$. Two étale neighborhoods (Y_1, b_1, ϕ_1) and (Y_2, b_2, ϕ_2) are dominated by a third one: take $Y = Y_1 \times_X Y_2$, take the point b in Y corresponding to some factor of $F_{b_1} \otimes_{F_a} F_{b_2}$ and $\phi = \phi_1 \otimes \phi_2$. In this way we see that the category of all étale neighborhoods of e give a filtered system.

The *stalk* S_e of a sheaf S on $X_{étale}$ at the étale point e is defined by the formula:

$$S_e := \varinjlim_{(Y,b,\phi)} S(Y)$$

The limit is take over the category of étale neighborhoods of e. If e' (given by $F_a \subset F_{e'}$) is another étale point lying over a, we get by choosing a continuous isomorphism $\psi : F_e \to F_{e'}$ a functor $(Y, b, \phi) \mapsto (Y, b, \psi \circ \phi)$ of the category of étale neighborhoods of e to the category of étale neighborhoods of e'. This gives an isomorphism of stalks

$$\begin{array}{ccc} S_e := \varinjlim_{(Y,b,\phi)} S(Y) & \xrightarrow{\psi^*} & S_{e'} := \varinjlim_{(Y,b,\phi)} S(Y) \\ s \in S(Y) \text{ w.r.t. } (Y,b,\phi) & \longmapsto & s \in S(Y) \text{ w.r.t. } (Y,b,\psi\circ\phi) \end{array}$$

In particular we get an action of $\mathcal{G}_e$ on the stalk functor $S \mapsto S_e$. It is clear that this action is continuous (with the discrete topology on S_e), since any ϕ is stabilized by an open subgroup of $\mathcal{G}_e$.

Let us construct some étale neighborhoods of e. Take a finite Galois extension $F_a \subset L$, say with group G, contained in F_e. By Remark 2.1.2 we can find an affinoid neighborhood U of a in X and a finite étale morphism $g : V \to U$, such that G acts on V over U, $g^{-1}(\{a\}) = \{v\}$ and $F_v \cong L$ (G-equivariant). We may also assume that V and U are connected. It is clear that $(V, v, F_v \to L \subset F_e)$ is an étale neighborhood of e. We claim that these étale neighborhoods are cofinal in the system of all étale neighborhoods of e.

Indeed, given an arbitrary (Y, b, ϕ) take L such that it contains $\phi(F_b)$. The fibre product $Y \times_X V$ contains a point c with $pr_1(c) = b$, $pr_2(c) = v$ and $F_b \to F_c \cong F_v \cong L \subset F_e$ equals ϕ. Using Lemma 3.1.5 we see that there is a commutative diagram

$$\begin{array}{ccc} V' & \longrightarrow & Y \\ \downarrow & & \uparrow \\ V & \longleftarrow & Y \times_X V \end{array}$$

where $V' \subset V$ is an affinoid subdomain containing v. By the Key Lemma we can find a smaller affinoid neighborhood U' of a in X such that $g^{-1}(U') \subset V'$. It is clear that $(g^{-1}(U'), v, F_v \to L \subset F_e)$ is of the form described above and dominates (Y, b, ϕ).

Lemma 3.3.1 *In the situation above.*

1. *The association $S \mapsto S_e$ is an exact functor of the category of étale sheaves on X to the category of continuous $\mathcal{G}_e$-sets.*
2. *For any étale neighborhood (Y, b, ϕ) we have*

   $$\left(S|_{Y_{rigid}}\right)_b = \left((r_{Y/X})_* S\right)_b = H^0(\mathrm{Gal}(\mathcal{H}_e/\phi(F_b)), S_e).$$

   In particular $S_a = (r_* S)_a = H^0(\mathcal{G}_e, S_e)$.
3. *The cohomological dimension $\mathrm{cd}(\mathcal{G}_e)$ of the pro-finite group $\mathcal{G}_e = \mathrm{Gal}(\mathcal{H}_e/F_a)$ is less than or equal to $\dim X + \mathrm{cd}(\mathrm{Gal}(k^{sep}/k))$.*

Proof. (1) We have to show that a surjection of étale sheaves $S \to T$ induces a surjection $S_e \to T_e$. Take an element $t \in T(Y)$ for some étale neighborhood (Y, b, ϕ). There is some étale covering $\{Z_i \to Y\}$ of Y and elements $s_i \in S(Z_i)$ such that s_i maps to the element $t|_{Z_i}$ in $T(Z_i)$. By Lemma 3.2.1 the family $\{(Z_i)_b \to b\}$ is a covering, hence there is some i_0 and analytic point $b_{i_0} \in Z_{i_0}$ mapping to b. Thus (Z_{i_0}, b_{i_0}, χ) is a neighborhood of e (here χ is some extension of ϕ) and s_{i_0} gives an element of S_e lifting t. (Another proof follows from the result $S_e \leftrightarrow i_a^* S$ below.)

(2) We only do the case $X = Y$, $a = b$. Take an element $s \in S_e$, which is fixed by the group $\mathcal{G}_e$. By our results above we may assume that $s \in S(V)$ for some special neighborhood $(V, v, F_v \to L \subset F_e)$ constructed above. Our assumption is that s is stable under the action of G acting on $S(V)$ via its action on V. If we show that $V \times_U V = V \times_X V$ is isomorphic to $V \times G$ then the sheaf property of S will imply that s comes from a unique section of S over U and hence we will be done. However, this again is a consequence of Remark 2.1.2 at least after shrinking U a bit.

(3) We use the notations of Lemma 2.1.1. By [S64, II 4.1] we may replace F_a by K, since this can at most increase the cohomological dimension. The field K is the completion of $k(\underline{t}) = k(t_1, \ldots, t_d)$ for some valuation and some $d \leq \dim X$. But then the group $\mathrm{Gal}(K^{sep}/K)$ is a closed subgroup of the group $\mathrm{Gal}(k(\underline{t})^{sep}/k(\underline{t}))$. We conclude by [S64, I Proposition 14, II Proposition 11]. □

There is a more canonical way to understand the étale stalks S_e. Consider the general morphism

$$i_a : a = \mathrm{Spm}(F_a) \longrightarrow X.$$

It is clear that the category of sheaves on $a_{étale}$ is equivalent to the category of discrete $\mathcal{G}_e$-sets. (Compare [M80, II 1.9].) Therefore i_a^* is a functor of sheaves on $X_{étale}$ to the category of discrete $\mathrm{Gal} - \mathrm{cont}(F_e/F_a)$-sets. This functor is precisely our functor $S \mapsto S_e$. The functor $(i_a)_*$ has the following description: if M is a set with a continuous $\mathcal{G}_e$-action, then

$$\Gamma(Y, (i_a)_* M) = \prod_{b \in Y_a} \big(\mathrm{Hom}_{F_a}(F_b, F_e) \times M\big)^{\mathcal{G}_e}$$

We leave it to the reader as a nice exercise that this functor is exact. It follows from the yoga of adjoint functors that $S \mapsto (i_a)^* S = S_e$ transforms injective sheaves into injective $\mathcal{G}_e$-modules.

Corollary 3.3.2 *There are canonical isomorphisms*

$$\big(R^q r_* S\big)_a \cong H^q(\mathcal{G}_e, S_e).$$

Proof. For $q = 0$ this is the lemma above. It follows for general q by the usual argument using that if S is injective then both sides are zero. (See above.) □

As in the rigid case we do not have enough étale points to separate étale sheaves. To overcome this difficulty we introduce the *étale overconvergent* sheaves: A sheaf S on $X_{étale}$ is said to be *(étale) overconvergent* if $S|_{Y_{rigid}}$ is overconvergent for all Y in $X_{étale}$. Before we can prove interesting properties of these sheaves we need some technical preparations; these will be done in the next section.

3.4 Étale overconvergent sheaves on affinoids

In this section X will be an affinoid variety. Let $f : Y \to X$ be a morphism with Y affinoid and let b be an analytic point of Y. We will say that Y is a *wide neighborhood* of b over X if there exists an affinoid generating system $f_1, \ldots, f_n$ of $O(Y)$ over $O(X)$ such that $|f_i|_b < 1$ for all $i = 1, \ldots, n$. Note that this agrees with our definition in §3.1 in the case that f is an open immersion.

Next we define the notion of relative compactness over X. Let us take a quasi-compact analytic variety Z over X and a quasi-compact open subvariety $Y \subset Z$. We say that Y is *relatively compact* in Z over X, or that Z is a *wide neighborhood* of Y over X, if for any analytic point b of Y there is an affinoid neighborhood $V \subset Z$ of b which is a wide neighborhood of b over X. Notation: $Y \subset\subset_X Z$. Remark that if $Y \subset\subset_Z Y'$ in this situation then also $Y \subset\subset_X Y'$. We note that if both Y and Z are affinoid then this agrees with the definition of [BGR, p. 394] (proof same as proof of [S93, Proposition 23], see also [B90, §2.5]).

Suppose that $f : Y \to X$ is an étale morphism with Y quasi-compact. We want to construct wide neighborhoods of f. We only do this in the case that f is an étale morphism of affinoids. Thus Y is affinoid and $O(Y)$ has a presentation:

$$O(Y) = O(X)\langle T_1, \ldots, T_n\rangle/(G_1, \ldots, G_n)$$

such that $\Delta = \det(\partial G_i/\partial T_j)$ generates the unit ideal of $O(Y)$. A fundamental property of special étale morphisms is that we may always choose this presentation such that $G_1, \ldots, G_n \in O(X)[T_1, \ldots, T_n]$. This follows immediately from the proposition below; in it we use $|R|$ for the supremum norm of an element $R \in O(X)\langle T_1, \ldots, T_n\rangle$.

Lemma 3.4.1 *In the situation above there exists an $\epsilon > 0$ such that if we take any $R_1, \ldots, R_n \in O(X)\langle T_1, \ldots, T_n\rangle$ with $|R_i| < \epsilon$ then we have:*

1. *The affinoid algebra $O(X)\langle T_1, \ldots, T_n\rangle/(G_1 + R_1, \ldots, G_n + R_n)$ defines a special étale morphism $f' : Y' \to X$.*
2. *There exists an isomorphism $Y \cong Y'$ of analytic varieties over X.*

Proof. Let us write $\Delta + R$ for the determinant of the matrix $\big(\partial(G_i + R_i)/\partial T_j\big)$. It is clear that if the R_i have small norm then R has small norm. Since Δ, G_i generate the unit ideal of $O(X)\langle T_1, \ldots, T_n\rangle$ it follows that $\Delta + R, G_i + R_i$ also generate the unit ideal if $|R_i|$ is small enough. This proves (1).

We claim there exists for any positive $\delta < 1$ an $\epsilon > 0$ such that for any affinoid $O(X)$-algebra A the following holds: If there are $a_1, ..., a_n \in A$ with all $|a_i| \le 1$ and all $|G_i(a_1, ..., a_n)| < \epsilon$, then there are $b_1, \ldots, b_n \in A$ such that all $|a_i - b_i| < \delta$ and all $G_i(b_1, ..., b_n) = 0$.

We suppose given $a_1, ..., a_n \in A$ with all $|a_i| \le 1$ and all $|G_i(a_1, ..., a_n)| < \epsilon$, the size of ϵ will be determined later. For an element $b = (b_1, ..., b_n) \in A^n$ we write $\|b\| = \max |b_i|$. Further $G = (G_1, ..., G_n)$ is seen as a map from $\{b \in A^n; \|b\| \le 1\}$ to A^n. Let $\partial G/\partial T$ denote the Jacobian matrix of G. Note that $|(\partial G/\partial T)|$ is bounded from below away from zero on Y, hence is bounded from below by $\eta > 0$ in a neighborhood of the form $|G_i| \le \epsilon_0$, some $\epsilon_0 > 0$. We apply Newton's method; consider the map $Z : b \mapsto b - (\partial G/\partial T(b))^{-1} G(b)$. By the remark above, and by considering a power series expansion of the map G, we see that for δ small enough (so that the quadratic and higher order terms of the power series are negligible) and $\epsilon < \eta\delta$ (for the constant terms) this defines a selfmap of the set $S := \{b \in A^n; \|b-a\| \le \delta\}$. Moreover it is then clear that the map $Z : S \to S$ is a contraction. The fixed point b of the contraction satisfies $G(b) = 0$.

We apply this claim to $A = O(X)\langle T_1, ..., T_n\rangle/(G_1 + R_1, ..., G_n + R_n)$ with the $|R_i| < \epsilon$ and where a_i, $i = 1, ..., n$ is the image of T_i in A. There results a morphism of affinoid $O(X)$-algebras $\alpha : O(X)\langle T_1, ..., T_n\rangle/(G_1, ..., G_n) \to O(X)\langle T_1, ..., T_n\rangle/(G_1 + R_1, ..., G_n + R_n)$, with $\alpha(T_i)$ close to T_i. We can do the same in the other direction to get $\beta : O(X)\langle T_1, ..., T_n\rangle/(G_1 + R_1, ..., G_n + R_n) \to O(X)\langle T_1, ..., T_n\rangle/(G_1, ..., G_n)$, with $\beta(T_i)$ close to T_i. The composition is an endomorphism of $O(Y) = O(X)\langle T_1, ..., T_n\rangle/(G_1, ..., G_n)$ as an $O(X)$-algebra which is close to the identity. It follows that this must be the identity by looking at the graph of it in the fibre product $Y \times_X Y$, where the diagonal is a union of connected components. □

Let us take an étale morphism of affinoids $f : Y \to X$ and take a presentation $O(Y) = O(X)\langle T_1, \ldots, T_n\rangle/(G_1, \ldots, G_n)$ with $G_i \in O(X)[T_1, \ldots, T_n]$. The functional determinant of this presentation $\Delta = \det(\partial G_i/\partial T_j)$ is viewed as a function on $X \times \mathbb{A}^{N,an}$. We define a morphism $f(r) : Y(r) \to X$ for $r \in \sqrt{|k^*|}, r > 1$ as follows:

$$Y(r) = \big\{(x, t_1, \ldots, t_n) \in X \times \mathbb{A}^{n,\mathrm{an}};\ |t_i| \le r \text{ and } G_i(x, t_1, \ldots, t_n) = 0\big\}$$

We claim that if our r is close to 1 then $f(r)$ will again be special étale. To see this we note that there is a presentation:

$$O(Y(r)) = O(X)\langle S_1, \ldots, S_n, T_1', \ldots, T_n'\rangle\big/\big((T_i')^m - \pi^{-m+1}S_i,\ G_i(\pi T_1', \ldots, \pi T_n')\big)$$

Here π is an element of k with $r^m = |\pi|$ and the relation of the coordinates is that $S_i = \pi^{-1}T_i^m$ and $T_i' = \pi^{-1}T_i$. The functional determinant of this presentation is $\pi^{-mn}\Delta|_{Y(r)}$. It is therefore clear that $Y(r) \to X$ is special étale as soon as $\Delta \in \Gamma(Y(r), O_{Y(r)})$ is invertible; this will be the case for r sufficiently close to 1 (the zero locus of Δ lies a positive distance away from Y!). Finally it is clear that $Y \subset\subset_X Y(r)$. We will use the notation $Y(r)$ even if no explicit presentation of $O(Y)$ is given, the number r will always denote an element of $\sqrt{|k^*|}$ bigger than 1 and small enough.

At this point we want to prove the analog of Lemma 2.3.1 in this situation. However, we need to be careful since any étale $U \to X$ has many *non-separated* wide neighborhoods, so the wide neighborhoods $Y \subset Y(r)$ can only be cofinal in the system of separated wide neighborhoods. Although this is in fact true, we restrict ourselves to the case of affinoid varieties.

Lemma 3.4.2 *With notations as above.*

1. *Let W, $U \subset\subset_X V$ be affinoid varieties étale over X and $f : V \to W$ a morphism over X. If V' is a wide neighborhood of U in V, then $f(V')$ is a wide neighborhood of $f(U)$ in W. For varying V' these give a cofinal system of wide neighborhoods of $f(U)$. If $f|_U$ is an isomorphism $U \to f(U)$ then for some $U \subset\subset_V V'$ f induces an isomorphism $V' \to f(V')$.*

2. *If $U \subset\subset_X V$ are affinoid varieties étale over X and $\varphi : Y \to U$ is a morphism then for some $r > 1$ there exists an extension $\varphi(r) : Y(r) \to V$ of φ. This*

extension is unique if r is sufficiently close to 1*. In particular the $Y(r)$, $r > 1$ form a cofinal system of affinoid wide neighborhoods of Y over X.*

3. *Suppose that $Y_i \to X$, $i = 1, \ldots, n$ is étale and Y_i affinoid. We have:*

$$Y_1(r) \times_X \ldots \times_X Y_n(r) = \big(Y_1 \times_X \ldots \times_X Y_m\big)(r)$$

Proof. Suppose we show that if b is an analytic point of U (with image a in W), then there is a wide neighborhood V_b of b in V' such that $f(V_b)$ is a wide neighborhood of a in W. This immediately implies the first statement of (1). The statement on co-finality then follows immediately by letting V' run through the inverse images of such a system of wide neighborhoods of $f(U)$. Let Z denote the complement of the diagonal in $V \times_W V$; it is a union of connected components and hence affinoid. Under the last assumption of (1) we have $pr_1(Z) \cap U = \emptyset$. Thus for some wide neighborhood V' of V we have $V' \times_W V' \cong V'$ and hence it will map isomorphically onto $f(V')$ (compare with Lemma 3.1.5).

Let us construct the neighborhood V_b. By assumption there exists an affinoid generating system $f_1, \ldots, f_r$ of $O(V)$ over $O(X)$ such that $|f_i|_b < 1$. Take a wide neighborhood V_b of b such that $||f_i||_{V_b} < 1$. By Lemma 3.1.6 we can find a wide neighborhood W_a of a in W such that $f^{-1}(W_a) = \bigcup V_i$ as in 3.1.6.1 and $V_1 \subset V_b$ is a wide neighborhood of b. Thus we may replace W by W_a and V by V_1 and assume that $||f_i||_V < 1$ for a generating system $f_1, \ldots, f_r$ of $O(V)$ over $O(W)$. But then V is finite over W [BGR, 9.6.3/6], so we get the existence of V_b by Lemma 3.1.5.

To prove (2) we apply (1) to the projection $Y(r) \times_X V \to Y(r)$. We see that there exists a wide neighborhood of the graph $\Gamma_\varphi \subset Y \times_X U \subset Y(r) \times_X V$ which maps isomorphically onto a wide neighborhood Y' of Y in $Y(r)$. Hence we can find some r', $1 < r' < r$ such that $Y(r') \subset Y'$ (see Lemma 2.3.1). This r' works.

The proof of (3) is formal. □

The lemma above allows us to work with étale morphisms of affinoids only. Therefore we introduce the *special étale site* of X. (Recall that X is affinoid.) It is denoted $X^{sp}_{\acute{e}tale}$ and is defined as follows:

1. Objects are étale morphisms $Y \to X$ with Y affinoid, i.e. special étale ones.

2. Morphisms are morphisms of analytic spaces over X.

3. Coverings are those finite families of morphisms $\{f_i : Y_i \to Y\}$ such that $\bigcup f_i(Y_i) = Y$.

It follows from the remarks made after the definition of special étale morphisms and Lemma 3.2.1 that this is indeed a site. It is functorial with respect to (general) morphisms of affinoids: $Z \to X$ induces a morphism of sites $Z^{sp}_{\acute{e}tale} \to X^{sp}_{\acute{e}tale}$.

The morphism of site $X_{\acute{e}tale} \to X^{sp}_{\acute{e}tale}$, given by the inclusion functor, induces an equivalence of associated topoi. (Use 3.1.2.)

Lemma 3.4.3 *The topos of sheaves on $X^{sp}_{\acute{e}tale}$ is coherent (see [SGA 4, Exposé VI]). In particular, étale cohomology of étale Abelian sheaves on X commutes with filtered direct limits, see [Ibid, 5.2].*

Proof. In the site $X^{sp}_{\acute{e}tale}$ finite fibered products are representable and any object is quasi-compact (see [Ibid, Definition 1.1]). Since it also has a final object we are done, see [Ibid, 2.4.1]. $\square$

The following lemma characterizes overconvergent étale sheaves on X in terms of the site $X^{sp}_{\acute{e}tale}$.

Lemma 3.4.4 *A sheaf S on $X^{sp}_{\acute{e}tale}$ corresponds to an overconvergent sheaf on $X_{\acute{e}tale}$ if and only if the natural map*

$$\varinjlim_{r>1} S(Y(r)) \longrightarrow S(Y)$$

is an isomorphism for all $Y \to X$ affinoid étale.

Proof. Suppose that S is overconvergent. In this case $S|_{Y(r_0)}$ is overconvergent for some $r_0 > 1$. Hence 3.4.4 is an isomorphism since $Y \subset\subset_{Y(r_0)} Y(r)$, $1 < r < r_0$ forms a cofinal system of wide neighborhoods of Y in $Y(r_0)$.

Conversely suppose 3.4.4 is an isomorphism always. Let $Y \to X$ be an étale morphism of affinoids. We have to show that $S|_Y$ is rigid overconvergent. Let $U \subset Y$ be a rational subset of Y. Choose some $r_0 > 1$ such that $Y \subset\subset_X Y(r_0)$ and $Y(r_0)$ is étale over X. Denote for $r > 1$ by $U(r)$ the wide neighborhood of U in $Y(r_0)$ defined in §3.3. It follows easily from Lemma 3.4.2 that these wide neighborhoods $U \subset\subset_X U(r)$ form a cofinal system of affinoid étale wide neighborhoods of U over X. Hence our assumption gives the isomorphism $\lim S(U(r)) = S(U)$.

However, we want to show that the map $\lim S(U(r) \cap Y) \to S(U)$ is an isomorphism. It is clear from the above that this is a surjection. Using for all the rational subdomains $U(r) \cap Y$ of Y the bijectivity of the map $S\big((U(r) \cap Y)(r')\big) \to S(U(r) \cap Y)$, it also follows that the map is injective. This proves our lemma. $\square$

We will say that a presheaf on $X^{sp}_{\acute{e}tale}$ is overconvergent if the map 3.4.4 is always an isomorphism. At this point we introduce a useful method to produce overconvergent (pre)-sheaves on $X^{sp}_{\acute{e}tale}$. Let S be a presheaf on $X^{sp}_{\acute{e}tale}$. We define the presheaf cS on $X^{sp}_{\acute{e}tale}$ as follows:

$$\Gamma(Y, cS) = cS(Y) := \varinjlim_{r>1} \Gamma(Y(r), S)$$

for any Y in $X^{sp}_{\acute{e}tale}$. Note that Lemma 3.4.2 implies that this is independent of the chosen representation of $O(Y)$ over $O(X)$ and that cS is indeed a presheaf. The construction c is a functor, there is an obvious functorial arrow $cS \to S$ and the map $ccS \to cS$ is an isomorphism. Hence the presheaf cS is overconvergent. It is therefore clear that the functor c is a right adjoint of the inclusion functor: overconvergent presheaves on $X^{sp}_{\acute{e}tale} \to$ presheaves on $X^{sp}_{\acute{e}tale}$.

Lemma 3.4.5 *With notations as above.*

1. *If S is a sheaf then cS is a (overconvergent) sheaf. The functor $S \mapsto cS$ is a right adjoint of the inclusion functor: overconvergent sheaves on $X \to$ sheaves on X. The functor $S \mapsto cS$ is left exact on the category of sheaves on $X^{sp}_{\acute{e}tale}$.*

2. *If $\mathcal{J}$ is an injective sheaf on $X^{sp}_{étale}$ and $\mathcal{U} = \{Y_i \to Y\}$ is a covering in $X^{sp}_{étale}$ then*

$$\check{H}^i(\mathcal{U}, c\mathcal{J}) = 0 \ \forall i > 0.$$

It follows that $c\mathcal{J}$ is a flabby sheaf on $X^{sp}_{étale}$.

3. *Any overconvergent sheaf can be embedded into a sheaf of the form $c\mathcal{J}$ with $\mathcal{J}$ injective.*

Proof. Let $\mathcal{U} = \{g_i : Y_i \to Y\}$ be a covering in $X^{sp}_{étale}$. We want to show the following: there exists a set of coverings $\mathcal{U}_\alpha$ such that for any (pre)sheaf S we have a canonical isomorphism:

$$(*) \qquad \varinjlim_{\alpha} \check{\mathcal{C}}(\mathcal{U}_\alpha, S) \cong \check{\mathcal{C}}(\mathcal{U}, cS)$$

(The symbols $\check{\mathcal{C}}$ denote Čech-complexes.) It is clear that this will prove that cS is a sheaf if S is a sheaf and it will prove the second assertion of the lemma. We leave the adjointness property to the reader, as well as the third part of the lemma.

We will only prove the above in the case that the covering $\mathcal{U} = \{g : Z \to Y\}$ is given by one map. Since for an arbitrary (and hence finite) covering in $X^{sp}_{étale}$ there exists a covering consisting of a single morphism giving an isomorphic Čech-Complex there is no loss of generality. To do this we fix $r_0 > 1$ small enough such that $Y(r_0)$ is étale over X and a $r_1 > 1$ small enough such that g extends to $\tilde{g} : Z(r_1) \to Y(r_0)$. Next, for any r_2, $1 < r_2 < r_1$, we choose a $r_3(r_2)$, $1 < r_3(r_2) < r_0$ such that $Y(r_3(r_2)) \subset \tilde{g}(Z(r_2))$. This is possible by Lemma 3.4.2, which also implies that we may choose $r_3(r_2)$ to be a decreasing function of r_2, decreasing to 1 in fact.

We put $Z_{r_2} = Z(r_2) \cap \tilde{g}^{-1}\big(Y(r_3(r_2))\big)$. The coverings we are looking for are $\mathcal{U}_{r_2} = \{Z_{r_2} \to Y(r_3(r_2))\}$. Note that there are commutative diagrams for $1 < r_2' < r_2$:

$$\begin{array}{ccccc} Z & \longrightarrow & Z_{r_2'} & \longrightarrow & Z_{r_2} \\ \downarrow & & \downarrow & & \downarrow \\ Y & \longrightarrow & Y(r_3(r_2')) & \longrightarrow & Y(r_3(r_2)) \end{array}$$

Hence we get the map (*). To show that (*) is an isomorphism we only need to prove that

$$Z \times_Y \ldots \times_Y Z \subset\subset_X Z_{r_2} \times_{Y(r_3(r_2))} \ldots \times_{Y(r_3(r_2))} Z_{r_2}$$

forms a cofinal system of wide neighborhoods of $Z \times_Y \ldots \times_Y Z$ as r_2 decreases to 1. This is clear from the following three facts: 1) $Z_{r_2} \times_X \ldots \times_X Z_{r_2}$ forms a cofinal system of wide neighborhoods of $Z \times_X \ldots \times_X Z$ (see Lemma 3.4.2), 2) $Z \times_Y \ldots \times_Y Z$ is a union of connected components of $Z \times_X \ldots \times_X Z$ and 3) the intersection of $Z_{r_2} \times_{Y(r_3(r_2))} \ldots \times_{Y(r_3(r_2))} Z_{r_2}$ with $Z \times_X \ldots \times_X Z$ is $Z \times_Y \ldots \times_Y Z$. □

Lemma 3.4.6 *(Properties of overconvergent sheaves on $X^{sp}_{étale}$.) In this lemma all (pre-)sheaves are (pre-)sheaves of Abelian groups.*

1. *The sheaf associated to an overconvergent presheaf is overconvergent.*

2. *For any overconvergent sheaf S the presheaves $Y \mapsto H^q(Y, S)$ are overconvergent; for any q the rigid sheaf $R^q(r_{X/Y})_*S$ is overconvergent on Y, in particular the sheaves $R^q r_* S$ are overconvergent on X_{rigid}.*

3. *The category of overconvergent sheaves is an exact subcategory of the category of all sheaves on* $X^{sp}_{étale}$.

4. *If* $f : Z \to X$ *is a general morphism of affinoids then* f^* *and* f_* *preserve overconvergent sheaves. The same holds for* $R^q f_*$ *for any* q.

5. *If* $\{f_i : X_i \to X\}$ *is a special étale covering of* X *then a sheaf on* $X^{sp}_{étale}$ *is overconvergent if and only if each* $f_i^*(S)$ *is overconvergent.*

6. *An overconvergent sheaf* S *is zero if and only if its étale stalks* S_e *are zero for all étale points* e *of* X.

Proof. 1) Let $a(S)$ denote the sheaf associated to S. The map $S \to a(S)$ factors as $S \to ca(S) \to a(S)$ since S is overconvergent. By the universal property of aS we get a section $a(S) \to ca(S)$ (as $ca(S)$ is a sheaf). It follows that $a(S)$ is a direct summand of the overconvergent sheaf $ca(S)$ and hence overconvergent.

2) Embed S in an overconvergent flabby sheaf as in the preceding lemma: $0 \to S \to c\mathcal{J}$. The quotient presheaf is overconvergent hence so is the quotient sheaf Q by 1). For any affinoid Y étale over X we get the exact sequence

$$0 \longrightarrow H^0(Y,S) \longrightarrow H^0(Y,c\mathcal{J}) \longrightarrow H^0(Y,Q) \longrightarrow H^1(Y,S) \longrightarrow 0$$

and isomorphisms $H^q(Y,S) \cong H^{q-1}(Y,Q)$ for $q > 1$. It follows immediately that the presheaf $Y \mapsto H^1(Y,S)$ is overconvergent and the usual induction on q does the rest.

3) Follows from 1) and 2) and the results on rigid overconvergent sheaves.

4) Remark that if $Y \to X$ is affinoid étale then $Y(r) \times_X Z \cong \big(Y \times_X Z\big)(r)$. The rest of the argument is completely analogous to the proof of Lemma 2.3.2 part 4.

5) Same argument as in the rigid case.

6) This is immediate from Lemma 3.3.1 combined with the result for rigid overconvergent sheaves. □

3.5 Étale overconvergent sheaves on general X

Let X be an arbitrary analytic variety over k. Recall that a sheaf S on $X_{étale}$ is overconvergent if $S|_Y$ is rigid overconvergent for any Y étale over X. It is clear from Lemma 3.4.6 that this condition is local in the étale topology on X.

There are now a number of easy consequences of the above which we list here:

1. If $f : Z \to X$ is an arbitrary (general) morphism then f^* preserves overconvergent sheaves.

2. If $f : Z \to X$ is quasi-compact then $R^q f_*$ preserves overconvergent sheaves. (Compare proof of Proposition 2.4.1.)

3. For any overconvergent sheaf S on X the rigid sheaves $R^q(r_{Y/X})_* S$ (in particular $R^q r_* S$) are overconvergent.

Finally, we have the following result.

Proposition 3.5.1 *If* X *is paracompact and* S *is an overconvergent torsion sheaf on* $X_{étale}$ *then* $H^q(X,S) = 0$ *for all* $q > 2\dim X + \mathrm{cd}(k)$, *where* $\mathrm{cd}(k)$ *denotes the cohomological dimension of* k.

Proof. Consider the spectral sequence with E_2-term $H^i(X_{rigid}, R^j r_* S)$ converging to $H^{i+j}(X, S)$. By Corollary 3.3.2 and Lemma 3.4.6 we get that the sheaves $R^j r_* S$ are zero for $j > \dim X + \mathrm{cd}(k)$. Hence we get the result from Corollary 2.5.10. □

3.6 Galois action on cohomology

Let us take a separable closure k^{sep} of k and let us denote by K the completion of k^{sep} with respect to the absolute value $|\,|$. Note that K is algebraically closed (see e.g. [BGR, 3.4]). We remark that the group $\mathcal{G} := \mathrm{Gal}(k^{sep}/k)$ can be identified with the group of continuous automorphisms of K over k.

Take an analytic variety X over k and an étale sheaf S on it. Consider the variety $X\hat{\otimes}K$ over K and the general morphism $\alpha : X\hat{\otimes}K \to X$. For any $\sigma \in \mathcal{G}$ there is an obvious general morphism $\varphi_\sigma : X\hat{\otimes}K \to X\hat{\otimes}K$. This is not a morphism of analytic varieties over K unless $\sigma = \mathrm{id}_K$; it lies over the continuous field homomorphism $\sigma : K \to K$. Since it is clear that $\alpha = \alpha \circ \varphi_\sigma$, we get an isomorphism $\alpha^*(S) \cong (\alpha \circ \varphi_\sigma)^*(S) \cong (\varphi_\sigma)^* \alpha^*(S)$. Thus we get

$$\varphi_\sigma^* : H^i(X\hat{\otimes}K, \alpha^* S) \longrightarrow H^i(X\hat{\otimes}K, \alpha^* S).$$

This defines an action of $\mathcal{G}$ on $H^i(X\hat{\otimes}K, \alpha^* S)$.

Another way to get a $\mathcal{G}$-module is to consider the morphism

$$p : X \to Sp(k)$$

As was noted above the sheaves $R^i p_* S$ correspond to $\mathcal{G}$-modules $(R^i p_* S)_e$. It will be shown below that these two Galois modules agree in the case that X is quasi-compact.

3.7 Étale base change

Let $f : Y \to X$ be a quasi-compact morphism of analytic varieties over k and S an étale sheaf on Y. The étale base change theorem compares the cohomology of S on the étale fibre Y_e with the étale stalks at e of the sheaves $R^q f_* S$. The étale fibre is just defined as $Y_a\hat{\otimes}F_e$, or as the fibre product of the general morphism $Sp(F_e) \to X$ with the morphism $Y \to X$. The result will be an isomorphism of $\mathcal{G}_e$-modules. As in the rigid case the theorem will follow formally from a lemma describing the étale site of the fibre Y_e in the affinoid case.

Therefore we suppose that $f : Y \to X$ is a morphism of affinoids over k and we fix an étale point e lying over the analytic point a of X. For any étale neighborhood (U, b, ϕ) of e with U affinoid we can consider the special étale site of $Y_U := Y \times_X U$. Using ϕ we can see e as an étale point of U lying over b and then it is clear that $(Y_U)_e = Sp(F_e) \times_U Y_U \cong Sp(F_e) \times_X Y = Y_e$. Thus a general morphism $Y_e \to Y_U$ which gives rise to the functor

$$\begin{array}{ccc} (Y_U)^{sp}_{\acute{e}tale} & \longrightarrow & (Y_e)^{sp}_{\acute{e}tale} \\ V \to Y_U & \longmapsto & V_e \to Y_e. \end{array}$$

On the other hand, if the affinoid étale neighborhood (U', b', ϕ') dominates (U, b, ϕ), there is clearly a functor $(Y_U)^{sp}_{\acute{e}tale} \to (Y_{U'})^{sp}_{\acute{e}tale}$ compatible with the functor described above.

Lemma 3.7.1 *The functors above define an equivalence of sites:*

$$\varinjlim_{(U,b,\phi)} (Y_U)^{sp}_{\acute{e}tale} \cong (Y_e)^{sp}_{\acute{e}tale}.$$

Proof. Note that all functors defined above underly morphisms of sites in the reverse directions. The statement follows from the following three assertions:

1. For any étale morphism $V \to Y_e$ with V affinoid there exists an affinoid étale neighborhood (U, b, ϕ) and an étale $W \to Y_U$ morphism of affinoids such that $W_e \cong V$ as varieties over Y_e.

2. Given two étale morphisms $W_i \to Y_U$, W_i affinoid ($i = 1, 2$ and U as above) and a morphism $\psi_e : W_{1,e} \to W_{2,e}$ there exists an affinoid étale neighborhood (U', b', ϕ') dominating (U, b, ϕ) and a morphism $\psi_{U'} : W_{1,U'} \to W_{2,U'}$ such that $\psi_{U',e} = \psi_e$. This $\psi_{U'}$ is unique if U' is small enough.

3. If $\{g_i : W_i \to W\}$ is a finite set of morphisms in $(Y_U)^{sp}_{\acute{e}tale}$ and $\{W_{i,e} \to W_e\}$ is an étale covering then $\{W_{i,U'} \to W_{U'}\}$ is an étale covering if U' is small enough.

Let us prove 1). By definition $O(V)$ has a presentation

$$O(V) \cong O(Y_e)\langle T_1, \ldots, T_n\rangle/(G_1, \ldots, G_n)$$

such that $\Delta := \det(\partial G_i/\partial T_j)$ is invertible. By Lemma 3.4.1 we may suppose that the G_i are polynomials. Since $O(Y_e) = O(Y)\hat{\otimes}_{O(X)}F_e$ we can approximate the G_i by polynomials with coefficients in $O(Y)\otimes_{O(X)}L$ for some finite separable field extension $F_a \subset L \subset F_e$. By Lemma 3.4.1 we may assume $G_i \in O(Y) \otimes_{O(X)} L[T_1, \ldots, T_n]$. We can construct (U, b, ϕ) such that $\phi(F_b) \supset L$ (see 3.3); for this U we can find polynomials $P_i \in O(Y)\hat{\otimes}_{O(X)}O(U)[T_1, \ldots, T_n]$ mapping to the G_i. The function $\Delta(P) := \det(\partial P_i/\partial T_j)$ on

$$W := Sp(O(Y)\hat{\otimes}_{O(X)}O(U)\langle T_1, \ldots, T_n\rangle/(P_1, \ldots, P_n)$$

is such that its restriction to W_e is invertible. Hence, $\Delta(P)$ is invertible on W_b, hence by Lemma 2.7.2 or 2.7.1 we get that $\Delta(P)$ is invertible on W after shrinking U. This gives that $W \to Y_U$ is special étale and $W_e \cong V$ by construction.

Next we do 2). Note that the morphism $\psi_e : W_{1,e} \to W_{2,e}$ gives rise to a graph morphism $\Gamma_e : W_{1,e} \to (W_1 \times_{Y_U} W_2)_e$ and that this morphism identifies $W_{1,e}$ with a union of connected components of $(W_1 \times_{Y_U} W_2)_e$. Hence Γ_e is an étale morphism of affinoids. By 1) there exists a smaller étale neighborhood (U', b', ϕ') and an étale morphism of affinoids $\Gamma_{U'} : W \to (W_1 \times_{Y_U} W_2) \times_U U' \cong W_{1,U'} \times_{Y_{U'}} W_{2,U'}$ with $W_e \cong W_{1,e}$ and $\Gamma_{U',e} = \Gamma_e$. We replace (U, b, ϕ) by (U', b', ϕ') and hence we have $\Gamma : W \to W_1 \times_{Y_U} W_2$. Consider $p_2 = pr_2 \circ \Gamma$. By the above we have that $(p_2)_e$ is an isomorphism. It follows that $(p_2)_b$ is a bijective (on analytic points) étale morphism of affinoids and hence an isomorphism. Thus for any analytic point $c \in W_{1,b}$ we have that $p_2^{-1}(c)$ consists of one analytic point $c' \in W$ with $F_c \cong F_{c'}$. Lemma 3.1.5 implies that p_2 is an open immersion in a wide open neighborhood $W_1(c)$ of c in W_1. Finitely many $W_1(c)_b$'s cover $W_{1,b}$ and $W_b = W_{1,b}$ hence by the key lemma for the rigid case we may shrink U and get that p_2 is an isomorphism (apply the key lemma to both W and W_1). Clearly, the morphism $pr_1 \circ \Gamma \circ (p_2)^{-1} : W_1 \to W_2$ does the job.

The uniqueness follows easily from the rigid key lemma by looking at graphs as above.

Finally, if the assumptions are as in 3) then $W_e = \bigcup_i g_{i,e}(W_{i,e})$ implies $W_b = \bigcup g_{i,b}(W_{i,b})$, since formation of image commutes with arbitrary base change, see Lemma 3.1.7. Thus the statement follows from the rigid case, i.e., if U small enough then $W = \bigcup g_i(W_i)$. □

This was the hard part of the proof of the base change theorem in the étale case. We can deduce the following analog of 2.7.3.

Corollary 3.7.2 *Consider the general morphism $\alpha : Y_e \to Y$.*

1. *The functor α^* preserves flabby sheaves.*
2. *For any sheaf S on $Y^{sp}_{étale}$, any (U, b, ϕ) and any $W \to Y_U$ as above we have:*

$$\Gamma(W_e, \alpha^* S) = \varinjlim_{(U',b',\phi')\geq(U,b,\phi)} \Gamma(W \times_U U', S)$$

Proof. For any S on $Y^{sp}_{étale}$ we have $H^q(Y_e, \alpha^* S) = \lim H^q(Y_U, S)$, by Remark 2.5.9 and the previous lemma. The same argument gives $H^q(W_e, \alpha^* S) = \lim H^q(W_U, S)$ for W as in 2). The results of the corollary follow directly from this. □

Theorem 3.7.3 *Let $f : Y \to X$ be a quasi-compact morphism of analytic spaces over k. Take an étale point e of X and denote by Y_e the (étale) fibre of f at e. The functors*

$$S \mapsto H^q(Y_e, \alpha^* S) \ \ \textit{resp.}\ S \mapsto (R^q f_* S)_e$$

are δ-functors of the category of Abelian sheaves on $Y_{étale}$ to the category of continuous $\mathcal{G}_e$-modules. These δ-functors are isomorphic.

Proof. Remark that $Y_e = Y_a \hat{\otimes} F_e$ and that $\alpha^* S$ is the pullback of $S|_{Y_{a,étale}}$ via the general morphism $Y_e \to Y_a$. Thus we see by 3.6 that the groups $H^q(Y_e, \alpha^* S)$ indeed have a Galois module structure. In the same way as in the proof of the rigid base change theorem it is proved that the functors under consideration form δ-functors. The maps

$$\left(R^q f_* S\right)_e \longrightarrow H^q(Y_e, \alpha^* S)$$

are defined similarly as in the proof of Theorem 2.7.4. These maps commute with Galois action since the action on both sides is defined through the action of $\mathcal{G}_e$ on F_e.

Let us prove that these maps are isomorphisms only in the case that both X and Y are affinoid. The general case then follows as it did in the rigid case. The result for $q = 0$ is just Corollary 3.7.2 part 2) with $W = Y$. The general result follows by induction on q and the fact that α^* preserves flabby sheaves. □

Corollary 3.7.4 *If $f : Y \to X$ is quasi-compact and has finite fibres then $R^q f_* S$ is zero for $q \geq 1$ and any overconvergent sheaf S on $Y_{étale}$. In particular the cohomology of S on Y is equal to the cohomology of $f_* S$ on X.*

Corollary 3.7.5 *(Hochschild-Serre spectral sequence.) Let K be a completion of a separable closure k^{sep} of k. Let $\mathcal{G} = \mathrm{Gal}(k^{sep}/k)$ denote the continuous Galois group of K over k. For any quasi-compact variety X over k and any Abelian sheaf S on $X_{étale}$ there is a spectral sequence*

$$H^i(\mathcal{G}, H^j(X\hat{\otimes}K, \alpha^*S)) \Rightarrow H^{i+j}(X, S).$$

Here $\alpha : X\hat{\otimes}K \to X$ is as in 3.6 and $H^q(\mathcal{G}, -)$ denotes continuous cohomology.

Proof. Let us write $p : X \to \mathrm{Sp}(k)$ as in 3.6. Let a denote the unique analytic point of $\mathrm{Sp}(k)$ and let e be an étale point lying over a. First we note that (if S is overconvergent)

$$H^q(X\hat{\otimes}K, \alpha^*S) \cong (R^q p_* S)_e$$

as $\mathcal{G}$-modules by the theorem above. This shows that the $\mathcal{G}$-module on the left has a continuous $\mathcal{G}$-action if it is given the discrete topology. It also proves that

$$H^0(\mathcal{G}, H^0(X\hat{\otimes}K, \alpha^*S)) = H^0(\mathcal{G}, (p_*S)_e) = (p_*S)_a = H^0(\mathrm{Sp}(k), p_*S) = H^0(X, S).$$

Here we used Lemma 3.3.2. Hence we only need to show that the functor which maps S to the Galois module $H^0(X\hat{\otimes}K, \alpha^*S)$ transforms an injective sheaf S on X into an acyclic $\mathcal{G}$-module. Since we are taking continuous cohomology we have:

$$H^q(\mathcal{G}, H^0(X\hat{\otimes}K, \alpha^*S)) = \varinjlim_{k\subset k'} H^q(\mathrm{Gal}(k'/k), H^0(X\hat{\otimes}K, \alpha^*S)^{\mathcal{G}'}),$$

where the limit runs over all finite Galois extensions $k \subset k'$ contained in K. By an argument as above this is the limit over the groups

$$H^q(\mathrm{Gal}(k'/k), H^0(X \otimes k', S)).$$

But since S is injective these groups compute the cohomology groups $H^q(X, S)$ (compare [M80, Theorem 2.20]) and these are zero for $q \geq 1$. □

4 Cohomology of varieties of dimension at most 1

In this section we suppose that the field k is algebraically closed. Let $p \geq 1$ denote the characteristic of the residue field of k. We put $p = 1$ if the residue field of k contains the field of rational numbers. Further X will denote an analytic space over k of dimension ≤ 1.

4.1 Some general results

For $n > 1$ which is not divisible by the characteristic of k, we consider the exact sequence

$$0 \longrightarrow \mu_n \longrightarrow \mathbb{G}_m \xrightarrow{n} \mathbb{G}_m \longrightarrow 0$$

of sheaves on $X_{étale}$. This sequence induces the following distinguished triangle of complexes on X_{rigid}:

$$\longrightarrow Rr_*\mu_n \longrightarrow Rr_*\mathbb{G}_m \longrightarrow Rr_*\mathbb{G}_m \longrightarrow Rr_*\mu_n[1]$$

We already know quite a lot about the homology sheaves of these complexes: We know that $R^q r_* \mu_n = 0$ for all $q \geq 2$ by Lemma 3.3.1 and Corollary 3.3.2 combined with the fact that $R^q r_* \mu_n$ are overconvergent (Lemma 3.4.6). Further we know that $R^1 r_* \mathbb{G}_m = 0$ by Corollary 3.2.4. The following result is a formal consequence of this.

Lemma 4.1.1 *In the derived category of Abelian sheaves on X_{rigid} we have the following isomorphism*

$$Rr_*\mu_n \cong \left(O_X^* \xrightarrow{n} O_X^*\right),$$

where the first term on the right is placed in degree 0.

Lemma 4.1.2 *Let X be connected paracompact (and still have dimension ≤ 1). We denote by $\mathbb{Z}/n\mathbb{Z}_X$ the constant sheaf with fibre $\mathbb{Z}/n\mathbb{Z}$ on $X_{\acute{e}tale}$, where n is prime to the characteristic of k. We have $H^0(X, \mathbb{Z}/n\mathbb{Z}_X) = \mathbb{Z}/n\mathbb{Z}$ and $H^q(X, \mathbb{Z}/n\mathbb{Z}_X) = 0$ for $q \geq 3$.*

1. *There is an exact sequence*

$$0 \longrightarrow H^1(X_{rigid}, \mathbb{Z}/n\mathbb{Z}) \longrightarrow H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \longrightarrow H^0(X_{rigid}, R^1 r_* \mathbb{Z}/n\mathbb{Z}_X) \to 0$$

and we have $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) = H^1(X_{rigid}, R^1 r_* \mathbb{Z}/n\mathbb{Z}_X)$.

2. *A choice of a primitive n^{th}-root of unity determines an exact sequence*

$$0 \longrightarrow O(X)^*/O(X)^{*\,n} \longrightarrow H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \longrightarrow \ker(n, Pic(X)) \longrightarrow 0$$

and an isomorphism $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) = Pic(X)/nPic(X)$.

Proof. The statement on H^0 is trivial. Consider the spectral sequence with E_2-terms $H^p(X_{rigid}, R^q r_* \mathbb{Z}/n\mathbb{Z}_X)$ abutting to $H^{p+q}(X, \mathbb{Z}/n\mathbb{Z}_X)$. Clearly 1) follows since $R^q r_* \mathbb{Z}/n\mathbb{Z}_X = (0)$ for $q \geq 2$ and cohomology of rigid sheaves is zero on X in dimensions ≥ 2 by Corollary 2.5.10.

A choice of a primitive n^{th}-root of unity determines an isomorphism of sheaves $\mathbb{Z}/n\mathbb{Z}_X \cong \mu_n$. Thus statement 2) follows from the lemma above and the vanishing of rigid cohomology in degrees ≥ 2 on X. □

4.2 The cohomology of $\mathbb{Z}/n\mathbb{Z}$ with $(n,p) = 1$

In this subsection we will determine the cohomology of X in certain cases where X is smooth and irreducible. We will use the word *curve* to denote a separated analytic variety of pure dimension 1. Recall that we are working over an algebraically closed field.

Proposition 4.2.1 *Let C be a nonsingular projective curve of genus g. We compute the cohomology with values in $\mathbb{Z}/n\mathbb{Z}$ for $(n,p) = 1$ of an open subvariety X of C as follows.*

1. *If $X = C$, then $H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \cong (\mathbb{Z}/n\mathbb{Z})^{2g}$ and $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) \cong \mathbb{Z}/n\mathbb{Z}$.*
2. *If X is the complement of finitely many points $c_1, \ldots, c_a$ $(a > 0)$ in C, then $H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \cong (\mathbb{Z}/n\mathbb{Z})^{2g+a-1}$ and $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) = 0$.*

3. Suppose X is $C \setminus (D_1 \cup \ldots \cup D_a)$ where the D_i are disjoint open discs in C. In this case $H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \cong (\mathbb{Z}/n\mathbb{Z})^{2g+a-1}$ and $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) = 0$.

In all of these cases, for any extension of algebraically closed complete valued fields $k \subset K$, the natural map $H^q(X, \mathbb{Z}/n\mathbb{Z}_X) \to H^q(X_K, \mathbb{Z}/n\mathbb{Z}_{X_K})$ is an isomorphism.

Proof. Part 1) follows from Lemma 4.1.2 part 2) and the fact that $Pic(C)$ corresponds to the algebraic Picard group of C by GAGA. Note that the isomorphism for H^2 is given by the isomorphism $Pic(C)/nPic(C) \cong \mathbb{Z}/n\mathbb{Z}$ induced by taking degrees of line bundles on C.

Note that in case 2) the space X is the admissible increasing union $X = \bigcup X_n$ of affinoid spaces X_n as in 3). Just take $D_{i,n}$ to be smaller and smaller open discs in C with center c_i. Thus if we prove 3) then 2) will follow by considering the Cartan-Leray spectral sequence associated to the covering $X = \bigcup X_n$. (Here we also need that the maps $H^i(X_{n+1}) \to H^i(X_n)$ are isomorphisms; this follows from the proof of 3) below.)

Let us assume X is as in 3). Any line bundle on X is the restriction of a line bundle of degree zero on C. In other terms, $Pic^0(C) \to Pic(X)$ is surjective. In particular $Pic(X)$ is a divisible group and by Lemma 4.1.2 we get $H^2(X, \mathbb{Z}/n\mathbb{Z}_X) = 0$. For the calculation of the group $H^1(X, \mathbb{Z}/n\mathbb{Z}_X)$ we start with the case where X is the closed unit disk $\mathbb{B} := \{z \in k; |z| \leq 1\}$.

Now $H^1(\mathbb{B}, \mathbb{G}_m) = Pic(\mathbb{B}) = 0$ and by 4.1.2 one has $H^1(\mathbb{B}, \mathbb{Z}/n\mathbb{Z}_{\mathbb{B}}) = O(\mathbb{B})^*/O(\mathbb{B})^{*n}$. The invertible functions on $\mathbb{B}$ are of the form $\lambda(1+f)$ with $\lambda \in k^*$ and $f \in O(\mathbb{B})$ has a norm < 1. The condition on n implies that such a function has an n-th root. Hence $H^1(\mathbb{B}, \mathbb{Z}/n\mathbb{Z}_{\mathbb{B}}) = 0$.

Next, we want to investigate a ring domain (or annulus) $\partial\mathbb{B} := \{z \in k; |z| = 1\}$. Again $Pic(\partial\mathbb{B}) = 0$. Further every invertible function on $\partial\mathbb{B}$ has uniquely the form $\lambda z^s(\sum_m a_m z^m)$ where $\lambda \in k^*$, $s \in \mathbb{Z}$ and where the Laurent series satisfies $a_0 = 1$, $|a_m| < 1$ for all $m \neq 0$ and $\lim |a_m| = 0$. It follows that $H^1(\partial\mathbb{B}, \mathbb{Z}/n\mathbb{Z}_{\partial\mathbb{B}}) = O(\partial\mathbb{B})^*/O(\partial\mathbb{B})^{*n} = \mathbb{Z}/n\mathbb{Z}$, a generator is given by the class of z. Clearly this is independent of the base field k.

Now we start proving the general statement. The pre-sheaves $U \mapsto H^i(U_{etale}, \mathbb{Z}/n\mathbb{Z}_C)$ are overconvergent on C_{rigid}. Hence it suffices to prove the statement for all wide neighborhoods X' of X in C. For such an X' we can find closed unit discs $B_i \subset D_i$ such that $X' \cap B_i$ is isomorphic to a ring domain $\partial\mathbb{B}$. If we have this then the covering

$$C = X' \cup \bigcup B_i$$

will be admissible. In particular it is also an étale covering of C. Therefore, we have the Mayer-Vietoris sequence exact sequence [M80, p. 110]

$$\begin{aligned} 0 \longrightarrow H^1(C, \mathbb{Z}/n\mathbb{Z}_C) \longrightarrow H^1(X', \mathbb{Z}/n\mathbb{Z}_C) \oplus \bigoplus H^1(B_i, \mathbb{Z}/n\mathbb{Z}_C) \\ \longrightarrow \bigoplus H^1(X' \cap B_i, \mathbb{Z}/n\mathbb{Z}_C) \longrightarrow H^2(C, \mathbb{Z}/n\mathbb{Z}_C) \longrightarrow 0. \end{aligned}$$

The zero on the right follows by the vanishing of H^2 on affinoid curves proved above and the zero on the left is trivial to establish. The result follows by the computation of cohomology of $\mathbb{B}$ and $\partial\mathbb{B}$ given above.

For a precise definition of the map $H^q(X, \mathbb{Z}/n\mathbb{Z}_X) \to H^q(X_K, \mathbb{Z}/n\mathbb{Z}_{X_K})$ see 5.1 below. The invariance of cohomology under extension of base field for X follows from the invariance of cohomology for the algebraic curve C and the spaces $\mathbb{B}$, resp. $\partial\mathbb{B}$. $\square$

Remark 4.2.2 Let $k_0 \subset k$ denote a complete subfield of k such that k is the completion of the separable algebraic closure of k_0. Suppose that the g.c.d.$(n, p) = 1$. Let $\partial\mathbb{B}$ be the ring domain $\{z \in k_0; |z| = 1\}$ over k_0. The Galois action (see 3.6) on $H^1(((\partial\mathbb{B})\hat{\otimes}_{k_0} k), \mu_n) = \mathbb{Z}/n\mathbb{Z}$ is trivial. By the proof above this cohomology group is canonically isomorphic to $O(\partial\mathbb{B}\hat{\otimes}_{k_0} k)^*/O(\partial\mathbb{B}\hat{\otimes}_{k_0} k)^{*\,n}$. The generator of this group is the class of the invertible function z. This is clearly invariant under the Galois group.

Remark 4.2.3 The open unit disc D is the increasing union of closed discs. Thus we see, by the argument that proved part 2 of the proposition, that $H^q(D, \mathbb{Z}/n\mathbb{Z}_D) = 0$ for $q \geq 1$. This result is partially generalized in the corollary below.

Corollary 4.2.4 *Let $\mathcal{L}$ be a compact subset of $\mathbb{P}^1_k$ and put $X = \mathbb{P}^1_k \setminus \mathcal{L}$. Then $H^1(X, \mathbb{Z}/n\mathbb{Z}_X)$ coincides with the group of $\mathbb{Z}/n\mathbb{Z}$-valued currents on the tree of X (or the tree of $\mathcal{L}$). More generally, for any connected open subspace X of $\mathbb{P}^1_k$, the group $H^1(X, \mathbb{Z}/n\mathbb{Z}_X)$ is equal to $O(X)^*/O(X)^{*n}$.*

Proof. The line bundles on any open subspace X of $\mathbb{P}^1_k$ are trivial (see [FP]) and hence $H^1(X, \mathbb{Z}/n\mathbb{Z}_X) \cong O(X)^*/O(X)^{*n}$. (Use Lemma 4.1.2.) The structure of the group $O(X)^*$ is well known if $X = \mathbb{P}^1_k - \mathcal{L}$. Namely, there is an exact sequence

$$0 \longrightarrow k^* \longrightarrow O(X)^* \longrightarrow C(T) \to 0,$$

where T denotes the tree of $\mathcal{L}$ and where $C(T)$ denotes the group of currents with values in $\mathbb{Z}$ on T. (See [FP].) It follows that $O(X)^*/O(X)^{*n} = C(T)/nC(T)$ is the group of currents on T with values in $\mathbb{Z}/n\mathbb{Z}$. $\square$

Proposition 4.2.5 *If X is a connected smooth affinoid curve then there is an embedding $X \subset C$ as in Proposition 4.2.1 part 3) above. We deduce from this the following results. Let A be an Abelian torsion group of exponent n, with $(n, p) = 1$.*

1. *There are natural isomorphisms $H^q(X, \mathbb{Z}/n\mathbb{Z}_X) \otimes A \cong H^q(X, A_X)$.*
2. *The cohomology groups $H^q(X, A_X)$ are invariant under algebraically closed extensions of base fields.*

Proof. The existence of such an embedding $X \to C$ is proved in [P80]. The group A is the direct limit of its finite subgroups. Taking cohomology commutes with direct limits (3.4.3), hence it suffices to do the case A is finite. Writing A as the direct sum of cyclic subgroups it follows that we may assume $A \cong \mathbb{Z}/n'\mathbb{Z}$ where $n'|n$. In this case both 1) and 2) follow easily from Proposition 4.2.1. $\square$

Remark 4.2.6 Other constant sheaves.

1. The cohomology of $\mathbb{Q}_X$. Since in this case the rigid sheaves $R^q r_* \mathbb{Q}_X$ for $q \geq 1$ are both torsion (by 3.3.2 and 3.4.6) and sheaves of $\mathbb{Q}$-vector spaces, they are zero. Hence we have
$$H^q(X, \mathbb{Q}_X) \cong H^q(X_{rigid}, \mathbb{Q})$$

for all q. If X is a separated quasi-compact smooth curve then we have

$$H^1(X, \mathbb{Q}_X) = H^1(X_{rigid}, \mathbb{Q}) = \mathbb{Q}^b,$$

where b is the Betti number of the graph of a semi-stable reduction of X. A semi-stable reduction of X is defined as follows: take a separated formal scheme $\mathfrak{X}$ of finite type over $\mathrm{Spf}(k^\circ)$, whose associated rigid space $\mathfrak{X}^{rig}$ is isomorphic to X. (See [R74] or [BL].) By blowing up $\mathfrak{X}$ a bit we may assume that the singularities of the special fibre are ordinary double points. This special fibre is a semi-stable reduction of X. Since any other such formal scheme $\mathfrak{X}'$ may be compared with $\mathfrak{X}$ by a sequence of blow ups and blow downs in points it follows that the associated graphs have the same homotopy type. The result now follows from Remark 2.5.9 and a computation of the Zariski cohomology of a constant sheaf on an algebraic semi-stable curve.

2. The constant sheaf $\mathbb{Z}/p\mathbb{Z}_X$.

   (a) Let the characteristic of k be $p > 1$. We will give a calculation of $H^1(\mathbb{B}, \mathbb{Z}/p\mathbb{Z}_X)$ where $\mathbb{B}$ is the closed unit disk. Consider the Artin-Schreier exact sequence

   $$0 \longrightarrow \mathbb{Z}/p\mathbb{Z}_{\mathbb{B}} \longrightarrow \mathbb{G}_a \xrightarrow{\phi} \mathbb{G}_a \longrightarrow 0$$

   on $X_{\acute{e}tale}$. Here of course $\phi(f) = f^p - f$. On cohomology we get an exact sequence

   $$0 \longrightarrow \mathbb{Z}/p\mathbb{Z} \longrightarrow O(\mathbb{B}) \xrightarrow{\phi} O(\mathbb{B}) \longrightarrow H^1(\mathbb{B}, \mathbb{Z}/p\mathbb{Z}_{\mathbb{B}}) \longrightarrow 0,$$

   since $H^1(\mathbb{B}, \mathbb{G}_a) = (0)$ by 3.2.5. The co-kernel of $\phi : O(\mathbb{B}) \to O(\mathbb{B})$ is a rather large group and is not invariant under algebraically closed extensions of base fields. This reflects the fact that the closed disk has many p-cyclic un-ramified coverings.

   (b) Here the characteristic of k is zero, but the characteristic of the residue field $\tilde{k}$ is $p > 1$. With the methods above it follows that $H^1(\mathbb{B}, \mathbb{Z}/p\mathbb{Z}_{\mathbb{B}}) = O(\mathbb{B})^*/O(\mathbb{B})^{*p}$. This is again a very large group not invariant under base field extensions. It can be shown that every (algebraic) finite étale covering of the affine line over $\tilde{k}$ lifts to a finite étale covering of $\mathbb{B}$. The conjecture of S.S.Abhyankar on the coverings of the affine line in characteristic p (proved by M.Raynaud) implies that the totality of nontrivial finite étale coverings of $\mathbb{B}$ is very large.

5 Base change revisited

In this section we prove a general base change theorem for quasi-compact morphisms and overconvergent étale sheaves. In order to be able to apply Theorem 3.7.3 we have to prove invariance of cohomology under extensions of algebraically closed base fields. This is done below for rigid and étale cohomologies and overconvergent sheaves.

5.1 A change of fields, étale case.

Let $k \subset K$ be an extension of complete and algebraically closed fields. For any analytic space X over k we denote by $p_K : X_K = X\hat{\otimes}_k K \to X$ the general morphism associated to the change of fields. For an étale sheaf S on X we have the étale sheaf $p_K^* S$ on X_K and comparison maps

$$H^q(X, S) \longrightarrow H^q(X_K, p_K^* S).$$

We would like to know when these are isomorphisms. As before we put $p \geq 1$ equal to the characteristic of the residue field of k (and $p = 1$ if $\mathrm{char}(\tilde{k}) = 0$).

Theorem 5.1.1 *The canonical maps $H^q(X, S) \longrightarrow H^q(X_K, p_K^* S)$ are isomorphisms if S satisfies the following conditions:*

1. *The sheaf S is overconvergent.*
2. *All étale stalks S_e of S are torsion groups, with torsion prime to p.*

Proof. By taking an admissible affinoid covering of X we see that it suffices to do the case that X is affinoid. Let us consider S as a sheaf on the site $X^{sp}_{\acute{e}tale}$. For any $n \in \mathbb{N}$ with $(n, p) = 1$ let $S_n \subset S$ be the subsheaf of S consisting of sections annihilated by n, i.e., $S_n := \mathrm{Ker}(S \stackrel{n}{\to} S)$ is also overconvergent. By our two conditions on S and Lemma 3.3.1 we see that any section $s \in S(Y)$ is torsion ($Y \to X$ is special étale, hence Y affinoid, hence quasi-compact). Thus we see that $S = \bigcup S_n$. By looking at stalks we see that $(p_K^* S)_n = p_K^* S_n$, hence also $p_K^* S = \bigcup p_K^* S_n$. Since cohomology commutes with direct limits (3.4.3) it suffices to do the case that S is a sheaf of $\mathbb{Z}/n\mathbb{Z}$-modules.

Consider fields L with $k \subset L \subset K$, which are complete and algebraically closed. We say that L has topological transcendence degree $\leq r$ over k if there exist elements $t_1, \ldots, t_r \in L$ such that L is the completion of the algebraic closure of $k(t_1, \ldots, t_r)$. The reasoning of Lemma 3.7.1 shows that the site $(X_K)^{sp}_{\acute{e}tale}$ is the direct limit of the sites $(X_L)^{sp}_{\acute{e}tale}$, taken over all L of finite topological transcendence degree over k. Therefore it suffices to prove $H^q(X, S) = H^q(X_L, p_L^* S)$ for $k \subset L$ of topological transcendence degree $\leq r$. By induction on r it suffices to do the case: $k \subset K$ of topological transcendence degree 1.

Take an element $t \in K$ such that K is the completion of the algebraic closure of $k(t)$. We may assume that $|t| \leq 1$. Consider the continuous k-algebra homomorphism $k\langle T\rangle \to K$ mapping T to t. This determines an étale point e of the closed unit disc $\mathbb{B}$ over k with $F_e = K$.

The problem we are studying may now be formulated with the help of the following diagram of analytic spaces and general morphisms.

$$\begin{array}{ccccc} X_K & \longrightarrow & X \times \mathbb{B} & \xrightarrow{p_1} & X \\ \downarrow & & \downarrow{\scriptstyle p_2} & & \downarrow{\scriptstyle q_2} \\ Sp(K) & \xrightarrow{e} & \mathbb{B} & \xrightarrow{q_1} & Sp(k). \end{array}$$

There is a general base change morphism (see [SGA 4, Exp. XVII 4.1.5])

$$q_1^* R^q(q_2)_* S \longrightarrow R^q(p_2)_* p_1^* S \qquad (1).$$

The comparison map from the theorem is the base change map for the big rectangle of the diagram. Our étale base change theorem 3.7.3 asserts that the base change map for the left square with p_1^*S is an isomorphism. By the functoriality properties of the base change morphism it will suffice to prove that (1) is an isomorphism.

Let $V \to \mathbb{B}$ be étale and V affinoid. Put $p : X \times V \to X$ equal to the projection. We write $H^m(V) = H^m(V, \mathbb{Z}/n\mathbb{Z}_V)$ and $H^m(V)_X$ is the constant sheaf with fibre $H^m(V)$ on $X_{étale}$. There is a natural map

$$H^m(V)_X \otimes_{\mathbb{Z}/n\mathbb{Z}} S \longrightarrow R^m p_* p^* S. \tag{2}$$

It is the composition

$$H^m(V)_X \otimes S \longrightarrow R^m p_*(\mathbb{Z}/n\mathbb{Z}_{X\times V}) \otimes S \longrightarrow R^m p_* p^* S,$$

the first map given by base change, the second deduced from $S \to p_*p^*S$ by the cup-product $R^m p_*(\mathbb{Z}/n\mathbb{Z}_{X\times V}) \otimes R^0 p_* p^* S \to R^m p_* p^* S$ associated to $\mathbb{Z}/n\mathbb{Z}_{X\times V} \otimes p^*S \to p^*S$. By étale base change 3.7.3 the stalk of $R^m p_* p^* S$ in the étale point f of X is $H^m(V\hat{\otimes}F_f, (S_f)_{V\hat{\otimes}F_f})$. Hence, by Proposition 4.2.5 (2) is an isomorphism on étale stalks for all f. Since both sides of (2) are overconvergent we get that (2) is an isomorphism. Thus we get that $R^m p_* p^* S = (0)$ for $n \geq 2$. Finally, since $p : X \times V \to X$ has a section, the maps $H^m(X, S) \to H^m(X \times V, p^*S)$ have sections. We conclude that the spectral sequence $H^i(X, R^j p_* p^* S) \Rightarrow H^{i+j}(X \times V, S)$ degenerates and gives:

$$\begin{array}{rclcl} H^m(X \times V, S) & \cong & H^m(X, H^0(V)_X \otimes S) & \oplus & H^{n-1}(X, H^1(V)_X \otimes S) \\ & \cong & H^0(V) \otimes H^m(X, S) & \oplus & H^1(V) \otimes H^{n-1}(X, S) \end{array}$$

Therefore, the sheaf associated to the presheaf $V \mapsto H^m(X \times V, S)$ (on $\mathbb{B}^{sp}_{étale}$) is the constant sheaf with fibre $H^m(X, S)$. Clearly this means that the right side of (1) is constant and hence that (1) is an isomorphism (look at fibres over $0 \in \mathbb{B}$). □

5.2 A change of fields, rigid case.

We think it is quite amusing that a similar theorem also holds for the rigid case. Notations are as in 5.1.

Theorem 5.2.1 *Let S be an overconvergent sheaf on X_{rigid}. The canonical maps*

$$H^q(X_{rigid}, S) \longrightarrow H^q((X_K)_{rigid}, p_K^* S)$$

are isomorphisms.

Proof. As in the proof of the étale case we may assume that X is affinoid and $k \subset K$ of topological transcendence degree 1 (using X^{rat}_{rigid} in stead of $X^{sp}_{étale}$). We consider subfields $L \subset K$, which are complete and are the completion of a function field of transcendence degree 1 over k. In this case we remark that $(X_K)^{rat}_{rigid}$ is the direct limit of the sites $(X_L)^{rat}_{rigid}$ for such fields L. Again it suffices to do the case $K = L$. (The field K is no longer algebraically closed!)

Suppose Z is a nonsingular projective irreducible curve over k, whose function field $k(Z)$ is a dense subfield of K. The embedding $k(Z) \to K$ defines an analytic point a of Z with $F_a = K$.

OBSERVATION 5.2.2 *There is an affinoid subdomain $U \subset Z$ in the filter of a with the following property: For every affinoid $V \subset U$ and any constant sheaf T on U the cohomology groups $H^n(V,T)$ are zero for $n \geq 1$.*

This follows quite easily from the stable reduction of Z. As was noted in Remark 4.2.6 part 1) the cohomology of a rigid constant sheaf on an affinoid smooth curve depends only on the Betti number of the graph of its stable reduction. Thus we take for U the pre-image of a Zariski open W of the stable reduction of Z, such that W contains no cycles. The assertion of the observation then holds for $V = U$. But also for any such $V \subset U$ it holds, since this corresponds to a Zariski open part in a blow up of the stable model of Z. Blow ups do not introduce extra cycles.

The rest of the proof of the theorem is similar to the proof of Theorem 5.1.1: just replace $\mathbb{B}$ by U and étale by rigid cohomology. □

REMARK 5.2.3 Both theorems are false when k is not algebraically closed. Just take $X = \mathrm{Sp}(k')$ where $k \subset k'$ is a finite Galois extension and $S = \mathbb{Z}/n\mathbb{Z}_X$. In this case $H^0(X,S) = \mathbb{Z}/n\mathbb{Z}$ and $H^0(X_K, p_K^* S) = (\mathbb{Z}/n\mathbb{Z})^{[k':k]}$. Even if X is a geometrically connected smooth projective curve and S is a constant sheaf the result is false in general. (Both rigid and étale case.)

COROLLARY 5.2.4 *Suppose that S is an overconvergent sheaf of $\mathbb{Z}[1/p]$-modules on $X_{\acute{e}tale}$. The canonical comparison maps $H^q(X,S) \to H^q(X_K, p_K^* S)$ are isomorphisms.*

Proof. There is an exact sequence

$$0 \longrightarrow S_{tors} \longrightarrow S \longrightarrow S \otimes \mathbb{Q} \longrightarrow Q \longrightarrow 0$$

By Theorem 5.1.1 the result is true for both S_{tors} and Q. Since $H^q(X, S \otimes \mathbb{Q})$ agrees with $H^q(X_{rigid}, r_* S \otimes \mathbb{Q})$ (compare Remark 4.2.6) we see the result is true for $S \otimes \mathbb{Q}$ also by the theorem above. The snake lemma gives the result for the sheaf S. □

5.3 QUASI-COMPACT BASE CHANGE.

By a combination of our previous results we can now prove a general base change theorem for quasi-compact morphisms.

THEOREM 5.3.1 *(Quasi-compact base change.) Consider a diagram*

$$\begin{array}{ccc} Z \times_X Y & \xrightarrow{g'} & Y \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ Z & \xrightarrow{g} & X \end{array}$$

and an overconvergent sheaf of $\mathbb{Z}[1/p]$-modules S on $Y_{\acute{e}tale}$. Here f is a quasi-compact morphism of analytic varieties over k and g is a (arbitrary) general morphism of analytic varieties. The base change morphism [SGA 4, Exposé XVII]

$$g^* Rf_* S \longrightarrow Rf'_*(g')^* S$$

is a quasi-isomorphism.

Proof. We only have to show that this morphism induces an isomorphism on the étale stalks of the overconvergent sheaves $g^*R^qf_*S$ and $R^qf'_*(g')^*S$. Any étale point e' of Z lies over an étale point e of X, i.e., such that $F_e \subset F_{e'}$. By the étale base change theorem 3.7.3 the map on the stalks is the map 5.1 between

$$\left(g^*R^qf_*S\right)_{e'} = \left(R^qf_*S\right)_e = H^q(Y_e, S|_{Y_e})$$

and

$$\left(R^qf'_*(g')^*S\right)_{e'} = H^q(Y_{e'}, (g')^*S|_{Y_{e'}}) = H^q(Y_e\hat{\otimes}F_{e'}, p^*_{F_{e'}}(S|_{Y_e})).$$

The statement follows from Corollary 5.2.4. □

6 The axioms for cohomology

Let k be a complete valued field. An 'abstract' cohomology theory for rigid analytic spaces over k is defined in [S-S, section 2] to be a cohomology theory $X \mapsto H^*(X)$ satisfying four axioms. There is also given a candidate for such a cohomology theory. Let A be a finite ring of order prime to the residue field of k. Let K be the completion of the algebraic closure of k. We put

$$H^*(X) := H^*(X\hat{\otimes}K, A_{X\hat{\otimes}K}).$$

As remarked in [S-S, p. 58], the nontrivial axioms to check in this case are the 'homotopy axiom' and the axiom concerning the cohomology of the projective space. In this section we will prove those axioms.

The homotopy axiom states that $H^*(X \times D) \cong H^*(X)$ for an open disc D. This follows immediately from the following theorem.

Theorem 6.0.2 *(The homotopy axiom.) Let X be an analytic space over k. Let S be an overconvergent sheaf of $\mathbb{Z}[1/p]$-modules on $X_{étale}$. Suppose D is an open or closed disc over k; let $p : X \times D \to X$ denote the projection. The canonical maps $H^q(X, S) \to H^q(X \times D, p^*S)$ are isomorphisms.*

Proof. If the disc D is open then it is the admissible union $D = \bigcup B_n$ of closed discs B_n of radius $\rho_n \in \sqrt{|k^*|}$. The covering $X \times D = \bigcup X \times B_n$ is the also admissible. Therefore, it suffices to prove the theorem for a closed disc B.

In this case we prove that $p_*p^*S \cong S$ and that $R^qp_*p^*S = (0)$ for $q \geq 1$. By the étale base change theorem the étale stalk at e of these sheaves are equal to $H^q(B\hat{\otimes}F_e, (S_e)_{B\hat{\otimes}F_e})$. Note that $B\hat{\otimes}F_e \cong \mathbb{B}$, the closed unit disc of radius 1 over F_e. If we prove that $H^q(\mathbb{B}, A_{\mathbb{B}}) = (0)$ for $q \geq 1$ for any $\mathbb{Z}[1/p]$-module A then we are done. A standard argument, compare with 4.2.5, reduces to the cases $A = \mathbb{Q}$ or $A = \mathbb{Z}/n\mathbb{Z}$. These cases where done in Remark 4.2.6 and Proposition 4.2.1. □

For the formulation of the following theorem, we need to be more precise about the Galois action on the cohomology groups. Let K denote the completion of the separable closure k^{sep} of k. The symbol $\mathcal{G} = \mathrm{Gal}(k^{sep}/k)$ denotes the continuous Galois group of K over k. See 3.6. The finite ring A is given the trivial $\mathcal{G}$ action. For any $i \in \mathbb{Z}$ we define $A(i) := A \otimes \left(\mu_n(K)\right)^{\otimes i}$ as a $\mathcal{G}$-module, where $n = \#\mathcal{G}$. The following result also follows from the comparison theorem in the following section.

THEOREM 6.0.3 *(Cohomology of $\mathbb{P}^d$.) Let $\mathbb{P}^d$ denote the d-dimensional projective space over k. We have as Galois modules*

$$H^q(\mathbb{P}^d) = H^q(\mathbb{P}^d_K, A_{\mathbb{P}^d_K}) = \begin{cases} A(-q/2) & \text{for } q \text{ even, } 0 \le q \le 2d, \\ (0) & \text{otherwise.} \end{cases}$$

Proof. The calculation of the cohomology is done by applying the Mayer-Vietoris sequence to the covering $\{U_0, U_1\}$ of $\mathbb{P}^d$ given by $U_0 = \{(z_0; ...; z_d)|\ |z_j| \le |z_0| \text{ for all } j\}$ and $U_1 = \{(z_0; ...; z_d)|\ |z_0| \le max(|z_1|, ..., |z_d|)\}$. The space U_0 is a product of disks and has therefore trivial cohomology. The spaces U_1 and $U_0 \cap U_1$ respectively, admit a surjective morphism to $\mathbb{P}^{d-1}$ given by $(z_0; ...; z_d) \mapsto (z_1; ...; z_d)$. The fibres are disks or ring domains respectively and the fiberings are locally trivial. Base change, with respect to the first map, yields $H^i(U_1) \stackrel{\sim}{\to} H^i(\mathbb{P}^{d-1})$. Base change applied to the second map gives rise to an exact sequence

$$0 \longrightarrow H^i(U_1) \longrightarrow H^i(U_0 \cap U_1) \longrightarrow H^{i-1}(\mathbb{P}^{d-1})(-1) \longrightarrow 0.$$

The (-1) in the last cohomology group is a consequence of the Galois action on the cohomology of a ring domain. See 4.2.2. Induction on d and the Mayer-Vietoris sequence imply that $H^i(\mathbb{P}^d) \cong H^{i-2}(\mathbb{P}^{d-1})(-1)$ for $i \ge 2$ and the expected values of H^0 and H^1. □

7 PURITY AND COMPARISON

Let X be a scheme of finite type over the complete valued field k. We write $X_{ét}$ for the small étale site of the scheme X. Further, X^{an} denotes the rigid analytic variety associated to X. There is a morphism of sites

$$\epsilon : X^{an}_{étale} \longrightarrow X_{ét}$$

comparing the algebraic and rigid étale sites. It is given by the functor that associates to the scheme Y étale over X the analytic space Y^{an} étale over X^{an}. We want to compare sheaves on both sides and their cohomology. It will turn out that if the characteristic of k is zero then we get results as proved in [SGA 4] comparing étale cohomology and classical cohomology over $\mathbb{C}$. However, if the characteristic is $p > 1$, only a weaker version holds. We will give counterexamples for the full statement.

In order to prove the statements above we use a purity result for rigid étale cohomology. It tells us what the cohomology of the complement of a smooth divisor in a smooth rigid analytic variety is.

7.1 A PRELIMINARY RESULT

We start by proving that sheaves of the form $\epsilon^* S$ are overconvergent.

LEMMA 7.1.1 *For any sheaf S on $X_{ét}$ the sheaf $\epsilon^* S$ is overconvergent.*

Proof. Let $Y \to X$ be an algebraic étale morphism, with Y affine. We also denote by Y the sheaf on $X_{ét}$ it defines. We only need to show that the sheaf $\epsilon^*(Y)$ is overconvergent. (The sheaves Y generate the category of sheaves on $X_{ét}$ and the direct limit of overconvergent sheaves is overconvergent.) By Zariski's main theorem we can embed Y as a Zariski open set in a scheme $\overline{Y}$ finite over X. Suppose $U \subset X^{an}$ is an affinoid subdomain and $V \to U$ is an étale morphism of affinoids. The notation $V(r) \to U$ is as in 3.4. We have to show that any morphism $\varphi : V \to Y^{an}$ over X^{an} extends (uniquely) to some $V(r) \to Y^{an}$. See 3.4.4. By Lemma 3.4.2 it suffices to show that $\varphi(V) \subset\subset_U Y^{an} \times_{X^{an}} U$. Clearly, we have that $\varphi(V) \subset\subset_U \overline{Y}^{an} \times_{X^{an}} U$, since the last space is finite over U. The result follows since $Y^{an} \times_{X^{an}} U$ is Zariski open in $\overline{Y}^{an} \times_{X^{an}} U$. See for example [S93, §3 Proposition 3]. □

7.2 Purity for rigid étale cohomology

Let $i : Z \to X$ be a closed immersion of analytic varieties over k. Let $U = X \setminus Z$ denote the admissible open subvariety of X which is the complement of Z. As usual $j : U \to X$ denotes the open immersion of U into X. We want to prove that sheaves on $Z_{étale}$ correspond to sheaves S on $X_{étale}$ such that j^*S is a final object in the category of sheaves on U, i.e., a sheaf which has exactly one section over each object of $U_{étale}$. This means that the category of sheaves on Z can be viewed as the closed sub-topos of $X^{\sim}_{étale}$ complementary to the open sub-topos $U^{\sim}_{étale}$. Compare [SGA 4, Exposé IV 9.3.5]. Although this follows easily for overconvergent sheaves, we need the result in general for the proof of purity below. It implies in particular that $Ri_*\mathcal{F} \cong i_*\mathcal{F}$ for any Abelian sheaf $\mathcal{F}$ on $Z_{étale}$.

Lemma 7.2.1 *The functor i_* identifies the category of sheaves (of sets) on Z with the category of sheaves S on X such that j^*S is a final object of $U^{\sim}_{étale}$.*

Proof. Let us take an admissible affinoid covering $X = \bigcup X_i$ of X and admissible affinoid coverings $X_i \cap X_j = \bigcup X_{ijk}$. Any sheaf on X is given by sheaves on X_i glued on the X_{ijk}, whereas a sheaf on Z (resp. U) is given by sheaves on $Z \cap X_i$ (resp. $U \cap X_i$) glued on the $Z \cap X_{ijk}$ (resp. $U \cap X_{ijk}$). In this way one reduces to the case that X is affinoid.

In this case we work with the sites $X^{sp}_{étale}$ and $Z^{sp}_{étale}$. The functor $X^{sp}_{étale} \to Z^{sp}_{étale}$ is denoted $W \mapsto W_Z = Z \times_X W$. Consider the following statements:

1. For any étale $W_0 \to Z$, W_0 affinoid, there exists an étale $W \to X$ morphism of affinoids such that $W_0 \cong W_Z$.

2. If $V, W \in X^{sp}_{étale}$ and $\phi_0 : W_Z \to V_Z$ is a morphism over Z then there is a Weierstrass domain $W' \subset W$ with $W'_Z = W_Z$ and a morphism $\phi : W' \to V$ lifting ϕ_0.

3. If $W \in X^{sp}_{étale}$ then any special étale covering $\{W_{i,0} \to W_Z\}$ may be lifted to a special étale covering of W.

Let us first prove that these imply the lemma.

We denote by e a final object of $(U_{étale})^{\sim}$. Further for any sheaf S on X we denote by $P(S)$ the presheaf on $Z^{sp}_{étale}$ defined by the formula:

$$P(S)(W_0) = \varinjlim_{V, \phi_0 : W_0 \to V_Z} \Gamma(V, S)$$

By definition, i^*S is the sheaf associated to the presheaf $P(S)$. It is clear from 1) and 2) above that $P(S)$ may be described as follows

$$P(S)(W_Z) = \varinjlim_{W' \subset W \text{ as in } 2)} \Gamma(W', S).$$

Such a subdomain W' is automatically a wide neighborhood of W_Z in W, since W_Z is closed in W. Therefore there exists a special subset $V \subset W$, disjoint with W_Z such that $W = W' \cup V$. This means that if S has the property that $j^*S \cong e$ then $\Gamma(W, S) = \Gamma(W', S)$ since both $\Gamma(V, S)$ and $\Gamma(V \cap W', S)$ consist of one element. In particular we see that for such S we have $P(S)(W_Z) = \Gamma(W, S)$. Property 3) above implies that $P(S)$ is a sheaf in this case and hence $i^*(S) = P(S)$. It follows immediately that $i_* i^* S \cong S$ for such sheaves S.

Conversely, if $\mathcal{F}$ is a sheaf on Z, it is immediate that $j^* i_* \mathcal{F} \cong e$. Hence by the above we have that $i^* i_* \mathcal{F} = P(i_* \mathcal{F})$ and

$$\Gamma(W_Z, i^* i_* \mathcal{F}) = \Gamma(W, i_* \mathcal{F}) = \Gamma(W_Z, \mathcal{F}).$$

We have proved that i_* and i^* are mutually inverse functors defining the desired equivalence of categories.

Let us prove 1). Let $Y_0 \to Z$ be an étale morphism of affinoids. We can choose a presentation (see Lemma 3.4.1)

$$O(Y_0) = O(Z)\langle T_1, \ldots, T_n \rangle / (G_1, \ldots, G_n)$$

with $G_1, \ldots, G_n \in O(Z)[T_1, \ldots, T_n]$ such that the determinant $\Delta_0 = \det\bigl(\partial G_j / \partial T_i\bigr)$ is invertible in $O(Z)$. Let us lift the polynomials G_i to polynomials $F_i \in O(X)[T_1, \ldots, T_n]$. Put $\Delta = \det\bigl(\partial F_j / \partial T_i\bigr)$. Take $\pi \in k^*$ such that $|\pi|$ is smaller than the infimum of $|\Delta_0|$ on Z. We consider the algebra

$$O(X)\langle T_0, T_1, \ldots, T_n \rangle / (F_1, \ldots, F_n, \Delta T_0 - \pi).$$

This defines a special étale morphism $Y \to X$ since the corresponding functional determinant is Δ^2, which is invertible. The isomorphism $Y_0 \cong Z \times_X Y$ follows by construction.

The proof of 2) is similar to the proof of 2) in Lemma 3.7.1. We consider the product $V \times_X W$ and the graph morphism $\Gamma_0 : W_Z \to (V \times_X W)_Z$. This morphism is étale. By 1) (with $W \times_X V$ in stead of X) we can find $\Gamma : Y \to V \times_X W$ such that $Y_Z \cong W_Z$ and $\Gamma_Z = \Gamma_0$. Next argue as in the proof of 3.7.1 to see that there is some wide neighborhood $W' \subset W$ of W_Z such that $\mathrm{pr}_1 \circ \Gamma : \Gamma^{-1}(V \times_X W') \to W'$ is an isomorphism. Thus we get $W' \to V$. Finally, the Weierstrass domains in W are cofinal in the set of neighborhoods of W_Z in W. To see this apply the rigid key lemma to a morphism $f : W \to \mathbb{B}^n$ with $W_Z = f^{-1}(0)$.

For 3) we first note that by 1) we may lift each of the special étale $W_{i,0} \to W_Z$ to special étale $f_i : W_i \to W$. The special subset $\bigcup f_i(W_i)$ is a neighborhood of W_Z in W, hence a wide neighborhood, hence there exists some special $V \subset W$ such that $V \cap W_Z = \emptyset$ and $W = V \cup \bigcup f_i(W_i)$. Write $V = \bigcup V_j$ as a finite union of rational subdomains of W, then the special étale covering of W we are looking for is the covering $\{W_i \to W, V_j \to W\}$. $\square$

Next, we prove some kind of purity in the rigid étale case. Let X be a smooth rigid variety over k. let $i : H \to X$ be a closed immersion, with H smooth over k and everywhere in X of co-dimension 1. Thus it is a smooth divisor in X. Let U denote the admissible open subset $X \setminus H$ of X and let j denote the open immersion $j : U \to X$.

Theorem 7.2.2 *(Purity.) With the notations as above and with n prime to the characteristic of k we have*

$$R^q j_* \mathbb{Z}/n\mathbb{Z}_U = \begin{cases} \mathbb{Z}/n\mathbb{Z}_X & q = 0 \\ i_*(\mu_n^{\otimes -1}) & q = 1 \\ (0) & q \geq 2. \end{cases}$$

Proof. The statement is local on X, hence we may assume X affinoid. Locally on X (in the Zariski topology) the ideal of H is generated by a single function, hence we may assume that H is given as $f = 0$ for some $f \in O(X)$. By [K68] we can find an affinoid neighborhood of H in X which has an admissible covering by affinoids of the form $H_i \times \mathbb{B}$. Here $\mathbb{B}$ is the closed unit ball over k with coordinate z. Thus we may assume that $X = H \times \mathbb{B}$ and $U = H \times \mathbb{B}^*$ where $\mathbb{B}^*$ is the punctured unit disc. Let us write $\bar{f} : X \to H$ for the projection and $f = \bar{f}|_U$ so that we have the following commutative diagram:

$$\begin{array}{ccccc} U & \xrightarrow{j} & X & \xleftarrow{i} & H \\ & f \searrow & \downarrow \bar{f} & \swarrow \mathrm{id} & \\ & & H & & \end{array}$$

We note that the sheaves $R^q j_* \mathbb{Z}/n\mathbb{Z}_U$ for $q \geq 1$ have are zero restricted to U, hence are of the form $i_* F_q$ for certain sheaves F_q on H (use lemma above). Further, it is clear that $j_* \mathbb{Z}/n\mathbb{Z}_U = \mathbb{Z}/n\mathbb{Z}_X$ on X. We study the spectral sequence associated to the isomorphism $Rf_* \cong R\bar{f}_* \circ Rj_*$. For the sheaf $\mathbb{Z}/n\mathbb{Z}_U$ its E_2-terms are $E_2^{ab} = R^a \bar{f}_* R^b j_* \mathbb{Z}/n\mathbb{Z}_U$ and it abuts to $R^{a+b} f_* \mathbb{Z}/n\mathbb{Z}_U$. In view of the fact that $R^b j_* \mathbb{Z}/n\mathbb{Z}_U = i_* F_b \cong Ri_* F_b$ for $b \geq 1$ (by 7.2.1), we see that $E_2^{ab} = 0$ for $a, b \geq 1$ and $E_2^{0b} = F_b$ for $b \geq 1$. Also we have $E_2^{a0} = R^a \bar{f}_* \mathbb{Z}/n\mathbb{Z}_X$. This is an overconvergent sheaf, whose étale stalks are $H^a(\mathbb{B}, \mathbb{Z}/n\mathbb{Z}_{\mathbb{B}})$, over various algebraically closed base fields. Hence, by Lemma 4.1.2 and since $Pic(\mathbb{B}) = (0)$, we see that $R^a \bar{f}_* \mathbb{Z}/n\mathbb{Z}_X = (0)$ for $a \geq 2$. The upshot of all of this is: 1) we have $R^0 f_* \mathbb{Z}/n\mathbb{Z}_U = \mathbb{Z}/n\mathbb{Z}_H$, 2) there is an exact sequence

$$0 \longrightarrow R^1 \bar{f}_* \mathbb{Z}/n\mathbb{Z}_X \longrightarrow R^1 f_* \mathbb{Z}/n\mathbb{Z}_U \longrightarrow F_1 \longrightarrow 0$$

and 3) there are isomorphisms $R^q f_* \mathbb{Z}/n\mathbb{Z}_U \cong F_q$ for $q \geq 2$.

We have already used the morphism

$$R\bar{f}_* \mathbb{Z}/n\mathbb{Z}_X \longrightarrow Rf_* \mathbb{Z}/n\mathbb{Z}_U.$$

In addition, there is a map

$$\mathbb{Z}/n\mathbb{Z}_H[-1] \longrightarrow Rf_* \mu_n,$$

which associates to $1 \in \mathbb{Z}/n\mathbb{Z}$ the section of $R^1 f_* \mu_n$ corresponding to the μ_n-torsor of $U = H \times \mathbb{B}^*$ given by the equation $y^n = z$. We claim that together these induce a quasi-isomorphism

$$R\bar{f}_* \mathbb{Z}/n\mathbb{Z}_X \oplus \mu_n^{\otimes -1}[-1] \longrightarrow Rf_* \mathbb{Z}/n\mathbb{Z}_U. \tag{1}$$

From the considerations above it follows that this implies the theorem.

To prove the claim we may assume that n is a prime power $n = p^r$. Let us treat the case that p equals the characteristic of k. In this case k is a p-adic field. The other cases are easier and similar arguments work.

Let us take $c \in \mathbb{N}$ large enough. For $m \in \mathbb{N}$ we put $R_m = \{x \in \mathbb{B}; |x| \geq |p^{cm}|\}$, a ring domain. Put $U_m = H \times R_m$, note that the covering $U = \bigcup U_m$ is admissible. Let us write $f_m : U_m \to H$ for the projection. We will study the overconvergent sheaf $R^q f_{m,*}\mu_n$. Its étale fibres are $H^q(R_m \hat{\otimes} F_e, \mu_n)$. These are zero for $q \geq 2$ and equal to μ_n for $q = 0$. Note that

$$\mathbb{P}^1 = \{|z| \geq |p^{cm}|\} \cup \{|z| \leq 1\},$$

hence that we have an exact sequence

$$0 \longrightarrow F_e^* \longrightarrow O^*(|z| \geq |p^{cm}|) \oplus O^*(|z| \leq 1) \longrightarrow O^*(R_m) \longrightarrow \mathbb{Z} \longrightarrow 0.$$

This follows by the computation of cohomology of $\mathbb{P}^1$, use for example Lemma 4.1.2. We see immediately that

$$H^1(R_m \hat{\otimes} F_e) = \mathbb{Z}/n\mathbb{Z} \oplus H^1(|z| \leq 1) \oplus H^1(|z| \geq |p^{cm}|).$$

This implies a corresponding decomposition of the overconvergent sheaf

$$R^1 f_{m,*}\mu_n = \mathbb{Z}/n\mathbb{Z}_H \oplus R^1 \bar{f}_*\mu_n \oplus Rest_m.$$

The maps $Rest_{m+1} \to Rest_m$ are zero, since by [L93, Theorem 2.1] the maps on the étale fibres $H^1(|z| \geq |p^{c(m+1)}|) \to H^1(|z| \geq |p^{cm}|)$ are zero if c is large enough.

This means that for any $V \to H$ affinoid étale we have a decomposition

$$H^q(V \times R_m) = H^{q-1}(V, \mu_n^{\otimes -1}) \oplus H^q(V \times \mathbb{B}) \oplus \mathrm{Rest}_m,$$

the transition maps $H^q(V \times R_{m+1}) \to H^q(V \times R_m)$ are the identity on the first two summands, zero on the last one. This proves that

$$\varprojlim H^q(V \times R_m) = H^{q-1}(V, \mu_n^{\otimes -1}) \oplus H^q(V \times \mathbb{B})$$

and the derived limit

$$\varprojlim{}^{(1)} H^q(V \times R_m) = (0).$$

We conclude that $H^q(V \times \mathbb{B}^*) = H^{q-1}(V, \mu_n^{\otimes -1}) \oplus H^q(V \times \mathbb{B})$, hence (1) is an isomorphism. □

7.3 Comparison

In this section X will denote a variety of finite type over over the complete valued field k. We state the results corresponding to [SGA 4, Exposé XI Theorem 4.4] in our case. Further, we will indicate the necessary changes in the proof given there so that it will work in our case also.

Theorem 7.3.1 *Suppose the characteristic of k is zero. There is an equivalence between the category of locally constant sheaves on $X_{ét}$ with finite stalks and the category of locally constant sheaves on $X^{an}_{étale}$ with finite stalks. The equivalence is given by the functors ϵ^* and ϵ_*.*

Proof. Since sheaves of this kind are representable by finite étale coverings we see that it suffices to prove the following statement: If $Y \to X^{an}$ is finite étale then there exists a (unique) finite étale morphism of schemes $Z \to X$ such that $Y \cong Z^{an}$. This was recently proved by Lütkebohmert, see [L93]. □

In the next theorem k is no longer of characteristic zero.

THEOREM 7.3.2 *Let X be smooth over* Spec(k) *and let k be algebraically closed. Suppose S is an Abelian locally constant sheaf on $X_{\acute{e}t}$ with finite stalks where all orders of torsion are prime to the characteristic of k. In this case we have $R^q\epsilon_*\epsilon^*S = (0)$ for $q \geq 1$. The canonical morphisms $H^q(X_{\acute{e}t}, S) \to H^q(X^{an}_{\acute{e}tale}, \epsilon^*S)$ are isomorphisms. In particular we have*

$$H^q(X_{\acute{e}t}, \mathbb{Z}/n\mathbb{Z}) \cong H^q(X^{an}, \mathbb{Z}/n\mathbb{Z}_{X^{an}}).$$

Proof. With the results proved above, we can use the proof of [SGA 4, Exposé XI Theorem 4.4 part (ii)]. In stead of the 'calcul direct' of line 1 on page 13 we use Theorem 7.2.2. The only other fact used in the proof which is not immediately clear is the following: Suppose $\bar{f} : \bar{X} \to S$ is a family of smooth projective curves over the scheme S, which is of finite type over k, suppose n is relatively prime to the characteristic of k. In this case $R^1\bar{f}^{an}_*\mathbb{Z}/n\mathbb{Z}_{\bar{X}^{an}}$ is a locally constant sheaf on $S^{an}_{\acute{e}tale}$. However, this is immediately clear from: 1) The corresponding fact in the algebraic case. 2) The base change map $\epsilon^*R^1f_*\mathbb{Z}/n\mathbb{Z} \to R^1\bar{f}^{an}_*\mathbb{Z}/n\mathbb{Z}_{\bar{X}^{an}}$ is an isomorphism (look at étale fibres). □

REMARK 7.3.3 The more general results proved in [SGA 4, Exposé XVI §4] should hold true for the rigid analytic case also. At least if the characteristic of k is zero then it should be possible with some effort to follow the reasoning of locus citatus in this case.

7.4 Counterexamples in characteristic $p > 0$.

Take k algebraically closed of characteristic $p > 0$. Riemann's existence theorem is no longer valid in this case. We give an example of this.

LEMMA 7.4.1 *Consider the covering $\psi : Y \to \mathbb{A}^{1\,an}$, given by the equation $T^p - T = F := \sum_{i\geq 0} a_i z^{p^i}$, where the series F converges on $\mathbb{A}^{1\,an}$. We suppose that there are infinitely many non zero a_i and that for every $k \geq 0$ one has*

$$|a_k + a^p_{k-1} + a^{p^2}_{k-2} + ... + a^{p^k}_0| = \ max\ _{0\leq i\leq k}(|a_i|^{p^{k-i}}).$$

Then Y is not isomorphic to Z^{an} for any covering $Z \to \mathbb{A}^1$.

Proof. Any p-cyclic (un-ramified) covering of the unit disk D is given by an equation $T^p - T = f$ with $f \in O(D)$. Two functions $f_1, f_2 \in O(D)$ define isomorphic coverings if and only if $\lambda_1 f_1 + \lambda_2 f_2 = h^p - h$ holds with $\lambda_1, \lambda_2 \in \mathbb{F}^*_p$ and $h \in O(D)$. Using that the structure sheaf O on the analytic space $\mathbb{A}^{1,an}$ has trivial cohomology, one finds for $\mathbb{A}^{1,an}$ similar results. Namely: Any p-cyclic analytic covering of $\mathbb{A}^{1\,an}$ is given by an equation $T^p - T = f$ with f a holomorphic function on $\mathbb{A}^{1\,an}$. Two holomorphic functions f_1, f_2 define the same p-cyclic extension if and only if there is a $\lambda \in \mathbb{F}_p$ such that the equation $T^p - T = \lambda_1 f_1 + \lambda_2 f_2$ has a holomorphic solution.

If $Y = Z^{an}$ then Z is a p-cyclic covering of $\mathbb{A}^1$ given by an equation of the form $U^p - U = g$ with $g \in zk[z]$. Let the equation $T^p - T = -G := F - \lambda g$ have a holomorphic solution $h = \sum_{i\geq 1} h_i z^i$. In some disk around 0 the spectral norm of G is less than 1. Therefore $\sum_{i\geq 0} G^{p^i}$ converges and is on this disk a solution of $T^p - T = -G$. So h coincides with $\sum_{i\geq 0} G^{p^i}$ on this disk and the power series expansion of h is equal to the power series expansion of $\sum_{i\geq 0} G^{p^i}$. One takes a disc $D(0,R)$ around 0 such that $1 < B := \|G\|_R = \|F\|_R > \|g\|_R$. After replacing z by $z\lambda$ for a suitable $\lambda \in k^*$, we may suppose that $R = 1$. First we look at $\sum_{i\geq 0} F^{p^i} = \sum_{k\geq 0} A_k z^{p^k}$ with $A_k = (a_k + a_{k-1}^p + a_{k-2}^{p^2} + ... + a_0^{p^k})$. A calculation shows that for $N >> 0$ one has $|A_N| = |A_{N-1}|^p$ and $|A_N| \geq B$. Then we look at $\sum_{i\geq 0} g^{p^i} = \sum_{k\geq 1} b_k z^k$. One can calculate that the absolute value of b_{p^k} grows less fast than $|A_k|$. This implies that the power series representing h is not convergent on $D(0,1)$. This contradiction ends the proof. □

Corollary 7.4.2 *The map*

$$H^1(\mathbb{A}^1_{et}, \mathbb{Z}/p\mathbb{Z}) \longrightarrow H^1(\mathbb{A}^{1\,an}_{etale}, \mathbb{Z}/p\mathbb{Z})$$

is injective but not surjective. In particular Theorem 7.3.2 is not valid for sheaves consisting of p-torsion.

Corollary 7.4.3 *Let $n > 1$ with $p|\phi(n)$ and with n not divisible by p. There is a locally constant sheaf S on $\mathbb{A}^{1\,an}_{étale}$ with stalk $\mathbb{Z}/n\mathbb{Z}$, which is not of the form $\epsilon_* T$.*

Proof. We consider the p-cyclic analytic covering $\psi : Y \to \mathbb{A}^{1\,an}$ of Lemma 7.4.1. Let σ denote the generator of the Galois group G of this extension. Let M denote the constant étale sheaf on Y with stalk $\mathbb{Z}/n\mathbb{Z}$. Let $a \in \mathbb{Z}/n\mathbb{Z}^*$ be an element of order p. We define an action G on $\mathbb{Z}/n\mathbb{Z}$ by $\sigma(i) = ai$. This induces an action of G on $\mathbb{Z}/n\mathbb{Z} \times Y$ by $\sigma((i,y)) = (ai, \sigma(y))$. The quotient by this group action is a sheaf S on $\mathbb{A}^{1\,an}_{étale}$ which is locally the constant sheaf with stalk $\mathbb{Z}/n\mathbb{Z}$. (And of course $\psi^* S$ is the constant sheaf on Y). However, there is no étale covering $\{Y_i\}$ of $\mathbb{A}^1$ which trivializes S. Indeed, such an étale covering would give a trivialization of $Y \to \mathbb{A}^{1\,an}$. □

References

[SGA 4] *M. Artin, A. Grothendieck* and *J.-L. Verdier*, Théorie des topos et cohomologie étale des schemas, Lecture Notes in Mathematics, 269, 270, 305, Springer Verlag, Berlin-Heidelberg-New York, 1972-1973.

[B90] *V.G. Berkovich*, Spectral theory and analytic geometry over non-archimedean fields, Mathematical surveys and monographs, vol. 33, American Mathematical Society, Providence 1990.

[B93] *V.G. Berkovich*, Étale cohomology for non-Archimedean analytic spaces, Publ. Math. I.H.E.S. 78, 5-161 (1993).

[B94a] *V.G. Berkovich*, Vanishing cycles for formal schemes, Invent. math. 115, 539-571 (1994).

[B94b] *V.G. Berkovich*, Vanishing cycles for non-Archimedean analytic spaces, Preprint.

[B94c] *V.G. Berkovich*, On the comparison theorem for étale cohomology of non-Archimedean analytic spaces, Preprint.

[BGR] *S. Bosch, U. Günther* and *R. Remmert*, Non-Archimedean analysis, Springer Verlag, Berlin-Heidelberg-New York 1984.

[BL] *S. Bosch* and *W. Lütkebohmert*, Formal and rigid geometry, Math. Ann. 295 (1993), 291-317.

[D92] *P. Deligne*, letter to Raynaud of 23/8/1992.

[FM] *J. Fresnel* and *M. Matignon*, Annali di Matematica pura ed applicata (4) 145 (1986), 159-210.

[FP] *J. Fresnel* and *M. van der Put*, Géométrie analytique rigide et applications, Progress in Mathematics, vol. 18, Birkhäuser, Boston 1981.

[GG] *L. Gerritzen* and *H. Grauert*, Die Azyklizität der affinoiden Überdeckungen, in: Global analysis, Papers in honor of K. Kodaira, 159-184, Princeton University Press, Princeton 1969.

[GP] *L. Gerritzen* and *M. van der Put*, Schottky groups and Mumford curves, Lecture Notes in Mathematics 817, Springer Verlag, Berlin-Heidelberg-New York 1980.

[H77] *R. Hartshorne*, Algebraic geometry, Springer Verlag, Berlin-Heidelberg-New York 1977.

[Hu] *R. Huber*, Étale cohomology of rigid analytic varieties and adic spaces, preprint, July 1994.

[K68] *R. Kiehl*, Die de Rham Kohomologie algebraischer Mannigfaltigkeiten über einem bewerteten Körper, Publ. Math. I.H.E.S., 33 (1968), 5-20.

[LP] *Q. Liu* and *M. van der Put*, On one dimensional separated rigid spaces. To appear.

[L93] *W. Lütkebohmert*, Riemann's existence problem for a p-adic field, Invent. math. 111 (1993), 309-330.

[M81] *F. Mehlmann*, Ein Beweis für einen Satz von Raynaud über flache Homomorphismen affinoider Algebren, Schriftenreihe des mathematischen Instituts der Universität Münster, 2. Serie, Heft 19 (1981).

[M80] *J.S. Milne*, Étale Cohomology, Princeton University Press, Princeton 1980.

[P80] *M. van der Put*, The class group of a one-dimensional affinoid space, Ann. Ist. Fourier 30 (1980), 155-164.

[P82] *M. van der Put*, Cohomology on affinoid spaces, Comp. Math. 45 (1982), 165-198.

[P92] *M. van der Put*, Serre duality for rigid analytic spaces, Indag. Mathem., N.S. 3(2) (1992), 219-235.

[R70] *M. Raynaud*, Anneaux locaux Henséliens, Lecture Notes in Mathematics 169, Springer Verlag, Berlin-Heidelberg-New York 1970.

[R74] *M. Raynaud*, Géométrie analytique rigide d'apres Tate, Kiehl,.. , Bull. Soc. Math. France, Mémoire 39-40 (1974), 319-327.

[S91] *P. Schneider*, The étale topology on rigid analytic spaces, informal notes, Kyoto 1991.

[S93] *P. Schneider*, Points of rigid analytic varieties, J. reine angew. Math. 434 (1993), 127-157.

[S-S] *P. Schneider* and *U. Stuhler*, The cohomology of p-adic symmetric spaces, Invent. math. 105 (1991), 47-122.

[S64] *J.P. Serre*, Cohomologie Galoisienne, Lecture Notes in Mathematics 5, Springer Verlag, Berlin-Heidelberg-New York, 1964.

Dr. A. J. de Jong
Department of Mathematics
Harvard University
One Oxford Street
Cambridge
Massachusetts O2138, USA
dejong@math.harvard.edu

Prof. M. van der Put
Department of Mathematics
University of Groningen
P.O.B. 800
9700 AV Groningen
The Netherlands
M.van.der.Put@math.rug.nl

Order of Torsion in CH^4 of Quadrics

Nikita A. Karpenko[1]

Received: November 24, 1995
Revised: January 30, 1996

Communicated by Alexander S. Merkurjev

Abstract. It is shown that the order of the torsion subgroup in the 4-codimensional Chow group $\mathrm{CH}^4(X_\varphi)$ of a projective quadric X_φ is at most 4 provided that the dimension of the corresponding quadratic form φ is greater than 8.

Consider a non-degenerate quadratic form φ over a field F of characteristic different from 2 and the corresponding projective quadric X_φ. We always assume that $\dim X_\varphi \geq 1$, i.e. $\dim \varphi \geq 3$. It is an open problem to describe the torsion subgroup of the Chow group $\mathrm{CH}^*(X_\varphi)$ (this is the group of algebraic cycles on X_φ modulo rational equivalence graded by co-dimension of cycles [1, 2]).

Generally speaking computation of the Chow group of an algebraic variety is an interesting and important problem of algebraic geometry. However the class of varieties for which this problem is solved is rather small. Chow groups and K-theory of quadrics were studied first by R. Swan. Although the K-theory was completely computed [13] the question on the Chow group remained open.

A new motivation grew out of the attempts to solve the norm residue homomorphism problem. During the work on this problem it became clear that a decisive progress could be achieved by computation of the so called K-cohomology groups [10, 12] for quadrics and in particular of their Chow groups.

In [4] Chow groups of small-dimensional quadrics were computed. An interesting phenomenon was found: some Chow groups have torsion and the problem of computing the whole Chow group reduces to finding the torsion.

Let us consider some first gradation components. The group $\mathrm{CH}^1(X_\varphi)$ is always torsion-free. The next group — $\mathrm{CH}^2(X_\varphi)$ is computed in [4]. In particular, it turns out that $\#\,\mathrm{Tors}\,\mathrm{CH}^2(X_\varphi) \leq 2$ for any form φ; moreover, $\mathrm{Tors}\,\mathrm{CH}^2(X_\varphi) = 0$ if $\dim\varphi > 8$ [4, theorem (6.1)]. In co-dimension 3 one has: $\#\,\mathrm{Tors}\,\mathrm{CH}^3(X_\varphi) \leq 2$ for any φ [5, theorem] and $\mathrm{Tors}\,\mathrm{CH}^3(X_\varphi) = 0$ if $\dim\varphi > 12$ [6, theorem 6.1]. As to co-dimension 4, it is known today that $\mathrm{Tors}\,\mathrm{CH}^4(X_\varphi) = 0$ if $\dim\varphi > 24$ [6, theorem 8.5]; however, one has an example of a 7-dimensional form φ (defined over an appropriate F) with infinite $\mathrm{Tors}\,\mathrm{CH}^4(X_\varphi)$ [7, theorem 6.5].

[1] Support and hospitality of Sonderforschungsbereich 343 (Bielefeld University) are gratefully acknowledged.

Here we prove that
$$\#\operatorname{Tors}\operatorname{CH}^4(X_\varphi) \le 4$$
for any φ of dimension greater than 8 (4.1). Notice that $\operatorname{Tors}\operatorname{CH}^4(X_\varphi) = 0$ if $\dim\varphi = 6$ [14], [4, (2.6)] and $\operatorname{CH}^4(X_\varphi) = 0$ if $\dim\varphi < 6$; so, the "exceptional" dimensions are only 7 and 8. We also reproduce (with small simplifications) the proof that $\#\operatorname{Tors}\operatorname{CH}^3(X_\varphi) \le 2$.

This note grew out from a remark of B. Kahn that $\operatorname{Tors}\operatorname{CH}^4(X_\varphi)$ is finite if $\dim\varphi > 8$.

1. An exact sequence

We consider (Quillen's) K-cohomology $H^p(X_\varphi, K_q)$ and the Grothendieck group $K_0'(X_\varphi)$ which we denote simply by $K(X_\varphi)$ and supply with the so called topological filtration
$$K(X_\varphi) = K(X_\varphi)^{(0)} \supset K(X_\varphi)^{(1)} \supset \dots .$$
We denote by $\tilde\varphi$ the form φ over a field extension $\tilde F$ of F which completely (so much as possible by the dimension reason) splits φ.

Proposition 1.1. *One has an exact sequence*
$$\operatorname{Ker}\big(H^2(X_\varphi, K_3) \to H^2(X_{\tilde\varphi}, K_3)\big) \to \operatorname{CH}^4(X_\varphi) \to K(X_\varphi)^{(4/5)} \to 0 .$$

Proof. The kernel of the canonical epimorphism $\operatorname{CH}^4(X_\varphi) \twoheadrightarrow K(X_\varphi)^{(4/5)}$ is controlled by certain differentials of the BGQ-spectral sequence $E_2^{p,q} = H^p(X_\varphi, K_{-q})$ [10, §7]. Since $\operatorname{CH}^4(X_\varphi) = E_2^{4,-4}$, the differentials in question start from $E_4^{0,-1}$, $E_3^{1,-2}$ and $E_2^{2,-3}$. Since
$$E_2^{0,-1} = H^0(X_\varphi, K_1) = F^\times \ \text{ and}$$
$$E_2^{1,-2} = H^1(X_\varphi, K_2) = F^\times \ \text{ (if } \dim\varphi > 4) \text{ [4, theorem (4.1)]}$$
all the differentials starting from $E_r^{0,-1}$ and $E_r^{1,-2}$ with $r \ge 2$ are 0. Hence we have an exact sequence
$$H^2(X_\varphi, K_3) \xrightarrow{d} \operatorname{CH}^4(X_\varphi) \to K(X_\varphi)^{(4/5)} \to 0 .$$

Using pull-back with respect to the embedding of X_φ in the enveloping projective space $\mathbb{P}$, one can define a homomorphism
$$F^\times = H^2(\mathbb{P}, K_3) \to H^2(X_\varphi, K_3)$$
which is easily checked to be an isomorphism in the case when φ splits and $\dim\varphi > 6$. For an arbitrary φ we obtain a commutative square
$$\begin{array}{ccc} F^\times & \longrightarrow & H^2(X_\varphi, K_3) \\ \big\downarrow & & \big\downarrow \\ \tilde F^\times & \longrightarrow & H^2(X_{\tilde\varphi}, K_3) \end{array}$$
which produce a decomposition
$$H^2(X_\varphi, K_3) = F^\times \oplus \operatorname{Ker}\big(H^2(X_\varphi, K_3) \to H^2(X_{\tilde\varphi}, K_3)\big)$$
provided that $\dim\varphi > 6$. Since $d|_{F^\times} = 0$, we are done in this case. The case $\dim\varphi \le 6$ is trivial and not of use for the consequent. $\square$

2. THE LEFT-HAND SIDE TERM

There is a description of the left-hand side term of (1.1).

PROPOSITION 2.1 ([8, prop. 1]). *Suppose that* $\dim\varphi \geq 5$ *and* φ *is not a 3-Pfister neighbor (i.e. not similar to a subform of an anisotropic 3-Pfister form). The kernel of the restriction*

$$H^2(X_\varphi, K_3) \longrightarrow H^2(X_{\tilde{\varphi}}, K_3)$$

is naturally isomorphic to the kernel of the Galois cohomology map

$$H^4(F, \mathbf{Z}/2) \longrightarrow H^4\big(F(\varphi), \mathbf{Z}/2\big)\,.$$

REMARK 2.2. The assumption that φ is not a 3-Pfister neighbor is likely superfluous.

DEFINITION 2.3. Denote by $P_4(\varphi)$ the subset of $H^4(F, \mathbb{Z}/2)$ *consisting* of 0 and all cup-products (a, b, c, d) with $a, b, c, d \in F^\times$ such that φ is similar to a subform of the 4-Pfister form $\langle\langle a, b, c, d\rangle\rangle$ (the latter means as usual the product $\langle 1, -a\rangle \otimes \langle 1, -b\rangle \otimes \langle 1, -c\rangle \otimes \langle 1, -d\rangle$).

PROPOSITION 2.4 ([3]). *If* φ *is any quadratic form with* $\dim\varphi \geq 5$ *then*

$$\operatorname{Ker}\Big(H^4(F, \mathbf{Z}/2) \to H^4\big(F(\varphi), \mathbf{Z}/2\big)\Big) = P_4(\varphi)\,.$$

COROLLARY 2.5. *If* $\dim\varphi > 8$ *one can rewrite the sequence (1.1) as follows:*

$$P_4(\varphi) \to \mathrm{CH}^4(X_\varphi) \to K(X_\varphi)^{(4/5)} \to 0\,.$$

□

3. THE RIGHT-HAND SIDE TERM

In order to control the right-hand side term of (2.5), we need some general facts on the subsequent quotients of the topological filtration on the Grothendieck group of a quadric. Most results of this § are from [5].

We are going to use the following notation.

We put for shortness $K = K(X_\varphi)$.

The quotient $K^{(p/p+1)}$ will be denoted by $\mathrm{G}^p K$.

We put forever $d = \dim X_\varphi = \dim\varphi - 2$.

Sometimes it is more convenient to use the lower indexes for the topological filtration by meaning dimension instead of co-dimension, i.e. $K_{(p)} = K^{(d-p)}$. All the graded groups appearing in this § are graded "by co-dimension"; by that reason the asterisk stays always as a superscript. However, sometimes it is more convenient to refer to a component of a graded group by giving its "dimension"; in this case we use the subscript. For instance, $\mathrm{G}_p K$ will stay for the p-dimensional component of the graded group $\mathrm{G}^* K$; it is the same as $\mathrm{G}^{d-p} K$.

Let $h \in K$ be the class of a hyperplane section of X_φ. This h does not depend on the choice of the hyperplane, moreover $h = 1 - [\mathcal{O}_{X_\varphi}(-1)]$.

For any $x \in K$ we define dimension $\dim x$ of x as the infimum of p such that $x \in K_{(p)}$. For instance, $\dim 0 = -\infty$, $\dim h = d - 1$. Any $0 \neq x \in K$ determines an element $0 \neq \bar{x} \in \mathrm{G}^* K$, namely the residue class in $\mathrm{G}_{\dim x} K$.

The subring of K generated by h will be denoted by H. It contains $[\mathcal{O}(n)]$ for all integers n. As a group, H is freely generated by $1,\ h,\ h^2,\ \ldots,\ h^d$. The filtration on

H induced from K is just the "filtration by powers of h". In particular, the adjoint graded group G^*H is torsion-free.

DEFINITION 3.1. Let us define an integer $s = s(\varphi)$ in the following way. If $\varphi \notin I^2(F)$ (where $I(F)$ stays for the ideal of the even-dimensional forms in the Witt ring of F) then the even Clifford algebra $C_0(\varphi)$ is simple, so it is isomorphic to the algebra $M_n(D)$ of $(n\times n)$-matrices over a skew-field D; in this case we take s such that $n = 2^s$. If $\varphi \in I^2(F)$, we take s such that $C_0(\varphi) \simeq M_{2^s}(D) \times M_{2^s}(D)$.

There is a trivial observation

LEMMA 3.2. *If $\varphi \notin I^2(F)$ then $K\big(C_0(\varphi)\big)$ is freely generated by the class of a (unique up to an isomorphism) simple $C_0(\varphi)$-module P; moreover,*

$$[C_0(\varphi)] = 2^{s(\varphi)} \cdot [P] \in K\big(C_0(\varphi)\big) \ .$$

If $\varphi \in I^2(F)$ then $K\big(C_0(\varphi)\big)$ is freely generated by the classes of two non-isomorphic simple $C_0(\varphi)$-modules P and P'; moreover,

$$[C_0(\varphi)] = 2^{s(\varphi)} \cdot \big([P] + [P']\big) \in K\big(C_0(\varphi)\big) \ .$$

□

LEMMA 3.3 ([4, lemma (3.6)]). *Let $\mathcal{U}$ be the Swan's sheaf on X_φ* [13, p. 126]. *Then in K*

$$[\mathcal{U}(d)] = h^d + 2h^{d-1} + \cdots + 2^{d-1}h + 2^d \ .$$

Since the sheaf $\mathcal{U}$ has a (right) action of $C_0(\varphi)$ the class $[\mathcal{U}] \in K$ is divisible by 2^s (3.2), so the following definition is correct (take also in account that the group K is torsion-free by [13, theorem 1] and (3.2)).

DEFINITION 3.4. For any $0 \le i < s$ we define an element $l_i \in K$ as

$$l_i = \frac{1}{2^{i+1}} \left(h^d + 2h^{d-1} + \cdots + 2^i h^{d-i}\right) \ ;$$

for a certain convenience reason we also put $l_{-1} = 0$.

What these elements are explains the following

LEMMA 3.5. *The element l_i is equal to the class of an i-dimensional linear subspace on X_φ if such a subspace lies on X_φ (i.e. if the form φ contains an $(i+1)$-dimensional totally isotropic subspace, i.e. if the Witt index of φ is at least $i+1$).*

Proof. Let $L_i \subset X_\varphi$ be an i-dimensional linear subspace of X_φ and $in : X_\varphi \hookrightarrow \mathbb{P}$ the embedding of X_φ into the projective space as a hypersurface. First assume that $\dim\varphi$ is odd. Then using [13, theorem 1] it is easy to see that the push-forward $in_* : K(X_\varphi) \to K(\mathbb{P})$ is injective, so it would be enough to check that $in_*([L_i]) = in_*(l_i)$. The left-hand side is just $[L_i] \in K(\mathbb{P})$ while the right-hand side can be rewritten with using the projection formula as

$$\frac{1}{2^{i+1}} \left(l^d + 2l^{d-1} + \cdots + 2^i l^{d-i}\right) \cdot [X_\varphi]$$

where l^i denotes the class of an i-co-dimensional linear subspace of $\mathbb{P}$. Computing

$$[X_\varphi] = 1 - [\mathcal{O}_{\mathbb{P}}(-2)] = 2l^1 - l^2$$

and multiplying we get l^{d-i+1} what is the same as the required $[L_i]$ because $\dim \mathbb{P} = d+1$.

Now assume that $\dim\varphi$ is even. Take any non-singular hyperplane section Y of X_φ containing L_i (it is really possible to find such a Y because $i \neq d/2$ (3.4)). Since Y is an odd-dimensional quadric we know from the previous paragraph that

$$[L_i] = \frac{1}{2^{i+1}}(h^{d-1} + 2h^{d-2} + \cdots + 2^i h^{d-1-i}) \in K(Y) .$$

Applying the push-forward with respect to the embedding $Y \hookrightarrow X_\varphi$ and using once again the projection formula for the right-hand side we get

$$[L_i] = \frac{1}{2^{i+1}}(h^{d-1} + 2h^{d-2} + \cdots + 2^i h^{d-1-i}) \cdot [Y] \in K(X_\varphi) .$$

Since $[Y] = h$ we are done. □

LEMMA 3.6. *For any* $0 \le i < s$ *one has:*

- $2l_i = h^{d-i} + l_{i-1}$;
- $hl_i = l_{i-1}$;
- $\dim l_i > \dim l_{i-1}$;
- *if* φ *is anisotropic then* $\dim l_i > i$.

Proof. The first two properties are obvious from the formula (3.4) defining l_i. Since the multiplication in K respects the filtration and $h \in K^{(1)}$ the second property implies the third one. If φ is anisotropic, the degree of any closed point on X_φ is even whence $l_0 \notin K_{(0)}$, i.e. $\dim l_0 > 0$; thus $\dim l_i \ge i + \dim l_0 > i$. □

COROLLARY 3.7. *If* φ *is anisotropic every element* $\bar{l_i} \in \mathrm{G}^*K$, $0 \le i < s$ *has order 2.*

Proof. By an agreement in the beginning of § we denote by $\bar{l_i}$ the class of $l_i \in K$ in $\mathrm{G}_{\dim l_i}K$. By (3.6) $2l_i = h^{d-i} + l_{i-1}$, $\dim l_i > \dim l_{i-1}$ and $\dim l_i > i = \dim h^{d-i}$. Thus $\dim l_i > \dim 2l_i$, i.e. $2\bar{l_i} = 0$. □

DEFINITION 3.8. Let us denote by $\mathcal{I}^* \subset \mathrm{Tors}\,\mathrm{G}^*K$ the subgroup generated by all $\bar{l_i}$, $0 \le i < s$. The quotient $\mathrm{Tors}\,\mathrm{G}^*K/\mathcal{I}^*$ will be denoted by $\mathcal{II}^*$.

THEOREM 3.9. *Assume that the quadratic form* φ *is anisotropic. There exits an exact sequence of graded groups*

$$0 \to \mathcal{I}^* \to \mathrm{Tors}\,\mathrm{G}^*K \to \mathcal{II}^* \to 0$$

where $\mathcal{I}^*$ *and* $\mathcal{II}^*$ *have the following properties:*

- $\#\mathcal{I}^p \le 2$ *for any* p;
- $\#\mathcal{I}^* = 2^s$ *where* $s = s(\varphi)$ *is defined in (3.1);*
- *if* $\varphi \notin I^2(F)$ *then* $\mathcal{II}^* = 0$;

moreover, in the case $\varphi \in I^2(F)$ *it holds:*

- *for every* p *the group* $\mathcal{II}^p$ *is cyclic;*
- $\mathcal{II}^p = 0$ *for* $p \ge d/2$;
- *if there exists a field extension of degree* 2^n *which completely splits* φ *then* $\#\mathcal{II}^*$ *divides* $2^{n+s-d/2}$;
- *if* $\mathcal{II}^0 = \mathcal{II}^1 = \cdots = \mathcal{II}^p = 0$ *for some* $p < d/2$ *then* $\mathcal{I}^0 = \mathcal{I}^1 = \cdots = \mathcal{I}^p = \mathcal{I}^{p+1} = 0$.

Proof. The graded groups $\mathcal{I}^*$ and $\mathcal{II}^*$ are defined in (3.8). The group $\mathcal{I}^*$ has exactly s non-trivial components: these are components of dimensions $\dim l_i$, $i = 0, 1, \dots, s-1$ (by (3.6) all the numbers $\dim l_i$ are distinct). Every non-trivial component has order 2 because it is generated by an element $\bar{l_i}$ (3.7). So, two first statements of the theorem hold by the very definition of $\mathcal{I}^*$.

Suppose that $\varphi \notin I^2(F)$. If we consider on H and K/H the filtrations induced from K the exact sequence $0 \to H \to K \to K/H \to 0$ will give an exact sequence of the adjoint graded groups:

$$0 \to \mathrm{G}^*H \to \mathrm{G}^*K \to \mathrm{G}^*(K/H) \to 0 .$$

Since G^*H is torsion-free we obtain an injection $\operatorname{Tors}\mathrm{G}^*K \hookrightarrow \mathrm{G}^*(K/H)$. Note that [13, theorem 1], (3.2) and (3.3) imply $\#K/H = 2^s$. Since $\operatorname{Tors}\mathrm{G}^*K \supset \mathcal{I}^*$, $\#\mathcal{I}^* = 2^s$ and $\#\mathrm{G}^*(K/H) = \#K/H = 2^s$ we obtain that $\operatorname{Tors}\mathrm{G}^*K = \mathcal{I}^*$, i.e. $\mathcal{II}^* = 0$.

Now suppose that $\varphi \in I^2(F)$. Denote by N the subgroup of K generated by H and $2^{-s}[\mathcal{U}]$. Considering on N and K/N the induced filtrations we get an exact sequence of the adjoint graded groups

$$0 \to \mathrm{G}^*N \to \mathrm{G}^*K \to \mathrm{G}^*(K/N) \to 0 .$$

So, the torsion subgroups are connected by the exact sequence:

$$0 \to \operatorname{Tors}\mathrm{G}^*N \to \operatorname{Tors}\mathrm{G}^*K \to \operatorname{Tors}\mathrm{G}^*(K/N) .$$

The same arguments as above show that $\operatorname{Tors}\mathrm{G}^*N = \mathcal{I}^*$. Thus the latter exact sequence produces an embedding $\mathcal{II}^* \hookrightarrow \mathrm{G}^*(K/N)$. Since the quotient K/N is a cyclic group every component $\mathrm{G}^p(K/N)$ is cyclic too; whence the fourth statement of the theorem.

Since $\operatorname{rk}\mathrm{G}^{d/2}K = 2$ [4, (3.1),(2.2),(2.7)] and $\operatorname{rk}\mathrm{G}^{d/2}N = 1$ we have

$$\operatorname{rk}\mathrm{G}^{d/2}(K/N) = 1 ;$$

thereby $\mathrm{G}^p(K/N) = 0$ for $p \geq d/2$ whence the fifth statement of the theorem.

Suppose that there exists a field extension of degree 2^n completely splitting φ, let $\tilde{\varphi}$ be the form φ over this extension. Let $\tilde{P}$ be a simple $C_0(\tilde{\varphi})$-module. Put $\tilde{u} = [\mathcal{U} \otimes \tilde{P}] \in K(X_{\tilde{\varphi}})$. The multiple $2^{d/2-s}\tilde{u}$ of $\tilde{u}$ lies in $K(X_\varphi)$ and generates the quotient $K(X_\varphi)/N(X_\varphi)$. Considering the element $\tilde{u}$ itself in the quotient $K(X_{\tilde{\varphi}})/N(X_{\tilde{\varphi}})$ one has: $\tilde{u} \in \big(K(X_{\tilde{\varphi}})/N(X_{\tilde{\varphi}})\big)^{(d/2)}$. Taking the transfer we get:

$$2^n\tilde{u} \in \big(K(X_\varphi)/N(X_\varphi)\big)^{(d/2)} .$$

Consequently, $\#\operatorname{Tors}\mathrm{G}^*\big(K(X_\varphi)/N(X_\varphi)\big)$ divides $2^{n+s-d/2}$ and we have proved the sixth statement.

Let us prove the seventh one. Denote by $l_{d/2} \in K(X_{\tilde{\varphi}})$ the class of a $(d/2)$-dimensional linear subspace $L_{d/2}$ lying on $X_{\tilde{\varphi}}$. Applying the projection formula to the embedding $L_{d/2} \hookrightarrow X_{\tilde{\varphi}}$ and using (3.5) one gets: $hl_{d/2} = l_{d/2-1}$. It follows from [13, theorem 1] that $2^{d/2-s}K(X_{\tilde{\varphi}}) \subset K(X_\varphi)$. In particular, $l := 2^{d/2-s}l_{d/2} \in K(X_\varphi)$.

LEMMA 3.10. *One has in $K(X_\varphi)$: $\dim l \geq m$, $\dim 2^n l = m$ (here 2^n is as above the degree of a field extension completely splitting φ) and $\dim l_{s-1} < \dim l$.*

Proof. Two first properties are evident. The last one holds since

$$hl = h(2^{d/2-s} l_{d/2}) = 2^{d/2-s} l_{d/2-1} \equiv l_{s-1} \mod H_{(d/2-1)} .$$

$\square$

Let $\mathcal{I}_p$ be the non-trivial component of $\mathcal{I}^*$ of maximal dimension and suppose that $p \geq d/2$. To prove the last statement of the theorem it suffices to find a number $q > p$ with $\operatorname{Tors} \mathrm{G}_q K(X_\varphi) \neq 0$. Put $q = \dim l$. Since $p = \dim l_{s-1}$ we have by the lemma: $q > p$. The group $\mathrm{G}_q K(X_\varphi)$ contains a non-zero element $\bar{l}$, moreover $2^n \bar{l} = 0$ by the lemma. Thus $\operatorname{Tors} \mathrm{G}_q K(X_\varphi) \neq 0$ and we are done. $\square$

Corollary 3.11. *If for some p*

$$\operatorname{Tors} \mathrm{G}^0 K = \operatorname{Tors} \mathrm{G}^1 K = \cdots = \operatorname{Tors} \mathrm{G}^p K = 0$$

then the group $\operatorname{Tors} \mathrm{G}^{p+1} K$ *is cyclic.*

Proof. According to the theorem a group $\operatorname{Tors} \mathrm{G}^{p+1} K$ might be non-cyclic only in the case when $\varphi \in I^2(F)$ and $p < d/2$. In this case we can apply the last statement of the theorem. $\square$

4. Torsion in CH^4

Theorem 4.1. *If* $\dim \varphi > 8$ *then* $\# \operatorname{Tors} \mathrm{CH}^4(X_\varphi) \leq 4$.

Proof. If φ is isotropic, say $\varphi \simeq \mathbb{H} \perp \psi$ then $\mathrm{CH}^4(X_\varphi) \simeq \mathrm{CH}^3(X_\psi)$ [11, proposition 1], [4, (2.2)]; by [5, theorem] (see also (5.1)) $\# \operatorname{Tors} \mathrm{CH}^3(X_\psi) \leq 2$ always.

Below in the proof we assume that φ is anisotropic.

Suppose that φ is not a 4-Pfister neighbor. Then by (2.5) we have an isomorphism $\mathrm{CH}^4(X_\varphi) \simeq \mathrm{G}^4 K(X_\varphi)$. If $\varphi \notin I^2(F)$ or $\dim \varphi \leq 10$ then $\# \operatorname{Tors} \mathrm{G}^4 K(X_\varphi) = \# \mathcal{I}^4 \leq 2$ by (3.9). So, only the case $\varphi \in I^2(F)$ and $\dim \varphi \geq 12$ is left.

If $\dim \varphi > 12$ all the groups $\mathrm{CH}^p(X_\varphi)$ with $p \leq 3$ are torsion-free. Hence the groups $\mathrm{G}^p K(X_\varphi)$ with $p \leq 3$ are torsion-free too and thereby $\mathrm{G}^4 K(X_\varphi)$ is cyclic (3.11). If $\dim \varphi > 14$ let us take a quadratic extension L/F such that φ_L is isotropic. Then $\mathrm{CH}^4(X_{\varphi_L}) \simeq \mathrm{CH}^3 X_\psi$ for a quadratic form ψ with $\dim \psi > 12$ whence $\operatorname{Tors} \mathrm{CH}^4(X_{\varphi_L}) \simeq \operatorname{Tors} \mathrm{CH}^3(X_\psi) = 0$. Applying the transfer we get $2 \operatorname{Tors} \mathrm{CH}^4(X_\varphi) = 0$, i.e. $\# \operatorname{Tors} \mathrm{CH}^4(X_\varphi) \leq 2$ in this case.

If $\dim \varphi = 14$ we take a biquadratic extension L/F such that the Witt index of φ_L is at least 2. Then $\mathrm{CH}^4(X_{\varphi_L}) \simeq \mathrm{CH}^2 X_\psi$ for a quadratic form ψ with $\dim \psi = 10$ whence $\operatorname{Tors} \mathrm{CH}^4(X_{\varphi_L}) \simeq \operatorname{Tors} \mathrm{CH}^2(X_\psi) = 0$ and by the transfer argument $4 \operatorname{Tors} \mathrm{CH}^4(X_\varphi) = 0$, i.e. $\# \operatorname{Tors} \mathrm{CH}^4(X_\varphi) \leq 4$.

For a 12-dimensional quadratic form φ lying in $I^2(F)$ let us compute the order of the second kind torsion $\mathcal{II}^* \subset \mathrm{G}^* K(X_\varphi)$. Let L/F be a field extension of degree $2^{d/2-s}$ ($d = 10$ now) splitting the Clifford invariant of the form φ. Since φ_L is a 12-dimensional from from $I^3(L)$ it (completely) splits in a quadratic extension E/L [9, Satz 14]. Putting $n = \log_2 [E : F] = d/2 - s + 1$ in the formula from (3.9) we get $\# \mathcal{II}^* \leq 2$. Since $\operatorname{Tors} \mathrm{G}^p K(X_\varphi) = 0$ for $p \leq 2$ we have: $\operatorname{Tors} \mathrm{G}^3 K(X_\varphi) = \mathcal{II}^3$ (3.9). Now we can argue as follows: if $\mathcal{II}^3 \neq 0$ then $\mathcal{II}^3 = \mathcal{II}^*$, in particular $\mathcal{II}^4 = 0$, so $\operatorname{Tors} \mathrm{G}^4 K(X_\varphi) = \mathcal{I}^4$ has the order at most 2; otherwise, if $\mathcal{II}^3 = 0$ the group $\mathcal{I}^4$ is zero (3.9) and so $\operatorname{Tors} \mathrm{G}^4 K(X_\varphi) = \mathcal{II}^4$ has the order at most 2 again.

We have completed the case when φ is not a 4-Pfister neighbor. Now assume the opposite. Since a Pfister neighbor uniquely determines the Pfister superform the left-hand side term of (2.5) has now the order 2. By this cause we have to show that the right-hand side term, i.e. the group $\operatorname{Tors} \mathrm{G}^4 K(X_\varphi)$ is of order at most 2. Looking at the previous part of the current proof we see that it is always the case except when $\varphi \in I^2(F)$ and $\dim \varphi = 14$. But since a 14-dimensional quadratic form of trivial discriminant is evidently not able to be an (anisotropic!) Pfister neighbor this exception does not occurs. □

Remark 4.2. The proof of the theorem contains in fact a more precise information on what $\operatorname{Tors} \mathrm{CH}^4(X_\varphi)$ for a particular φ can be. One can also handle the case of $\dim \varphi = 7, 8$ if φ is not similar to a subform of an anisotropic 4-Pfister form — see (2.2).

5. Torsion in CH^3

Theorem 5.1 ([6]). *For any φ, one has $\# \operatorname{Tors} \mathrm{CH}^3(X_\varphi) \leq 2$.*

Proof. If φ is isotropic, say $\varphi = \mathbb{H} \perp \psi$, then $\mathrm{CH}^3(X_\varphi) \simeq \mathrm{CH}^2(X_\psi)$. Since $\# \operatorname{Tors} \mathrm{CH}^2 \leq 2$ for any quadric [4, theorem (6.1)] we are done in this case. From now on we suppose that φ is anisotropic.

Arguments like (1.1) show that $\mathrm{CH}^3(X_\varphi) \simeq \mathrm{G}^3 K(X_\varphi)$ [4, corollary (4.5)]. If $\varphi \notin I^2(F)$ or $\dim \varphi \leq 8$ then

$$\# \operatorname{Tors} \mathrm{G}^3 K(X_\varphi) \leq 2$$

by (3.9). From now on we consider only the case $\varphi \in I^2(F)$ and $\dim \varphi \geq 10$.

Since $\dim \varphi \geq 10$, the groups $\mathrm{G}^p K(X_\varphi)$ for $p \leq 2$ are torsion-free (for $p = 2$ it holds according to the computation of $\mathrm{CH}^2(X_\varphi)$ [4, theorem (6.1)]). Hence $\operatorname{Tors} \mathrm{G}^3 K(X_\varphi) = \mathit{II}^3$ (3.9) which is a cyclic group. The last we need to show is $2 \operatorname{Tors} \mathrm{CH}^3(X_\varphi) = 0$. For this it would suffice to find a quadratic extension L/F such that the group $\mathrm{CH}^3(X_{\varphi_L}) = 0$ is torsion-free (then one can use the transfer argument).

Take simply an arbitrary quadratic extension L/F which partially splits (i.e. makes isotropic) the form φ, say $\varphi_L = \mathbb{H} \perp \psi$. We have: $\mathrm{CH}^3(X_{\varphi_L}) \simeq \mathrm{CH}^2(X_\psi)$. If $\operatorname{Tors} \mathrm{CH}^2(X_\psi) = 0$ we are done.

If not then according to the computation of CH^2 the form ψ is similar to a 3-Pfister form. In this case we can compute the order of the second kind torsion $\mathit{II}^* \subset \mathrm{G}^* K(X_\varphi)$ by using the formula from (3.9). We have: $d = 8$, $s(\varphi) = 3$ (if $s(\varphi) = 4$ then φ should be isotropic as a 10-dimensional form from I^3) and since one can split φ by a field extension of degree 4 we can put $n = 2$. Thus $\# \mathit{II}^* \leq 2^{2+3-8/2} = 2$. In particular, $\# \operatorname{Tors} \mathrm{G}^3 K(X_\varphi) = \# \mathit{II}^3 \leq 2$. □

References

[1] Fulton, W. *Intersection theory.* Springer-Verlag, 1984.
[2] Hartshorne, R. *Algebraic Geometry.* Springer-Verlag, 1977.
[3] Kahn, B., Rost, M., Sujatha, R. *Unramified cohomology of quadrics.* To appear.
[4] Karpenko, N. A. *Algebro-geometric invariants of quadratic forms.* Leningrad (St. Petersburg) Math. J. 2 (1991), no. 1, 119–138.

[5] Karpenko, N. A. *Cycles of codimension 3 on a projective quadric.* J. Soviet Math. 63 (1993), no. 6, 678–682.
[6] Karpenko, N. A. *Chow groups of quadrics and index reduction formula.* Nova J. of Algebra and Geometry 3 (1995), no. 4, 357–379.
[7] Karpenko, N. A., Merkurjev, A. S. *Chow groups of projective quadrics.* Leningrad (St. Petersburg) Math. J., 2, 1991, no. 3, 655–671.
[8] Merkurjev, A. S. *K-theory of simple algebras.* Proc. Symp. Pure Math. 58.1 (1995), 65–83.
[9] Pfister, A. *Quadratische Formen in beliebigen Körpern.* Invent. Math. 1 (1966), 116–132.
[10] Quillen, D. *Higher algebraic K-theory: I.* Lecture Notes in Mathematics, 341 (1973), 85–147.
[11] Rost, M. *Some new results on the Chowgroups of quadrics.* Preprint, Regensburg, 1990.
[12] Rost, M. *Chow groups with coefficients.* Second degree thesis, 1995 (available on the WWW K-theory preprint archives, http://www.math.uiuc.edu/K-theory/).
[13] Swan, R. G. *K-theory of quadric hypersurfaces.* Ann. of Math. 122 (1985), no. 1, 113–154.
[14] Swan, R. G. *Zero cycles on quadric hypersurfaces.* Proc. Amer. Math. Soc. 107 (1989), no. 1, 43–46.

Nikita A. Karpenko
Westfälische Wilhelms-Universität
Mathematisches Institut
Einsteinstr. 62
D-48149 Münster, Germany
karpenk@math.uni-muenster.de

Twisted Pfister Forms

Detlev W. Hoffmann[1]

Received: November 29, 1995
Revised: February 27, 1996

Communicated by Alexander S. Merkurjev

Abstract. Let F be a field of characteristic $\neq 2$. In this paper we investigate quadratic forms φ over F which are anisotropic and of dimension 2^n, $n \geq 2$, such that in the Witt ring WF they can be written in the form $\varphi = \sigma - \pi$ where σ and π are anisotropic n- resp. m-fold Pfister forms, $1 \leq m < n$. We call these forms twisted Pfister forms. Forms of this type with $m = n-1$ are of great importance in the study of so-called good forms of height 2, and such forms with $m = 1$ also appear in Izhboldin's recent proof of the existence of n-fold Pfister forms τ over suitable fields F, $n \geq 3$, for which the function field $F(\tau)$ is not excellent over F. We first derive some elementary properties and try to give alternative characterizations of twisted Pfister forms. We also compute the Witt kernel $W(F(\varphi)/F)$ of a twisted Pfister form φ. Our main focus, however, will be the study of the following problems: For which forms ψ does a twisted Pfister form φ become isotropic over $F(\psi)$? Which forms ψ are equivalent to φ (i.e., the function fields $F(\varphi)$ and $F(\psi)$ are place-equivalent over F)? We also investigate how such twisted Pfister forms behave over the function field of a Pfister form of the same dimension which then leads to a generalization of the result of Izhboldin mentioned above.

1991 Mathematics Subject Classification: Primary 11E04; Secondary 11E81, 12F20.

1 Introduction

Let F be a field of characteristic $\neq 2$. WF denotes the Witt ring of non-degenerate quadratic forms over F (which we will simply call forms over F). P_nF (resp. GP_nF) denotes the set of all forms isometric (resp. similar) to n-fold Pfister forms, i.e., forms

[1] This research has been carried out during stays at the Mathematics Department of the University of Kentucky, Lexington and the Sonderforschungsbereich 343: "Diskrete Strukturen in der Mathematik", Universität Bielefeld. I would like to thank both institutions for their hospitality.

of the type $\langle\langle a_1, \cdots, a_n\rangle\rangle = \langle 1, a_1\rangle \otimes \cdots \langle 1, a_n\rangle$. We say that φ is a Pfister neighbor if there exists $\pi \in P_nF$ for some n such that φ is similar to a subform of π and $\dim\varphi > \frac{1}{2}\dim\pi = 2^{n-1}$. In this case, we say that φ is a Pfister neighbor of π. Any $\pi \in P_nF$ can be written as $\pi \simeq \langle 1\rangle \perp \pi'$. The form π' is called the pure part of π and it is uniquely determined up to isometry.

An important part of the algebraic theory of quadratic forms deals with the behavior of forms over F under a field extension K/F. Of particular interest is the case where $K = F(\psi)$ is the function field of a form ψ over F. If ψ is isotropic then $F(\psi)/F$ is purely transcendental. This situation is of not much interest with regard to the questions we will consider since one of our main goals lies in determining whether an anisotropic form φ over F becomes isotropic over K, something which cannot happen if K/F is purely transcendental. The extension K/F is said to be excellent if for any form φ over F the anisotropic part $(\varphi_K)_{\mathrm{an}}$ of φ over K is defined over F, i.e., there exists a form $\tilde\varphi$ over F such that $(\varphi_K)_{\mathrm{an}} \simeq \tilde\varphi_K$. Knebusch has shown in [K 2, Theorem 7.13] that if $F(\psi)/F$ is excellent where ψ is an anisotropic form, then ψ is a Pfister neighbor. As for the converse of this statement, it suffices to consider Pfister forms. This is because if ψ is a Pfister neighbor of τ then $F(\psi)$ and $F(\tau)$ are (place-)equivalent over F which implies that $F(\psi)$ is excellent iff $F(\tau)$ is excellent. So let $K = F(\tau)$ for some anisotropic $\tau \in P_nF$. It is easy to show that K/F is excellent for $n = 1$, and for $n = 2$ this was shown by Arason in [ELW 1, Appendix II]. It was an open problem whether K/F is always excellent for $n \geq 3$ until recently, when Izhboldin [I] gave a negative answer. In fact, he proved the even stronger result that to any anisotropic $\tau \in P_nF$, $n \geq 3$, there always exists a field extension E/F, some $\sigma \simeq \langle 1\rangle \perp \sigma' \in P_nE$ and some $d \in \dot E = E \setminus \{0\}$ not a square such that $\sigma' \perp \langle d\rangle$ is anisotropic, it becomes isotropic over $E(\tau)$, but its anisotropic part over $E(\tau)$ is not defined over E. In particular, $E(\tau)/E$ is not excellent.

Let us now turn to a seemingly unrelated problem. It is well-known that if φ is an anisotropic form over F then $\varphi_{F(\varphi)}$ is hyperbolic iff $\varphi \in GP_nF$ for some n. Going one step further, what can one say about an anisotropic form φ over F for which $\varphi_1 \simeq (\varphi_{F(\varphi)})_{\mathrm{an}}$ does not vanish but where $(\varphi_1)_{F_1(\varphi_1)}$ becomes hyperbolic where $F_1 = F(\varphi)$. Such a form is said to be of height 2. By the above, we know that $\varphi_1 \in GP_mF_1$ for some $m \geq 1$ and we say that φ has degree m. We call φ good if there exists some $\rho \in P_mF$ such that $\varphi_1 \simeq a\rho_{F_1}$ for some $a \in \dot F_1$. If one can choose $a \in \dot F$ already then φ is an excellent form in the sense of Knebusch [K 2, Section 7], and in this case one knows how φ has to look like (cf. [K 2, Lemma 10.1(i)]). An open problem is to classify anisotropic good non-excellent forms of height 2. It is believed that if φ is of that type and of degree $n-1$ then there exists some $\alpha \in P_{n-2}F$ and some 4-dimensional form β over F such that $\varphi \simeq \alpha \otimes \beta$ and $\alpha \otimes \langle\langle -d\rangle\rangle$ is anisotropic where $d = d_\pm\beta$ is the signed discriminant of β. This conjecture has been proved for $n = 2$ (cf. [K 2, Theorem 10.3]), $n = 3$ (cf. [F 2, Theorem 1.6]), and $n = 4$ (cf. [Ka, Théorème 2.12]). (It is easy to show that if φ is of this type $\alpha \otimes \beta$ then φ is good non-excellent of height 2.)

What do these forms $\alpha \otimes \beta$ of height 2 and Izhboldin's examples $\sigma' \perp \langle d\rangle$ have in common? In both cases we are dealing with anisotropic forms φ of dimension 2^n. If $\varphi \simeq \alpha \otimes \beta$ and if we write $\beta \simeq \langle d, u, v, uv\rangle$ (possibly after scaling), then in WF we have

$$\varphi = \alpha \otimes \langle\langle u, v\rangle\rangle - \alpha \otimes \langle\langle -d\rangle\rangle \ .$$

If $\varphi \simeq \sigma' \perp \langle d\rangle$ then in WF we have

$$\varphi = \sigma - \langle\langle -d\rangle\rangle \ .$$

We observe that in these situations φ can be written as the difference of an n-fold Pfister form and an $(n-1)$-fold resp. 1-fold Pfister form. We aim at a unifying concept which includes both types of forms $\alpha \otimes \beta$ and $\sigma' \perp \langle d\rangle$. This leads us quite naturally to what we call *twisted Pfister forms*. φ is said to be a twisted Pfister form if φ is anisotropic of dimension 2^n for some n, such that in WF it can be written as $\varphi = \sigma - \pi$ for some anisotropic forms $\sigma \in P_nF$ and $\pi \in P_mF$, $1 \leq m < n$. The above examples represent twisted Pfister forms at the extreme ends of the spectrum: $m = n-1$ and $m = 1$. These examples also serve as a motivation for our in-depth study of these forms. It should be emphasized that Izhboldin's striking results in [I] and his clever constructions there gave the initial impulse to our present investigations.

As simple as the structure of twisted Pfister forms appears, this class of forms leads in our opinion to a wealth of interesting results and new problems as the above examples indicate. This is somewhat surprising considering their "proximity" to ordinary Pfister forms.

In the next section, we will recall some of the important facts about function fields and generic splitting of quadratic forms which we will need rather extensively in what will follow. Starting with the basic notion of linkage of Pfister forms in Section 3, we will then give the precise definition of twisted Pfister forms and derive some of their fundamental properties as well as some alternative characterizations. For completeness' sake, we included a short Section 4 in which we compute the Witt kernel $W(F(\varphi)/F)$ of a twisted Pfister form φ. These results have been previously obtained by Fitzgerald [F 1]. In Section 5 we attack the problem of determining those forms ψ for which a twisted Pfister form φ becomes isotropic over $F(\psi)$. In Section 6 we will determine in some cases the equivalence class of a twisted Pfister form φ (here, we mean that φ is equivalent to ψ, $\varphi \sim \psi$, if $\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are isotropic). Some of our results in Sections 5 and 6 apply to an even bigger class of forms than twisted Pfister forms. The results in these two sections can be regarded as an extension and generalization of our earlier work in [H 4]. In Section 7, we consider the case of a twisted Pfister form φ of dimension 2^n and an anisotropic $\tau \in P_nF$. We generalize Izhboldin's results in [I] on when $(\varphi_{F(\tau)})_{\mathrm{an}}$ is defined over F and add some remarks about so-called $F(\tau)$-minimal forms. Finally, in Section 8, we explicitly construct $\tau \in P_nF$, $n \geq 3$, such that $F(\tau)/F$ is not excellent where F is purely transcendental of degree $n-1$ over $\mathbb{Q}$. We also generalize Izhboldin's construction of a field extension E/F such that $E(\tau)/E$ is not excellent where one starts with an arbitrary field F permitting an anisotropic Pfister form $\tau \in P_nF$, $n \geq 3$. Our construction is still based on Izhboldin's original ideas used in [I].

2 Some basic facts

In our notations and terminology we follow Lam's book [L 1] and Scharlau's book [S]. $\varphi \simeq \psi$ denotes isometry of the forms φ and ψ over F, whereas $\varphi = \psi$ stands for equality in the Witt ring WF. We write φ_{an} for the anisotropic part of φ and $i_W(\varphi)$ for its Witt index. Thus, if we denote the hyperbolic plane $\langle 1, -1\rangle$ by $\mathbb{H}$ and put

$i = i_W(\varphi)$, we have $\varphi \simeq \varphi_{\mathrm{an}} \perp (i \times \mathbb{H})$. If φ is a subform of ψ, i.e., if there exists a form η such that $\psi \simeq \varphi \perp \eta$, then we write $\varphi \subset \psi$ for short.

If K/F is a field extension and if φ is a form over F, then we denote the form which one obtains from φ by scalar extension by φ_K. The Witt kernel $W(K/F)$ is the kernel of the natural map $WF \to WK$ induced by scalar extension. We put $D_K(\varphi) = \{a \in \dot{K} \mid \langle a \rangle \subset \varphi_K\}$ and $G_K(\varphi) = \{a \in \dot{K} \mid a\varphi_K \simeq \varphi_K\}$ (we omit the subscript if $K = F$). K/F is said to be excellent if for any form φ over F there exists a form $\tilde{\varphi}$ over F such that $(\varphi_K)_{\mathrm{an}} \simeq \tilde{\varphi}$, i.e., the anisotropic part of φ over K is defined over F. A form φ over F is called K-minimal if φ is anisotropic, φ_K is isotropic, and η_K is anisotropic for any $\eta \subset \varphi$ with $\dim \eta < \dim \varphi$. Two field extensions K and L of F are called equivalent if there exist F-places $\lambda : K \to L \cup \infty$ and $\mu : L \to K \cup \infty$, we write $K \sim L$ for short. In this situation, K/F is excellent iff L/F is excellent (this follows from [K 1, Proposition 3.1], see also [ELW 1, Corollary 2.8]), and K-minimal forms are exactly the L-minimal forms. φ is said to be round (or multiplicative) if $D(\varphi) = G(\varphi)$. If φ is a Pfister form then φ is multiplicative and either anisotropic or hyperbolic (cf. [L 1, Ch. 10, Corollaries 1.6, 1.7] or [S, Ch. 4, Corollary 1.5]).

Let now φ be a form over F such that $\dim \varphi \geq 2$ and $\varphi \not\simeq \mathbb{H}$. The function field $F(\varphi)$ of φ is the function field of the projective quadric defined by $\varphi = 0$. To avoid case distinctions, we put $F(\varphi) = F$ if $\dim \varphi \leq 1$ or $\varphi \simeq \mathbb{H}$. If $\dim \varphi = n \geq 2$ then $F(\varphi)/F$ is a purely transcendental extension of degree $n - 2$ over F followed by a quadratic extension, and $F(\varphi)/F$ is purely transcendental iff φ is isotropic ([S, Ch. 4, Remark 5.2(vi)]). $F(\varphi)$ is a generic zero (or isotropy) field of φ over F, i.e., if K is any field extension of F with φ_K isotropic then there exists a place $\lambda : F(\varphi) \to K \cup \infty$ over F. We say that two forms φ and ψ are equivalent if $F(\varphi) \sim F(\psi)$, and we write $\varphi \sim \psi$. In the following proposition, we collect some more results about function fields of quadratic forms which we will need later on.

Proposition 2.1 *Let φ and ψ be anisotropic forms over F.*

(i) ([K 1, Theorem 3.3].) *$\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are both isotropic iff $F(\varphi) \sim F(\psi)$, i.e., iff $\varphi \sim \psi$.*

(ii) ([S, Ch. 4, Theorem 5.4(i)].) *$\varphi_{F(\varphi)}$ is hyperbolic iff $\varphi \in GP_nF$ for some $n \geq 1$.*

(iii) ([L 1, Ch. 7, Lemma 3.1], [S, Ch. 2, Lemma 5.1].) *If $\dim \psi = 2$ then $\varphi_{F(\psi)}$ is isotropic iff $a\psi \subset \varphi$ for some $a \in \dot{F}$.*

(iv) (Cassels-Pfister subform theorem, [S, Ch. 4, Theorem 5.4(ii)].) *If $\varphi_{F(\psi)}$ is hyperbolic then $a\psi \subset \varphi$ for any $a \in D(\varphi) \cdot D(\psi)$.*

(v) ([S, Ch. 4, Theorem 5.4(iv)].) *If ψ is a Pfister neighbor of the Pfister form π, then $\varphi_{F(\psi)}$ is hyperbolic iff there exists a form γ over F such that $\varphi \simeq \pi \otimes \gamma$. In particular, $W(F(\psi)/F) = \pi WF$.*

(vi) ([H 3, Theorem 1].) *If $\dim \varphi \leq 2^n < \dim \psi$ for some n then $\varphi_{F(\psi)}$ stays anisotropic.*

(vii) ([H 3, Proposition 2].) *If ψ is a Pfister neighbor of the Pfister form π then $\varphi \sim \psi$ iff φ is a Pfister neighbor of π.*

(viii) ([L 2, Theorem 10.1].) *Let ρ be another form over F. If $\varphi_{F(\psi)}$ is isotropic and if $\psi_{F(\rho)}$ is isotropic then $\varphi_{F(\rho)}$ is isotropic.*

(ix) ([K 2, Theorem 7.13] or [ELW 1, Examples 2.2(i)].) *If $F(\psi)/F$ is excellent then ψ is a Pfister neighbor.*

(x) (Cf. part (iii) of this proposition and [ELW 1, Appendix II by Arason].) *If ψ is a Pfister neighbor of an n-fold Pfister form, $n = 1$ or 2, then $F(\psi)/F$ is excellent.*

Let now φ be a form over F which is not hyperbolic. We define inductively fields F_i, $i \geq 0$, and forms φ_i over F_i as follows. Let $\varphi_0 \simeq \varphi_{\mathrm{an}}$ and $F_0 = F$. For $i \geq 1$ we put $F_i = F_{i-1}(\varphi_{i-1})$ and $\varphi_i \simeq ((\varphi_{i-1})_{F_{i-1}})_{\mathrm{an}}$. The smallest h for which $\dim\varphi_h \leq 1$ is called the height of φ. The tower $F_0 \subset F_1 \subset \cdots \subset F_h$ is called a generic splitting tower of φ over F, F_h is a generic splitting field of φ over F, and F_{h-1} is called the leading field of φ over F. φ_j is called the j-th kernel form of φ and $i_W(\varphi_{F_j}) = i_j(\varphi)$ the j-th Witt index of φ. By the splitting pattern of φ we mean the sequence $\{\dim\varphi_0, \dim\varphi_1, \cdots, \dim\varphi_h\}$ (this definition is different from the one given in [HuR]). The degree of φ is defined as follows. If $\dim\varphi$ is odd we put $\deg\varphi = 0$. Otherwise, we know by Proposition 2.1(ii) that $\varphi_{h-1} \in GP_nF_{h-1}$ for some $n \geq 1$. In this case we put $\deg\varphi = n$. Let $\tau \in P_nF_{h-1}$ such that φ_{h-1} is similar to τ. Then τ is called the leading form of φ. If the leading form is defined over F we say that φ is a good form (in this case there actually exists $\sigma \in P_nF$ such that $\tau \simeq \sigma_{F_{h-1}}$, cf. [K 2, Proposition 9.2]).

There are two natural filtrations of the Witt ring. One is given by the n-th powers I^nF of the ideal IF of even-dimensional forms in WF. I^nF is additively generated by the n-fold Pfister forms. One has $I^2F = \{\varphi \in IF \mid d_\pm\varphi = 1 \in \dot{F}/\dot{F}^2\}$, where $d_\pm\varphi$ denotes the signed discriminant of a form φ, and by Merkurjev's theorem [M] one has $I^3F = \{\varphi \in I^2F \mid c(\varphi) = 1\}$ where $c(\varphi)$ denotes the Clifford invariant of φ which is an element in the Brauer group BrF of F. The other filtration is given by the ideals $J_nF = \{\varphi \in WF \mid \deg\varphi \geq n\}$ (cf. [K 1, Theorem 6.4] for the fact that these sets are ideals, see also [S, Ch. 4, Theorem 7.3]). One has $I^nF \subset J_nF$ for all $n \geq 0$ (cf. [K 1, Corollary 6.6], [S, p. 164]). This is essentially the Arason-Pfister Hauptsatz which in its original form states that if $0 \neq \varphi \in I^nF$ is anisotropic then $\dim\varphi \geq 2^n$, and furthermore, if $\varphi \in I^nF$ is anisotropic and $\dim\varphi = 2^n$ then $\varphi \in GP_nF$ (see [AP, Hauptsatz and Korollar 3]). If we define $\deg'\varphi = n$ if $\varphi \in I^nF \setminus I^{n+1}F$, we thus have $\deg'\varphi \leq \deg\varphi$. It is still an open problem whether $I^nF = J_nF$ for all n and all F. This is known to be true for $n \leq 4$ (cf. [Ka, Théorème 2.8] and the references there). We will mainly work with the ideals J_nF.

Proposition 2.2 *Let φ and ψ be forms over F with φ not hyperbolic, and let $F = F_0 \subset F_1 \subset \cdots \subset F_h$ be a generic splitting tower of φ as defined above.*

(i) ([K 1, Proposition 6.9 and Corollary 6.10], see also [S, Ch. 4, Theorem 7.5].) *$I^mFJ_nF \subset J_{m+n}F$ for all $m, n \geq 0$. Furthermore, $\deg(\varphi \otimes \psi) = \deg\varphi$ iff $\dim\psi$ is odd.*

(ii) ([AK, Satz 18].) *If $\deg\varphi_{F(\psi)} > \deg\varphi$ then $\dim\psi \leq 2^n$, and if furthermore $\dim\psi = 2^n$ then $\psi \in GP_nF$ and $\varphi \equiv \psi \pmod{J_{n+1}F}$. In particular, $\psi_{F_{h-1}}$ is similar to the leading form of φ.*

(iii) ([K 1, Corollary 3.9 and Proposition 5.13].) *Let K/F be a field extension. Let $K \cdot F_j$ be the free composite of K and F_j over F. If $i_W(\varphi_K) \geq i_j(\varphi)$ then $K \cdot F_j$ is purely transcendental over K.*

3 Twisted Pfister forms

The following result is well-known (cf. [EL, Theorem 1.4]). Since we will use it quite often without always referring to it explicitly, we will include a proof at this point for the reader's convenience.

Lemma 3.1 *Let $\alpha \in WF$ be a round form, i.e., $G(\alpha) = D(\alpha)$. Let $\varphi \in WF$. If $\alpha \otimes \varphi$ represents $a \in \dot{F}$, then there exists $\psi \in WF$ with $a \in D(\psi)$ such that $\alpha \otimes \varphi \simeq \alpha \otimes \psi$. Furthermore, if $\dim \varphi \geq 2$ and $\alpha \otimes \varphi$ is isotropic, then then there exists an isotropic $\psi \in WF$ such that $\alpha \otimes \varphi \simeq \alpha \otimes \psi$.*

Proof. Let $\varphi \simeq \langle a_1, \cdots, a_n \rangle$ so that $\alpha \otimes \varphi \simeq a_1\alpha \perp \cdots \perp a_n\alpha$. Since $a \in D(\alpha \otimes \varphi)$ there are $x_i \in D(\alpha)$, not all 0, such that $a = a_1x_1 + \cdots + a_nx_n$. Say, $x_1, \cdots, x_m \neq 0$, $m \leq n$. As α is round, $x_i\alpha \simeq \alpha$ for $1 \leq i \leq m$. Thus, $\alpha \otimes \langle a_1, \cdots, a_m \rangle \simeq \alpha \otimes \langle a_1x_1, \cdots, a_mx_m \rangle$. By the above, a is represented by $\langle a_1x_1, \cdots, a_mx_m \rangle$. Hence, $\langle a_1x_1, \cdots, a_mx_m \rangle \simeq \langle a, a_2', \cdots, a_m' \rangle$ and thus $\alpha \otimes \varphi \simeq \alpha \otimes \langle a, a_2', \cdots, a_m', a_{m+1}, \cdots, a_n \rangle$.

Now suppose that $\dim \varphi \geq 2$ and that $\alpha \otimes \varphi$ is isotropic. Write $\varphi \simeq \varphi' \perp \langle -x \rangle$. By assumption, there exists $y \in \dot{F}$ such that y is represented both by $\alpha \otimes \varphi'$ and by $x\alpha$. This is clear if both forms are anisotropic because their difference is isotropic. If either one of them is isotropic, it is universal and therefore represents any non-zero element represented by the other form. By the above, $\alpha \otimes \varphi' \simeq \alpha \otimes \psi'$ with $y \in D(\psi')$ and $x\alpha \simeq y\alpha$. In particular, the form $\psi \simeq \psi' \perp \langle -y \rangle$ is isotropic and $\alpha \otimes \varphi \simeq \alpha \otimes \psi$. □

To gain a better understanding of the definition of twisted Pfister forms which we will give later it seems useful to recall another well-known result due to Elman and Lam [EL, Theorem 4.5].

Lemma 3.2 *Let $\sigma \in P_nF$ and $\pi \in P_mF$ be anisotropic with $m \leq n$. Let $a, b \in \dot{F}$. Then $i := i_W(a\sigma \perp b\pi) = 0$ or 2^r for some integer r with $0 \leq r \leq m$. Furthermore, $i \geq 1$ iff there exists $x \in \dot{F}$ such that $(a\sigma \perp b\pi)_{\mathrm{an}} \simeq x(\sigma \perp -\pi)_{\mathrm{an}}$. If $i = 2^r \geq 1$ then there exist $\alpha \in P_rF$, $\sigma_1 \in P_{n-r}F$, and $\pi_1 \in P_{m-r}F$ such that $\sigma \simeq \alpha \otimes \sigma_1$ and $\pi \simeq \alpha \otimes \pi_1$.*

Proof. We may assume that $i \geq 1$. Then there exist $u \in D(\sigma)$ and $v \in D(\pi)$ such that $au + bv = 0$. The roundness of σ and π implies that $u\sigma \simeq \sigma$ and $v\pi \simeq \pi$. Thus, with $x = au = -bv$, we have

$$a\sigma \perp b\pi \simeq au\sigma \perp bv\pi \simeq x\sigma \perp -x\pi.$$

Thus, $(a\sigma \perp b\pi)_{\mathrm{an}} \simeq x(\sigma \perp -\pi)_{\mathrm{an}}$.

Now if $i = 1$ there is nothing else to show. So let us assume that $i \geq 2$ and let α' be a common Pfister neighbor of σ and π of maximal dimension. Since $i \geq 2$ we have $\dim \alpha' \geq 2$ as both forms have at least a common 2-dimensional form, and every such 2-dimensional form is trivially a Pfister neighbor. Say, α' is a Pfister neighbor of $\alpha \in P_nF$. Since α' becomes isotropic over $F(\alpha)$, it follows that $\sigma_{F(\alpha)}$ and $\pi_{F(\alpha)}$ are also isotropic and hence hyperbolic. By the Cassels-Pfister subform theorem and because 1 is represented by σ, π, and α, there exist forms σ_0 and π_0 such that $\sigma \simeq \alpha \perp \sigma_0$ and $\pi \simeq \alpha \perp \pi_0$, cf. Proposition 2.1(iv). The maximality of $\dim \alpha'$ implies that $\dim \alpha = \dim \alpha'$. Suppose $i > \dim \alpha$. Then $\sigma_0 \perp -\pi_0$ is isotropic and there exists a $w \in \dot{F}$ which is represented both by σ_0 and π_0. In particular, the Pfister neighbor $\alpha \perp \langle w\rangle$ of $\alpha \otimes \langle\langle w\rangle\rangle$ is a common subform of σ and π, a contradiction to the maximality of $\dim \alpha' = \dim \alpha$. Thus, $i = \dim \alpha = 2^r$ for some $r \geq 1$. As for the remaining statement, there is nothing else to show if $\dim \alpha = \dim \sigma$. So suppose $\dim \sigma_0 > 0$ and let $v \in D(\sigma_0)$. Thus, the Pfister neighbor $\alpha \perp \langle v\rangle$ of $\alpha \otimes \langle\langle v\rangle\rangle$ is a subform of σ, and by an argument similar to above, we get that $\sigma \simeq \alpha \otimes \langle\langle v\rangle\rangle \perp \tilde{\sigma}$. The existence of $\sigma_1 \in P_{n-r}F$ now follows by an easy induction on $\dim \sigma_0$. The existence of π_1 can be shown in the same way. □

Definition 3.3 Let σ, π be anisotropic Pfister forms. If $i_W(\sigma \perp -\pi) = 2^r$, $r \geq 0$, then r is called the *linkage number* of σ and π. We write $\ln(\sigma, \pi) = r$. A form $\alpha \in P_r F$ such that $\sigma \simeq \alpha \otimes \sigma_1$ and $\pi \simeq \alpha \otimes \pi_1$ for suitable Pfister forms σ_1, π_1 is called a *link* of σ and π.

It should be remarked that a link α is generally not uniquely determined up to isometry.

We now consider the case of an anisotropic form φ of dimension 2^n such that in WF we have $\varphi = a\sigma + b\pi$, where $a, b \in \dot{F}$ and σ ,π are Pfister forms with $\dim \sigma \geq \dim \pi$. In view of Lemma 3.2, we then have that $\varphi \simeq x(\sigma \perp -\pi)_{\mathrm{an}}$ for some $x \in \dot{F}$. We want to exclude the case where $\varphi \in GP_nF$. An easy check then shows that we may assume $\sigma \in P_nF$ and $\pi \in P_mF$ are both anisotropic with $1 \leq m < n$ and we have $\ln(\sigma, \pi) = m - 1$. We now come to the definition of twisted Pfister forms and of what we will call weakly twisted Pfister forms, a type of form which will also appear frequently throughout the paper.

Definition 3.4 (i) Let $1 \leq m < n$. A form φ over F is called a *twisted* (n, m)-*Pfister form* (or simply (n, m)-*Pfister form*) if there exist anisotropic forms $\sigma \in P_nF$ and $\pi \in P_mF$ such that $\ln(\sigma, \pi) = m - 1$ and such that $\varphi \simeq (\sigma \perp -\pi)_{\mathrm{an}}$. In this case we say that the (n, m)-Pfister form φ is defined by (σ, π). The set of all forms isometric (resp. similar) to (n, m)-Pfister forms is denoted by $P_{n,m}F$ (resp. $GP_{n,m}F$). φ is called a *twisted Pfister form* if $\varphi \in GP_{n,m}F$ for some (n, m) with $1 \leq m < n$.

(ii) Let $1 \leq m < n$. A form φ over F is called a *weakly twisted* (n, m)-*Pfister form* if φ is anisotropic, $\dim \varphi = 2^n$, and $\varphi \equiv \pi \otimes \eta \pmod{J_nF}$ for some anisotropic $\pi \in P_mF$ and some odd-dimensional $\eta \in WF$. We call π the *twist* of φ. The set of all weakly twisted (n, m)-Pfister forms will be denoted by $P^w_{n,m}F$.

Remark 3.5 (i) Let $1 \leq m < n$. In view of Lemma 3.2 and by the remarks preceeding the definition, $\varphi \in GP_{n,m}F$ iff φ is anisotropic, $\dim \varphi = 2^n$, and there exist anisotropic $\sigma \in GP_nF$ and $\pi \in GP_mF$ such that $\varphi = \sigma + \pi$ in WF. If this is the case, then $\varphi \notin GP_nF$. In fact, $\varphi \equiv \sigma + \pi \equiv \pi \not\equiv 0 \pmod{J_nF}$ because $\sigma \in GP_nF \subset J_nF$ and $\pi \in GP_mF$ is anisotropic and thus, since $\dim \pi = 2^m < 2^n$ and by the Arason-Pfister Hauptsatz, $\pi \notin J_nF$. Similarly, $\varphi \equiv \pi \not\equiv 0 \pmod{I^nF}$.

(ii) Let now $\varphi \in P_{n,m}F$ be defined by (σ, π). Let α be a link of σ and π, i.e., $\alpha \in P_{m-1}F$ and there exist $\sigma_1 \in P_{n-m+1}F$ and $d \in \dot{F}$ such that $\sigma \simeq \alpha \otimes \sigma_1$ and $\pi \simeq \alpha \otimes \langle\langle -d\rangle\rangle$. Let σ_1' denote the pure part of σ_1, i.e., $\sigma_1 \simeq \langle 1\rangle \perp \sigma_1'$. Then $\varphi \simeq \alpha \otimes (\langle d\rangle \perp \sigma_1')$.

(iii) If $\varphi \in P_{n,m}F$ is defined by (σ, π) then $\varphi \equiv \sigma - \pi \equiv -\pi \pmod{J_nF}$. Hence, $GP_{n,m}F \subset P^w_{n,m}F$. This is generally a proper inclusion if $m \leq n - 3$ (see, e.g., Example 5.13), but it is an equality if $1 \leq n-2 \leq m \leq n-1 \leq 3$ (cf. Proposition 3.17 below).

Before we continue, let us mention some of the properties of twisted Pfister forms which will be useful later. In fact, we state these results for a possibly wider class of forms (see also Conjecture 3.9 below).

PROPOSITION 3.6 *Let $1 \leq m < n$. Let $\varphi \in WF$ be anisotropic and $\dim\varphi = 2^n$. Suppose that $\varphi \equiv x\pi \pmod{J_nF}$ for some anisotropic $\pi \in P_mF$ and some $x \in \dot{F}$. Then the following holds.*

(i) *$\varphi_{F(\pi)}$ is anisotropic and in $GP_nF(\pi)$. In particular, if $\varphi \in P_{n,m}F$ is defined by (σ, π), then $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$ is anisotropic.*

(ii) *φ is good with leading form defined by π. We have*

$$\mathrm{ht}(\varphi) = \begin{cases} 2 & \text{if} \quad m = n-1, \\ 3 & \text{if} \quad m < n-1. \end{cases}$$

In particular, $i_1(\varphi) = 2^{m-1}$ and the splitting pattern of φ is $\{2^n, 2^{n-1}, 0\}$ if $m = n-1$ and $\{2^n, 2^n - 2^m, 2^m, 0\}$ if $m < n-1$.

These statements hold in particular if $\varphi \in GP_{n,m}F$.

Proof. (i) First note that $\pi_{F(\pi)} = 0$ and thus $\varphi_{F(\pi)} \in J_nF(\pi)$. If $\varphi_{F(\pi)}$ is anisotropic then, since $\dim\varphi = 2^n$, this implies $\varphi_{F(\pi)} \in GP_nF(\pi)$. So suppose $\varphi_{F(\pi)}$ is isotropic. Then $\dim(\varphi_{F(\pi)})_{\mathrm{an}} < 2^n$ and by the Arason-Pfister Hauptsatz, $\varphi_{F(\pi)}$ is hyperbolic. Thus, there exists $\gamma \in WF$, $\dim\gamma = 2^{n-m}$, such that $\varphi \simeq \pi \otimes \gamma$. Since $n > m$ we have that $\dim\gamma$ is even, i.e., $\gamma \in IF$. But $\pi \in P_mF \subset I^mF$. Hence, $\varphi \simeq \pi \otimes \gamma \in I^{m+1}F \subset J_{m+1}F$. But clearly, $\varphi \equiv x\pi \not\equiv 0 \pmod{J_{m+1}F}$, a contradiction.

If $\varphi \in P_{n,m}F$ is defined by (σ, π), then in WF we have $\varphi = \sigma - \pi$ and thus, in $WF(\pi)$, $\varphi_{F(\pi)} = \sigma_{F(\pi)}$. Now $\varphi \equiv -\pi \pmod{J_nF}$ and by Remark 3.5(i) and by the above, it is clear that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$ is anisotropic.

(ii) Since $\dim\varphi = 2^n$ but $\varphi \notin GP_nF$, we have $\mathrm{ht}(\varphi) \geq 2$. Let $F_0 = F$, F_1, and F_2 be the first three fields in a splitting tower of φ, and let φ_1 and φ_2 be the first two kernel forms of φ. Clearly, $0 < \dim\varphi_1 < 2^n$ and $(\varphi_1)_{F_1(\pi)} \equiv 0 \pmod{J_nF_1(\pi)}$. Hence, by the Arason-Pfister Hauptsatz, $(\varphi_1)_{F_1(\pi)}$ is hyperbolic. Thus, there exists $\gamma \in WF_1$ such that $\varphi_1 \simeq \pi_{F_1} \otimes \gamma$. Comparing dimensions shows that $1 \leq \dim\gamma \leq 2^{n-m} - 1$. This shows in particular that $i_1(\varphi) \geq 2^{m-1}$. Define $\psi \in WF_1$ by $\psi \simeq \varphi_1 \perp -x\pi_{F_1} \simeq \pi_{F_1} \otimes (\gamma \perp \langle -x \rangle)$. Note that $\dim\psi \leq 2^n$ and

$$\psi \equiv \varphi_1 - x\pi_{F_1} \equiv \varphi_{F_1} - x\pi_{F_1} \equiv x\pi_{F_1} - x\pi_{F_1} \equiv 0 \pmod{J_nF}.$$

Thus, by the Arason-Pfister Hauptsatz, either ψ is hyperbolic or ψ is anisotropic and in GP_nF_1.

Suppose that ψ is hyperbolic. Let $\mu \simeq (\varphi \perp -x\pi)_{\mathrm{an}}$ over F. By definition, $\mu \equiv 0 \pmod{J_nF}$. Note that φ and π are anisotropic and $\dim\varphi = 2^n > \dim\pi = 2^m$. Hence, $0 < 2^n - 2^m \leq \dim\mu \leq 2^n + 2^m < 2^{n+1}$. Therefore, by the Arason-Pfister Hauptsatz, $2^n \leq \dim\mu < 2^{n+1}$ and we must have $\deg\mu = n$. Over $F_1 = F(\varphi)$ we have $\mu_{F_1} = \varphi_{F_1} - x\pi_{F_1} = \psi_{F_1} = 0$ and thus $\deg\mu_{F_1} = \infty > \deg\mu = n$. Now $\dim\varphi = 2^n$ and Proposition 2.2(ii) yields $\varphi \in GP_nF$, obviously a contradiction.

It follows that $\psi \simeq \varphi_1 \perp -x\pi_{F_1} \in GP_nF_1$ is anisotropic. Furthermore, $\dim\varphi_1 = 2^n - 2^m$. In particular, $\psi_{F_1(\pi)}$ is isotropic and hence hyperbolic and thus $(\varphi_1)_{F_1(\pi)}$ is hyperbolic as well. If $m = n-1$ then $\dim\varphi_1 = \dim\pi = 2^{n-1}$ which immediately yields that φ_1 is similar to π_{F_1}, which in turn implies that φ is good of height 2 with leading form defined by π. Now if $m < n-1$ then $\dim\varphi_1 = 2^n - 2^m > 2^{n-1}$. Thus, φ_1 is a Pfister neighbor with complementary form $-x\pi_{F_1}$. It follows readily that $\varphi_2 \simeq x\pi_{F_2}$ and that φ is a good form of height 3 with leading form defined by π. In fact, the second kernel form is defined by $x\pi$ already over F. □

It should be remarked that the fact that the leading form of φ is defined by π also follows directly from $\varphi \equiv x\pi \equiv \pi \pmod{J_{m+1}F}$ by [K 2, Theorem 9.6]. In our proof, we also wanted to determine the height of φ explicitly.

Corollary 3.7 *Let $1 \le m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Then $\varphi_{F(\pi)}$ is anisotropic and in $GP_nF(\pi)$, $i_1(\varphi) = 2^{m-1}$, and φ is good with leading form defined by π. Furthermore, if $F(\pi)/F$ is excellent, then there exists $\sigma \in GP_nF$ such that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$.*

Proof. Write $\varphi \equiv \pi \otimes \eta \pmod{J_nF}$ with $\dim\eta$ odd. That $\varphi_{F(\pi)}$ is anisotropic and in $GP_nF(\pi)$ can be shown as in Proposition 3.6, using the fact that $\pi \otimes \eta \not\equiv 0 \pmod{J_nF}$ as π is anisotropic and $\dim\eta$ is odd and hence $\deg(\pi\otimes\eta) = \deg\pi = m < n$ (see Proposition 2.2(i)). Similarly as before, we get that $i_1(\varphi) \ge 2^{m-1}$. We want to show that we have equality and also that φ is good with leading form defined by π. We may assume that after scaling $d_\pm\eta = 1$. It is clear that $\varphi \equiv \pi\otimes\eta \equiv \pi \pmod{J_{m+1}F}$ and it follows from [K 2, Theorem 9.6] that φ is good with leading form defined by π. Let L be the leading field of $\pi \otimes \eta$ and $K = F(\pi)$. Since $(\pi \otimes \eta)_K = 0$ we have that the free composite KL is purely transcendental over K (see Proposition 2.2(iii)). Since φ_K is anisotropic, we therefore have that φ_{KL} is anisotropic and hence φ_L is anisotropic as well. Now $(\pi\otimes\eta)_L = \pi_L$ in WL by [K 1, Proposition 6.12]. Hence, φ_L is anisotropic, $\dim\varphi_L = 2^n$, and $\varphi_L \equiv \pi_L \pmod{J_nL}$. By Proposition 3.8, we have $i_1(\varphi_L) = 2^{m-1}$. But $i_1(\varphi_L) \ge i_1(\varphi) \ge 2^{m-1}$. Hence, $i_1(\varphi) = 2^{m-1}$.

Finally, if $F(\pi)/F$ is excellent, then the existence of some $\sigma \in GP_nF$ such that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$ follows from [ELW 1, Proposition 2.11]. □

Corollary 3.8 *Let $1 \le m < n$. Let $\varphi \in P^w_{n,m}F$. Let $\psi \subset a\varphi$ for some $a \in \dot F$ and $\dim\psi > 2^n - 2^{m-1}$. Then $\psi_{F(\varphi)}$ is isotropic and $\varphi \sim \psi$.*

Proof. We have $i_1(\varphi) = i_W(\varphi_{F(\varphi)}) = 2^{m-1}$ by the previous proposition. Since ψ is similar to a subform of φ and $\dim\psi > \dim\varphi - i_1(\varphi)$, it follows readily that $\psi_{F(\varphi)}$ is isotropic. Clearly, $\psi_{F(\psi)}$ and hence $\varphi_{F(\psi)}$ are isotropic as well. Thus, $\varphi \sim \psi$. □

We finish this section with some conjectures and a characterization of forms in $GP_{n,m}F$. As already remarked, if $\varphi \in P_{n,m}F$ is defined by (σ,π), then $\varphi \equiv -\pi \pmod{J_nF}$. It would be interesting to know whether a converse of this also holds, i.e., is the following conjecture always true?

Conjecture 3.9 *Let $1 \le m < n$. Let φ be an anisotropic form over F with $\dim\varphi = 2^n$. If there exists an anisotropic $\pi \in GP_mF$ such that $\varphi \equiv \pi \pmod{J_nF}$ then $\varphi \in GP_{n,m}F$.*

This conjecture is related to the following well-known conjecture (see, for example, [Ka, Conj. 9]).

Conjecture 3.10 *Let $\psi \in J_nF$ be anisotropic and $\dim\psi < 2^n + 2^{n-1}$. Then $\dim\psi = 2^n$ and $\psi \in GP_nF$.*

By the definition of degree it is clear that $\dim\psi \geq 2^n$ and that $\psi \in GP_nF$ if $\dim\psi = 2^n$. Note also that if $\psi \simeq (\pi_1 \perp -\pi_2)_{\mathrm{an}}$ where $\pi_i \in P_nF$ and $\ln(\pi_1,\pi_2) = n-2$, then $\dim\psi = 2^n + 2^{n-1}$. So the conjecture essentially states that there is a gap in the dimensions of anisotropic forms in J_nF between 2^n and $2^n + 2^{n-1}$. We have the following results concerning these two conjectures.

Proposition 3.11 (i) *Conjecture 3.10 implies Conjecture 3.9.*
(ii) *Conjecture 3.10 holds for $n \leq 4$.*
(iii) *Conjecture 3.9 holds for $n \leq 4$.*

Proof. (i) Let φ be anisotropic, $\dim\varphi = 2^n$, and $\varphi \equiv \pi \pmod{J_nF}$ for some anisotropic $\pi \in GP_mF$ where $1 \leq m < n$. If $m = n-1$ then, for all $x \in \dot{F}$, $\pi \equiv x\pi \pmod{J_nF}$ so that in this case we may assume (after possibly scaling) that there exists $u \in D(\varphi) \cap D(\pi)$. Consider $\sigma \simeq (\varphi \perp -\pi)_{\mathrm{an}}$. We clearly have $\sigma \in J_nF$. Furthermore, $0 < 2^n - 2^m \leq \dim\sigma \leq 2^n + 2^{n-1} - 2$. The last inequality is obvious if $m < n-1$, and it follows for $m = n-1$ since we assumed that φ and π represent a common element $u \in \dot{F}$. If Conjecture 3.10 holds, we have that $\sigma \in GP_nF$. Hence, in WF, $\varphi = \sigma - \pi$ with $\sigma \in GP_nF$ and $\pi \in GP_mF$. By Remark 3.5(i) we have $\varphi \in GP_{n,m}F$.

(ii) The case $n = 2$ is trivial and the case $n = 3$ is essentially due to Pfister (cf. [P, Satz 14] or [S, Ch. 2, Theorem 14.4], the result is usually given in terms of I^3F). The case $n = 4$ can be found in [H 7], again in terms of I^4F. Here, we use that $I^nF = J_nF$ for $n \leq 4$.

(iii) follows from (ii) and (i). □

The next little result shows that Conjecture 3.9 is at least "stably" true.

Proposition 3.12 *Let $1 \leq m < n$. Let φ be an anisotropic form over F with $\dim\varphi = 2^n$. If there exists an anisotropic $\pi \in GP_mF$ such that $\varphi \equiv \pi \pmod{J_nF}$ then $\varphi \in GP_{n,m}K$ for some field extension K/F.*

Proof. Let K be the leading field of $\varphi \perp -\pi$. Since $0 \neq \varphi \perp -\pi \in J_nF$ and $\dim(\varphi \perp -\pi) < 2^{n+1}$, we have that $\deg(\varphi \perp -\pi) = n$, i.e., $((\varphi \perp -\pi)_K)_{\mathrm{an}} \simeq \sigma \in GP_nK$. Since K is obtained by taking function fields of dimension $> 2^n$ (in case $K \neq F$), φ_K and π_K are anisotropic by [H 3, Theorem 1]. Also, $\varphi_K = \sigma + \pi_K$ in WK. It is now obvious by Remark 3.5(i) that $\varphi \in GP_{n,m}K$. □

Corollary 3.13 *Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Then $\varphi \in GP_{n,m}K$ for some field extension K/F.*

Proof. Write $\varphi \equiv \pi \otimes \eta \pmod{J_nF}$ with $\dim\eta$ odd. Without loss of generality, we may assume that $d_{\pm}\eta = 1$. Let L be the leading field of $\pi \otimes \eta$. As in the proof of Corollary 3.7, we have that φ_L is anisotropic and $\varphi_L \equiv \pi_L \pmod{J_nL}$. The claim now follows immediately from Proposition 3.12. □

We know by Proposition 3.11(i) that Conjecture 3.10 implies Conjecture 3.9. It would be interesting to know whether the converse holds as well, i.e., whether these two conjectures are equivalent. At least a partial answer is given by the following.

PROPOSITION 3.14 (i) *If Conjecture 3.9 holds for* $(n,m) = (n,1)$, $n \geq 3$, *then there are no anisotropic forms of dimension* $2^n + 2$ *in* J_nF.

(ii) *If Conjecture 3.9 holds for* $(n,m) = (n,1)$ *and for* $(n,m) = (n,2)$, $n \geq 4$, *then there are no anisotropic forms of dimension* $2^n + 2$ *and* $2^n + 4$ *in* J_nF.

Proof. (i) Let $\psi \in J_nF$ and suppose that $\dim \psi = 2^n + 2$ and that ψ is anisotropic. Write $\psi \simeq \varphi \perp -\pi$ with $\dim \pi = 2$. Obviously, both $\pi \in GP_1F$ and φ are anisotropic, $\dim \varphi = 2^n$ and $\varphi \equiv \pi \pmod{J_nF}$. By assumption, this implies that $\varphi \in GP_{n,1}F$. Thus, there exist $\sigma \in GP_nF$ and $\tau \in GP_1F$ such that in WF we have $\varphi = \sigma + \tau$. Hence, $\psi = \sigma + \tau - \pi \in WF$. Now ψ and $\sigma \in J_nF$. Therefore, $\tau - \pi \in J_nF$. But $\dim \tau + \dim \pi = 4 < 2^n$ which, by the Arason-Pfister Hauptsatz, yields that $\tau - \pi = 0$ in WF. Hence, in WF we have $\psi = \sigma$. But $\dim \sigma = 2^n < \dim \psi$. Thus, ψ is isotropic, a contradiction.

(ii) By part (i), forms of dimension $2^n + 2$ in J_nF are isotropic. So let $\psi \in J_nF$ and suppose that $\dim \psi = 2^n + 4$ and that ψ is anisotropic. Write $\psi \simeq \psi' \perp -\delta$ with $\dim \delta = 2$. Let $d = d_\pm \delta$ so that δ is similar to $\langle 1, -d \rangle$. Let $L = F(\sqrt{d})$. We have that $\psi_L = \psi'_L \in J_nL$. Now $\dim \psi' = 2^n + 2$ and by assumption and part (i), we have that ψ'_L is isotropic. Hence, ψ' contains a subform similar to δ, say, $\psi' \simeq \varphi \perp x\delta$. Let $-\pi \simeq \delta \perp x\delta \in GP_2F$. Then we have $\psi \simeq \varphi \perp -\pi$ and thus, $\varphi \equiv \pi \pmod{J_nF}$ with anisotropic $\pi \in GP_2F$, anisotropic φ, $\dim \varphi = 2^n$. By assumption, this implies that $\varphi \in GP_{n,2}F$. With a reasoning analogous to the one in the proof of part (i), we conclude again that ψ is isotropic, a contradiction. □

The following result shows that the existence of a large enough Pfister neighbor as a subform of the form φ in Conjecture 3.9 is equivalent to φ being in $GP_{n,m}F$.

PROPOSITION 3.15 *Let* $1 \leq m < n$. *Let* φ *be an anisotropic form over* F *with* $\dim \varphi = 2^n$. *Suppose there exists an anisotropic* $\pi \in GP_mF$ *such that* $\varphi \equiv \pi \pmod{J_nF}$. *Then* $\varphi \in GP_{n,m}F$ *iff* φ *contains a Pfister neighbor of dimension* $2^{n-1} + 1$.

Proof. Say, $\varphi \in P_{n,m}F$. Then it follows readily from Remark 3.5(ii) (and with the notations there) that φ contains the Pfister neighbor $\alpha \otimes \sigma'_1$ of dimension $2^n - 2^{m-1} \geq 2^{n-1} + 1$.

Conversely, let $\mu \subset \varphi$ be a Pfister neighbor of dimension $2^{n-1} + 1$ of, say, $\sigma \in P_nF$, and let $x \in \dot{F}$ such that $\mu \subset x\sigma$. Define $\psi \simeq (\varphi \perp -x\sigma)_{\mathrm{an}}$. Then $\dim \psi \leq \dim \varphi + \dim \sigma - 2 \dim \mu = 2^n - 2$. Note that $\psi \equiv \varphi - x\sigma \equiv \varphi \equiv \pi \pmod{J_nF}$ and hence $\psi_{F(\pi)} \equiv 0 \pmod{J_nF(\pi)}$. By the Arason-Pfister Hauptsatz, this implies that $\psi_{F(\pi)} = 0$ in $WF(\pi)$ and there exists $\eta \in WF$ such that $\psi \simeq \pi \otimes \eta$. Since $\dim \psi \leq 2^n - 2$ we must therefore have $\dim \psi \leq 2^n - 2^m$. As $\psi \perp -\pi \in J_nF$ and $\dim(\psi \perp -\pi) \leq 2^n$, the Arason-Pfister Hauptsatz yields two cases. Either $\psi \perp -\pi = 0$ in WF. Then $\varphi = \psi + x\sigma = \pi + x\sigma$ in WF and thus $\varphi \in GP_{n,m}F$ by Remark 3.5(i). Or $\psi \perp -\pi \simeq \tau \in GP_nF$ is anisotropic. In this case, $\varphi = \psi - x\sigma = \tau - x\sigma + \pi = \gamma + \pi$, where $\gamma \simeq (\tau \perp -x\sigma)_{\mathrm{an}}$. Now $\dim \varphi = 2^n$ and $\dim \pi = 2^m$, $\tau, x\sigma \in GP_nF$, and φ, π, and γ are all anisotropic. By Lemma 3.2, $\dim \gamma = 0$, 2^n, or $\geq 2^n + 2^{n-1}$. We consider two cases. If $m \leq n-2$ then in order to have $\varphi = \gamma + \pi$, i.e., $\gamma = \varphi - \pi$, we must have $\dim \gamma = 2^n$ as

$$0 < 2^n - 2^m \leq \dim(\varphi \perp -\pi)_{\rm an} \leq 2^n + 2^m < 2^n + 2^{n-1}\ .$$

But $\gamma \in I^n F$ and thus necessarily $\gamma \in GP_nF$. Since $\varphi = \gamma + \pi$ we therefore have $\varphi \in GP_{n,m}F$. If $m = n-1$ then $\psi \perp -\pi \in GP_nF$ implies that $\psi \simeq y\pi$ for some $y \in \dot{F}$. Hence, $\varphi = \psi + x\sigma = y\pi + x\sigma$ in WF which readily implies $\varphi \in GP_{n,m}F = GP_{n,n-1}F$. □

Let us finish this section by showing that in certain cases "weakly twisted" implies "twisted" as mentioned already in Remark 3.5(iii). First, we show a very easy lemma.

Lemma 3.16 *Let $n \geq 2$, $\pi \in P_{n-2}F$ and $\tilde{\eta} \in WF$. Then there exists a form $\eta \in WF$ with $\dim \eta \leq 2$ and $\dim \eta \equiv \dim \tilde{\eta} \pmod 2$ such that $\pi \otimes \tilde{\eta} \equiv \pi \otimes \eta \pmod{J_nF}$.*

Proof. We may assume that $\dim \tilde{\eta} \geq 3$. Let us first consider the case where $\dim \tilde{\eta}$ is even. Let $a = d_{\pm}\tilde{\eta}$. Then $\tilde{\eta} \perp -\langle\langle -a\rangle\rangle \in I^2F$ and thus, since $\pi \in J_{n-2}F$, $\pi \otimes (\tilde{\eta} \perp -\langle\langle -a\rangle\rangle) \in J_nF$ or $\pi \otimes \tilde{\eta} \equiv \pi \otimes \langle\langle -a\rangle\rangle \pmod{J_nF}$ and we put $\eta \simeq \langle\langle -a\rangle\rangle$.

Let us now consider the case where $\dim \tilde{\eta}$ is odd. After scaling, we may assume that $\tilde{\eta} \simeq \langle 1\rangle \perp \eta'$. Let now $a = d_{\pm}\eta'$. By a similar argument as above, we get

$$\pi \otimes \tilde{\eta} \equiv \pi + \pi \otimes \eta' \equiv \pi + \pi \otimes \langle\langle -a\rangle\rangle \equiv \pi \otimes \langle\langle 1, -a\rangle\rangle + a\pi \equiv a\pi \pmod{J_nF}$$

because $\pi \otimes \langle\langle 1, -a\rangle\rangle \in GP_nF \subset J_nF$. Here, we put $\eta \simeq \langle a\rangle$. □

The final result in this section now follows readily from this lemma together with Proposition 3.11.

Proposition 3.17 *Let $1 \leq n-2 \leq m \leq n-1$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Then there exists $x \in \dot{F}$ such that $\varphi \equiv x\pi \pmod{J_nF}$. In particular, if Conjecture 3.9 holds for (n,m) (m as above) then $P^w_{n,m}F = GP_{n,m}F$. Thus, this equality holds whenever $1 \leq n-2 \leq m \leq n-1 \leq 3$.*

4 The Witt kernel of the function field of a twisted Pfister form

We already used several times that if $\pi \in P_nF$ is anisotropic and if $\varphi \in W(F(\pi)/F)$ is anisotropic, then there exists a form γ over F such that $\varphi \simeq \pi \otimes \gamma$. In particular, $W(F(\pi)/F)$ is a *strong n-Pfister ideal*, i.e., every anisotropic form in $W(F(\pi)/F)$ is isometric to an orthogonal sum of forms similar to n-fold Pfister forms in $W(F(\pi)/F)$, in this case forms similar to π itself. We will show that if $\varphi \in P_{n,m}F$ then $W(F(\varphi)/F)$ is a strong $(n+1)$-Pfister ideal, and we will determine the $(n+1)$-fold Pfister forms in $W(F(\varphi)/F)$. These results are implicitly contained in the work of Fitzgerald [F 1]. We will nevertheless provide a proof for the reader's convenience.

Theorem 4.1 *Let $\varphi \in P_{n,m}F$ be defined by (σ, π). Let $\alpha \in P_{m-1}F$ be a link of σ and π and let $d \in \dot{F}$ such that $\pi \simeq \alpha \otimes \langle\langle -d\rangle\rangle$. Let η be an anisotropic form over F. Then $\eta \in W(F(\varphi)/F)$ if and only if there exist an integer $k \geq 1$, $r_i, s_i \in \dot{F}$, $1 \leq i \leq k$, such that $s_i \in D(\langle d\rangle \perp -\alpha)$ and*

$$\eta \simeq \mathop{\perp}_{i=1}^{k} r_i\sigma \otimes \langle\langle s_i\rangle\rangle\ .$$

This theorem follows from the following more general result.

Theorem 4.2 *Let $\sigma \simeq \langle 1\rangle \perp \sigma' \in P_nF$, $n \geq 2$, be anisotropic and let $\gamma_1, \gamma_2 \in WF$ such that $\gamma_1 \subset \sigma'$, $\dim \gamma_1 > 2^{n-1}$, and $\gamma_1 \perp \gamma_2 \simeq \sigma$. Let $d \in \dot{F}$ such that $\psi \simeq \gamma_1 \perp \langle d\rangle$ is anisotropic and not a Pfister neighbor. Let η be an anisotropic form over F. Then $\eta \in W(F(\psi)/F)$ if and only if there exist an integer $k \geq 1$, $r_i, s_i \in \dot{F}$, $1 \leq i \leq k$, such that $s_i \in D(\langle d\rangle \perp -\gamma_2)$ and*

$$\eta \simeq \perp_{i=1}^{k} r_i\sigma \otimes \langle\langle s_i\rangle\rangle \ .$$

Proof. To show the "if"-part, it suffices to show that if $s \in D(\langle d\rangle \perp -\gamma_2)$ then $\sigma \otimes \langle\langle s\rangle\rangle \in W(F(\psi)/F)$. Now s being represented by $\langle d\rangle \perp -\gamma_2$ is equivalent to d being represented by $\langle s\rangle \perp \gamma_2$ (Witt cancellation!). Now clearly $s\sigma$ represents s. Hence,

$$\psi \simeq \gamma_1 \perp \langle d\rangle \subset \gamma_1 \perp \gamma_2 \perp \langle s\rangle \subset \sigma \perp s\sigma \simeq \sigma \otimes \langle\langle s\rangle\rangle \ .$$

It is now obvious that $\sigma \otimes \langle\langle s\rangle\rangle$ is isotropic and hence hyperbolic over $F(\psi)$.

As for the converse, let $\eta \in W(F(\psi)/F)$ be anisotropic. Since $\gamma_1 \subset \psi$ we have that η also becomes hyperbolic over $F(\gamma_1)$. But γ_1 is a Pfister neighbor of σ, i.e., $\gamma_1 \sim \sigma$ and thus $\eta \in W(F(\sigma)/F)$. Hence, there exists a form τ over F with $\eta \simeq \sigma \otimes \tau$. After scaling, we may assume that τ represents 1, i.e., $\tau \simeq \langle 1\rangle \perp \tau'$ and $\eta \simeq \sigma \perp \sigma \otimes \tau'$. Now $\dim \psi > 2^{n-1}$ and ψ is not a Pfister neighbor. In particular, ψ is not similar to a subform of $\sigma \in P_nF$ and therefore $\sigma_{F(\psi)}$ stays anisotropic. Hence, we must have $\dim \tau' \geq 1$. As $\eta_{F(\psi)} = 0$, the Cassels-Pfister subform theorem yields that for every $a \in D(\eta)\cdot D(\psi)$ we have $a\psi \subset \eta$. Since ψ and σ and therefore also η have the subform γ_1 in common, they represent common elements. Hence, we may choose $a = 1$ and we get that $\psi \subset \eta$, i.e.,

$$\psi \simeq \gamma_1 \perp \langle d\rangle \subset \eta \simeq \sigma \perp \sigma \otimes \tau' \simeq \gamma_1 \perp \gamma_2 \perp \sigma \otimes \tau' \ .$$

Hence, there exists $u \in D(\gamma_2) \cup \{0\}$ and $s \in D(\sigma \otimes \tau') \cup \{0\}$ such that $d = u + s$. Note that $d \notin D(\gamma_2)$ because otherwise $\psi \simeq \gamma_1 \perp \langle d\rangle \subset \gamma_1 \perp \gamma_2 \simeq \sigma$, i.e., ψ is a Pfister neighbor of σ, in contradiction to the definition of ψ. Hence, we must have $s \neq 0$, i.e., $s \in D(\sigma \otimes \tau')$. By Lemma 3.1, we may in fact assume that $s \in D(\tau')$ so that $\tau' \simeq \langle s\rangle \perp \tau''$. Hence, we get

$$\eta \simeq \sigma \perp s\sigma \perp \sigma \otimes \tau'' \simeq \sigma \otimes \langle\langle s\rangle\rangle \perp \sigma \otimes \tau'' \ .$$

Now $s = d - u \in D(\langle d\rangle \perp -\gamma_2)$. We have already shown that in this case, $\sigma \otimes \langle\langle s\rangle\rangle$ becomes hyperbolic over $F(\psi)$. Therefore, $\sigma \otimes \tau''$ also has to become hyperbolic over $F(\psi)$ because η does. The proof can now easily be finished by induction on $\dim \tau$. □

Proof of Theorem 4.1. As in Remark 3.5(ii), we write $\sigma \simeq \alpha \otimes \sigma_1$ for some $\sigma_1 \simeq \langle 1\rangle \perp \sigma_1' \in P_{n-m+1}F$, so that we get $\varphi \simeq \alpha \otimes (\langle d\rangle \perp \sigma_1')$. Let $\psi \simeq \alpha \otimes \sigma_1' \perp \langle d\rangle \subset \varphi$. We have $\dim \psi = 2^n - 2^{m-1} + 1$. Hence, by Corollary 3.8, $\varphi \sim \psi$ and therefore $W(F(\varphi)/F) = W(F(\psi)/F)$. Note that ψ is not a Pfister neighbor because φ is not a Pfister neighbor and the only forms equivalent to Pfister neighbors are Pfister neighbors themselves (cf. Proposition 2.1(vii)). Note also that $\sigma \simeq \alpha \otimes \sigma_1' \perp \alpha$. Now $\alpha \simeq \langle 1\rangle \perp \alpha'$ and hence $\sigma' \simeq \alpha \otimes \sigma_1' \perp \alpha'$ contains $\alpha \otimes \sigma_1'$ as a subform. The assumptions in Theorem 4.2 on ψ are then fulfilled by putting $\gamma_1 \simeq \alpha\otimes\sigma_1'$ and $\gamma_2 \simeq \alpha$. The claim of the theorem now follows immediately from Theorem 4.2 with σ, d, ψ, γ_1 and γ_2 given as above. □

5 Isotropy of twisted Pfister forms over function fields of quadratic forms

Let φ be an anisotropic Pfister form over F and $\psi \in WF$ be anisotropic with $\dim\psi \geq 2$. The fact that Pfister forms are either anisotropic or hyperbolic plus the Cassels-Pfister subform theorem imply that φ becomes isotropic over $F(\psi)$ iff ψ is similar to a subform of φ. Now suppose φ is a twisted Pfister form and ψ is as above. When is φ isotropic over $F(\psi)$? The problem turns out to be considerably more complicated and we are only able to obtain partial results. Let us start with a useful observation.

Proposition 5.1 *Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Let $\psi \in WF$ be anisotropic with $\dim\psi \geq 2$ and assume that $D(\varphi) \cap D(\psi) \neq \emptyset$. Then $\varphi_{F(\psi)}$ is isotropic iff $\psi_{F(\pi)} \subset \varphi_{F(\pi)}$. In particular, if $\varphi \in P_{n,m}F$ is defined by (σ, π) and if $D(\varphi) \cap D(\psi) \neq \emptyset$ or $D(\sigma) \cap D(\psi) \neq \emptyset$, then $\varphi_{F(\psi)}$ is isotropic iff $\psi_{F(\pi)} \subset \sigma_{F(\pi)}$.*

Proof. The second statement clearly follows from the first one since if $\varphi \in P_{n,m}F$ is defined by (σ, π) then $\varphi \equiv \sigma - \pi \equiv -\pi \pmod{J_nF}$ and $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$, see Proposition 3.6(i).

To prove the first statement, we note that by Corollary 3.7 we have that $\varphi_{F(\pi)} \in GP_nF(\pi)$ is anisotropic. If $\varphi_{F(\psi)}$ is isotropic then $\varphi_{F(\pi)(\psi)}$ is also isotropic and hence hyperbolic, and the Cassels-Pfister subform theorem together with $D(\varphi) \cap D(\psi) \neq \emptyset$ implies that $\psi_{F(\pi)} \subset \varphi_{F(\pi)}$.

Conversely, suppose that $\psi_{F(\pi)} \subset \varphi_{F(\pi)}$. Clearly, $\varphi_{F(\pi)(\psi)}$ is isotropic and hence hyperbolic because $\varphi_{F(\pi)} \in GP_nF(\pi)$. Note that $F(\pi)(\psi) \simeq F(\psi)(\pi)$. Suppose $\varphi_{F(\psi)}$ is anisotropic. Then, by Proposition 2.1(v) and since $\varphi_{F(\psi)(\pi)} = 0$, there exists $\gamma \in WF(\psi)$ such that $\varphi_{F(\psi)} \simeq \gamma \otimes \pi_{F(\psi)}$. Now $\dim\gamma = 2^{n-m}$ is even and $\pi \in GP_mF$. In particular, $\gamma \otimes \pi_{F(\psi)} \in J_{m+1}F(\psi)$ by Proposition 2.2(1). Now if we write $\varphi \equiv \pi \otimes \eta \pmod{J_nF}$ with $\dim\eta$ odd, we readily get $\varphi \equiv \pi \otimes \eta \equiv \pi \pmod{J_{m+1}F}$. Hence we have

$$\varphi_{F(\psi)} \equiv \pi_{F(\psi)} \equiv \gamma \otimes \pi_{F(\psi)} \equiv 0 \pmod{J_{m+1}F(\psi)}$$

which yields $\pi_{F(\psi)} = 0$ in $WF(\psi)$. But then $F(\psi)(\pi)/F(\psi)$ is purely transcendental. Thus, the anisotropic form $\varphi_{F(\psi)}$ stays anisotropic over $F(\psi)(\pi)$, a contradiction to $\varphi_{F(\psi)(\pi)} = 0$. $\square$

This result gives us a criterion to decide whether φ becomes isotropic over $F(\psi)$, however, it only works over $F(\pi)$. Although function fields of Pfister forms have a somewhat nicer behavior than function fields of arbitrary forms, it seems desirable to find criteria which, at least in principle, work over F itself. What we would like to have is some sort of descent from $F(\pi)$ to F where π is an anisotropic Pfister form. This can easily be achieved if $F(\pi)/F$ is an excellent field extension which is always the case for $m = 1$ and 2, but generally not for $m \geq 3$, see also Proposition 2.1(x) and Corollary 8.4.

Proposition 5.2 *Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Let $\psi \in WF$ be anisotropic with $\dim\psi \geq 2$ and assume that $D(\varphi) \cap D(\psi) \neq \emptyset$. Suppose furthermore that $F(\pi)/F$ is excellent. Then $\varphi_{F(\psi)}$ is isotropic iff there exists a form $\tilde{\psi} \in WF$, $\dim\tilde{\psi} = 2^n$, such that $\psi \subset \tilde{\psi}$ and $\tilde{\psi}_{F(\pi)} \simeq \varphi_{F(\pi)}$. In particular, if $\varphi \in P_{n,m}F$ is defined by (σ, π) and if $D(\varphi) \cap D(\psi) \neq \emptyset$ or $D(\sigma) \cap D(\psi) \neq \emptyset$, then*

$\varphi_{F(\psi)}$ is isotropic iff there exists a form $\tilde{\psi} \in WF$, $\dim \tilde{\psi} = 2^n$, such that $\psi \subset \tilde{\psi}$ and $\tilde{\psi}_{F(\pi)} \simeq \sigma_{F(\pi)} \simeq \varphi_{F(\pi)}$.

Proof. The "if"-part follows directly from Proposition 5.1 even without the excellence assumption. As for the converse, consider $\varphi \perp -\psi$. Since $\varphi_{F(\psi)}$ is isotropic, we know by Proposition 5.1 that $\psi_{F(\pi)} \subset \varphi_{F(\pi)}$. This plus the excellence of $F(\pi)/F$ imply that there exists $\chi \in WF$, $\dim \chi = 2^n - \dim \psi$, such that $\varphi_{F(\pi)} \simeq \psi_{F(\pi)} \perp \chi_{F(\pi)}$, and with $\tilde{\psi} \simeq \psi \perp \chi$ over F, we get $\tilde{\psi}_{F(\pi)} \simeq \varphi_{F(\pi)}$. □

Remark 5.3 The previous proposition provides indeed a criterion which, at least in principle, can be checked over F. This is because $\tilde{\psi}_{F(\pi)} \simeq \varphi_{F(\pi)}$ means that $\tilde{\psi} \perp -\varphi \in W(F(\pi)/F)$. In other words, with φ and ψ as above, $\varphi_{F(\psi)}$ is isotropic iff there exist forms $\tilde{\psi}$ and τ in WF with $\dim \tilde{\psi} = 2^n$ such that $\psi \subset \tilde{\psi}$ and $(\tilde{\psi} \perp -\varphi)_{\mathrm{an}} \simeq \pi \otimes \tau$.

We do not know whether Proposition 5.2 holds in general without the assumption on $F(\pi)/F$ being excellent. However, this result at least indicates that in order to decide if φ becomes isotropic over $F(\psi)$, it seems to be important to characterize those forms ψ over F of dimension 2^n for which $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$. This will be the focus of most of the remainder of this section.

We will eventually be interested in characterizing those forms $\psi \in WF$ of dimension 2^n which become isometric to some $\varphi \in P_{n,m}F$ over $F(\pi)$, where φ is defined by (σ, π). Our aim is to make this description as precise as possible, something which, in general, doesn't seem to be easy and which we will only do in the cases $m = n-1$ and $m = n-2$. In fact, the case $m = n-1$ has been dealt with in [H 4, Theorem 3.3]. We have the following result.

Theorem 5.4 *Let $\varphi \in P_{n,n-1}F$ be defined by (σ, π). Let $\psi \in WF$ with $\dim \psi = 2^n$. Then the following are equivalent.*

(i) *$\varphi_{F(\psi)}$ is isotropic.*

(ii) *$\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$.*

(iii) *Either ψ is similar to φ or ψ is similar to some $\tau \in P_nF$ and φ contains a Pfister neighbor of τ.*

Proof. The equivalence of (i) and (ii) is clear from Proposition 5.1. Clearly, (iii) implies (i). That (iii) follows from any of the other statements was shown in [H 4, Theorem 3.3] under the additional assumption that ψ contains a Pfister neighbor of dimension $2^{n-1} + 2^{n-2}$. By Proposition 5.8, (ii) implies that there exist $\alpha \in P_{n-2}F$ and $\psi_1 \in WF$, $\dim \psi_1 = 4$, such that $\psi \simeq \alpha \otimes \psi_1$. Let $\psi' \subset \psi_1$ with $\dim \psi' = 3$. Then ψ' is a Pfister neighbor of some $\beta \in P_2F$ and $\alpha \otimes \psi' \subset \psi$ is a Pfister neighbor of dimension $2^{n-1} + 2^{n-2}$ of $\alpha \otimes \beta \in P_nF$ and we can apply [H 4, Theorem 3.3] as desired. □

Corollary 5.5 *Let $\varphi \in P_{n,n-1}F$ be defined by (σ, π) and suppose that $F(\pi)/F$ is excellent (which always holds if $n-1 = 1$ or 2). Let $\psi \in WF$ with $\dim \psi \geq 2$. Then $\varphi_{F(\psi)}$ is isotropic iff ψ is similar to a subform of φ or ψ is similar to a subform of some $\tau \in P_nF$ and φ contains a Pfister neighbor of τ.*

Proof. This is a direct consequence of Proposition 5.2 and the previous theorem. □

Part (iii) of the previous theorem tells us that in order to decide whether $\varphi_{F(\psi)}$ is isotropic (where $\dim\psi = 2^n$), it suffices to look only at φ and ψ and how they relate to each other over F. The form π isn't really needed explicitly. It turns out that if $\varphi \in P_{n,m}F$ with $m < n-1$ then the situation is not quite so nice anymore as the form π will play a more prominent role. Let us start with a simple observation.

Proposition 5.6 *Let $1 \le m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Let $\psi \in WF$ be anisotropic with $\dim\psi = 2^n$. Suppose that $\varphi_{F(\psi)}$ is isotropic. Then there exists $\mu \in WF$ such that $\psi \equiv \pi\otimes\mu \pmod{J_nF}$. In particular, $\deg\psi \ge m$, and $\deg\psi = m$ iff $\dim\mu$ is odd (i.e., iff $\psi \in P^w_{n,m}F$ with twist π). Furthermore, If $m = n-1$ (resp. $m = n-2$) then there exists such μ with $\dim\mu \le 1$ (resp. $\dim\mu \le 2$).*

Proof. Write $\varphi \equiv \pi\otimes\eta \pmod{J_nF}$ with $\dim\eta$ odd. After scaling, we may assume that $D(\psi)\cap D(\varphi) \ne \emptyset$. Since $\varphi_{F(\psi)}$ is isotropic we have by Proposition 5.1 that $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$. Let $\chi \simeq (\psi \perp -\varphi)_{\mathrm{an}}$. Then $\chi \in W(F(\pi)/F)$ and there exists $\tilde\mu \in WF$ with $\chi \simeq \pi\otimes\tilde\mu$. Let us put $\mu \simeq \tilde\mu \perp \eta$. Then we have

$$\psi \equiv \chi + \varphi \equiv \pi\otimes\tilde\mu + \pi\otimes\eta \equiv \pi\otimes\mu \pmod{J_nF}\ .$$

We have $\deg(\pi\otimes\mu) = \deg\pi = m$ if $\dim\mu$ is odd (cf. Proposition 2.2(i)), in which case $\deg\psi = \deg(\pi\otimes\mu) = m$ as $m < n$. If $\dim\mu$ is even, we have $\deg(\pi\otimes\mu) \ge m+1$. Since $n \ge m+1$ we thus also have $\deg\psi \ge m+1$.

The remaining statements for $n-2 \le m \le n-1$ follow readily from Lemma 3.16. □

Remark 5.7 In the above proof, we have $\dim\chi \le 2^{n+1}-2$ and one obtains $\dim\tilde\mu \le 2^{n+1-m}-1$ or $\dim\mu \le 2^{n+1-m}-1+\dim\eta$. If $\varphi \in GP_{n,m}F$ or if $F(\pi)/F$ is excellent, then we know that there exists $\sigma \in GP_nF$ such that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)} \in GP_nF(\pi)$ (cf. Corollary 3.7 in the case where $F(\pi)/F$ is excellent). We can slightly improve the estimate of $\dim\mu$ for general m in this case. After scaling, we may assume that $\sigma \in P_nF$ and $D(\sigma)\cap D(\psi) \ne \emptyset$. Since $\varphi_{F(\psi)}$ is isotropic we then have by Proposition 5.1 that $\psi_{F(\pi)} \simeq \sigma_{F(\pi)}$. Let $\chi \simeq (\psi \perp -\sigma)_{\mathrm{an}}$. Then $\dim\chi \le 2^{n+1}-2$ as $D(\psi)\cap D(\sigma) \ne \emptyset$, and also $\chi \in W(F(\pi)/F)$. Hence, there exists $\mu \in WF$ with $\chi \simeq \pi\otimes\mu$. Since $2^m\dim\mu = \dim\chi \le 2^{n+1}-2$ we have $\dim\mu \le 2^{n+1-m}-1$. Furthermore, $\sigma \in P_nF$ and we get

$$\chi \equiv \psi - \sigma \equiv \psi \equiv \pi\otimes\mu \pmod{J_nF}\ .$$

In the case where φ is a twisted Pfister form, we can be more precise about how ψ has to look like.

Proposition 5.8 *Let $\varphi \in P_{n,m}F$ be defined by (σ,π). Let $\alpha \in P_{m-1}F$, $\sigma_1 \in P_{n-m+1}F$, and $d \in \dot F$ such that $\sigma \simeq \alpha\otimes\sigma_1$ and $\pi \simeq \alpha\otimes\langle\langle -d\rangle\rangle$ (see Remark 3.5(ii)). Let $\psi \in WF$ with $\dim\psi = 2^n$. Then the following holds.*

(i) *If $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$ then there exists $\psi_1 \in WF$, $\dim\psi_1 = 2^{n-m+1}$, such that $\psi \simeq \alpha\otimes\psi_1$. In particular, $\psi \in I^mF$, i.e., $\deg'\psi \ge m$.*

(ii) *If $\psi \in P_nF$ and $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$ then there exist $s \in \dot F$, σ_2, $\psi_2 \in P_{n-m}F$, such that $\sigma \simeq \alpha\otimes\langle\langle s\rangle\rangle\otimes\sigma_2$ and $\psi \simeq \alpha\otimes\langle\langle s\rangle\rangle\otimes\psi_2$. In particular, $\mathrm{ln}(\psi,\sigma) \ge m$ and $\mathrm{ln}(\psi,\pi) = m-1$.*

Proof. First, let us recall that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$, so that in both parts we actually assume that $\psi_{F(\pi)} \simeq \sigma_{F(\pi)}$ and therefore, we may assume that ψ represents 1 already over F after possibly scaling (this is of course always true if $\psi \in P_nF$).

(i) If $m = 1$, i.e., $\dim \alpha = 1$, there is nothing to show. So let us assume that $m \geq 2$ so that we have $\alpha_{F(\alpha)} = 0$. Since $\psi_{F(\pi)} \simeq \sigma_{F(\pi)}$ we clearly have $\psi_{F(\alpha)(\pi)} \simeq \sigma_{F(\alpha)(\pi)} \simeq (\alpha \otimes \sigma_1)_{F(\alpha)(\pi)} = 0$. Similarly, $\pi_{F(\alpha)} = 0$ which implies that $F(\alpha)(\pi)/F(\alpha)$ is purely transcendental. Hence, we must already have $\psi_{F(\alpha)} = 0$. Note that ψ is anisotropic as $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$ is anisotropic. Hence, by Proposition 2.1(v), there exists $\psi_1 \in WF$, $\dim \psi_1 = 2^{n-m+1}$, such that $\psi \simeq \alpha \otimes \psi_1$. Since $\alpha \in I^{m-1}F$ and $\dim \psi_1$ even, i.e., $\psi_1 \in IF$, we have $\psi \in I^mF$, i.e., $\deg' \psi \geq m$.

(ii) Let $\gamma \simeq (\psi \perp -\sigma)_{\mathrm{an}} \in I^nF$. By assumption, $\gamma \in W(F(\pi)/F)$. Also, $\dim \gamma \leq 2^{n+1} - 2$ as $1 \in D(\psi) \cap D(\sigma)$. Thus, by Proposition 2.1(v), there exists $\beta \in WF$, $\dim \beta < 2^{n+1-m}$, such that $\gamma \simeq \pi \otimes \beta$. Since $\gamma \in I^nF$ and $\pi \in P_mF$ with $1 \leq m < n$, we must have that $\dim \beta$ is even. Therefore, $\dim \beta \leq 2^{n+1-m} - 2$, i.e., $\dim \gamma \leq 2^{n+1} - 2^{m+1}$. Hence, $i_W(\psi \perp -\sigma) \geq 2^m$ and the existence of $s \in \dot{F}$, σ_2, $\psi_2 \in P_{n-m}F$ such that $\sigma \simeq \alpha \otimes \langle\langle s \rangle\rangle \otimes \sigma_2$ and $\psi \simeq \alpha \otimes \langle\langle s \rangle\rangle \otimes \psi_2$ now follows from Lemma 3.2 (cf., in particular, the proof of Lemma 3.2 and use the fact that we already have $\sigma \simeq \alpha \otimes \sigma_1$ and $\psi \simeq \alpha \otimes \psi_1$ by part (i)).

Clearly, $\ln(\psi, \sigma) \geq m$ since $\alpha \otimes \langle\langle s \rangle\rangle \in P_mF$ divides both ψ and σ. It is also obvious that $m \geq \ln(\psi, \pi) \geq m - 1$ as $\alpha \in P_{m-1}F$ divides both ψ and π. Now $\ln(\psi, \pi) = m$ would imply that $\pi \subset \psi$ and thus $\psi_{F(\pi)} \in P_nF(\pi)$ would be hyperbolic, a contradiction to $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$ being anisotropic. $\square$

Before we state our theorem about forms in $P_{n,n-2}F$ which parallels in a certain sense Theorem 5.4, we will provide a lemma which we will need in the proof of this theorem.

Lemma 5.9 *Let $1 \leq m \leq n-2$ and let $\varphi \in WF$ be anisotropic with $\dim \varphi = 2^n$ such that $\varphi \equiv \pi \pmod{J_nF}$ for some anisotropic $\pi \in GP_mF$. Assume furthermore that $\varphi \simeq \alpha \otimes \beta$ for some $\alpha \in P_{m-1}F$ and some $\beta \in WF$, $\dim \beta = 2^{n-m+1}$. If Conjecture 3.9 holds for $(n, m+1)$ then $\varphi \in GP_{n,m}F$.*

Proof. After scaling, we may assume that $\pi \in P_mF$. First, we note that there exists $d \in \dot{F}$ such that $\pi \simeq \alpha \otimes \langle\langle -d \rangle\rangle$. This is obvious if $m = 1$, i.e., $\alpha \simeq \langle 1 \rangle \in P_0F$. If $m > 1$ we have that $\alpha_{F(\alpha)} = 0$, thus $\varphi_{F(\alpha)} = 0$ as well, which in turn yields $\pi_{F(\alpha)} \in J_nF(\alpha)$. Since $\pi \in P_mF$ and $m < n$ we must have $\pi_{F(\alpha)} = 0$. The existence of $d \in \dot{F}$ such that $\pi \simeq \alpha \otimes \langle\langle -d \rangle\rangle$ follows immediately from Proposition 2.1(v). Write $\beta \simeq \langle x \rangle \perp \beta'$ and define $\tilde{\beta} \simeq \langle xd \rangle \perp \beta'$ and $\tilde{\varphi} \simeq \alpha \otimes \tilde{\beta}$. Then $\dim \tilde{\varphi} = \dim \varphi = 2^n$ and $\tilde{\varphi} = \varphi - x\pi$ in WF. In particular, one gets $\tilde{\varphi}_{F(\pi)} \simeq \varphi_{F(\pi)}$ which is anisotropic by Proposition 3.6. Hence, $\tilde{\varphi}$ is anisotropic. Furthermore,

$$\tilde{\varphi} \equiv \varphi - x\pi \equiv \pi - x\pi \equiv \pi \otimes \langle\langle -x \rangle\rangle \pmod{J_nF} .$$

We have two cases. If $\pi \otimes \langle\langle -x \rangle\rangle$ is isotropic and hence hyperbolic, then $\tilde{\varphi} \in J_nF$ and thus $\tilde{\varphi} \in GP_nF$ as $\dim \tilde{\varphi} = 2^n$. In WF, we get $\varphi = \tilde{\varphi} + x\pi$ which readily implies that $\varphi \in GP_{n,m}F$ as $\tilde{\varphi} \in GP_nF$ and $x\pi \in GP_mF$ (cf. Remark 3.5(i)). If $\pi \otimes \langle\langle -x \rangle\rangle$ is anisotropic then by our assumption $\tilde{\varphi} \in GP_{n,m+1}F$ and there exists $\sigma \in GP_nF$ and $\rho \in GP_{n,m+1}F$ such that $\tilde{\varphi} = \sigma + \rho$ in WF. In particular, $\tilde{\varphi} \equiv \rho \equiv \pi \otimes \langle\langle -x \rangle\rangle$

(mod J_nF), and it readily follows that there exists $y \in \dot{F}$ such that $\rho \simeq y\pi \otimes \langle\langle -x \rangle\rangle$ (recall that $m+1 < n$). Hence, in WF,

$$\begin{aligned} \varphi - \pi &= \tilde{\varphi} + x\pi - \pi = \sigma + y\pi \otimes \langle\langle -x \rangle\rangle + x\pi - \pi \\ &= \sigma + \pi \otimes \langle y, -xy, x, -1 \rangle = \sigma - \pi \otimes \langle\langle -x, -y \rangle\rangle \ . \end{aligned}$$

Suppose first that $m+2 < n$. Then $\varphi - \pi \equiv 0 \equiv -\pi \otimes \langle\langle -x, -y \rangle\rangle \pmod{J_nF}$, i.e., $\pi \otimes \langle\langle -x, -y \rangle\rangle \in J_nF$, which implies $\pi \otimes \langle\langle -x, -y \rangle\rangle = 0$ as $\pi \otimes \langle\langle -x, -y \rangle\rangle \in P_{m+2}F$ and $m+2 < n$. We then have $\varphi = \sigma + \pi$ with $\sigma \in GP_nF$ and $\pi \in P_mF$ which yields that $\varphi \in GP_{n,m}F$ as desired.

Finally, if $m+2 = n$, we have that σ, $\pi \otimes \langle\langle -x, -y \rangle\rangle \in GP_nF$. By Lemma 3.6 we then get $\dim(\sigma \perp -\pi \otimes \langle\langle -x, -y \rangle\rangle)_{\mathrm{an}} = 0$, 2^n, or $\geq 2^n + 2^{n-1}$. On the other hand,

$$\begin{aligned} 0 < 2^n - 2^m = \dim\varphi - \dim\pi \leq \dim(\varphi \perp -\pi)_{\mathrm{an}} \leq \\ \leq \dim\varphi + \dim\pi = 2^n + 2^m < 2^n + 2^{n-1} \ . \end{aligned}$$

Now $(\sigma \perp -\pi \otimes \langle\langle -x, -y \rangle\rangle)_{\mathrm{an}} \simeq (\varphi \perp -\pi)_{\mathrm{an}}$ and we therefore must have $\dim(\sigma \perp -\pi \otimes \langle\langle -x, -y \rangle\rangle)_{\mathrm{an}} = 2^n$. As $\sigma \perp -\pi \otimes \langle\langle -x, -y \rangle\rangle \in J_nF$ it follows that $(\sigma \perp -\pi \otimes \langle\langle -x, -y \rangle\rangle)_{\mathrm{an}} \simeq \tau \in GP_nF$. Hence, $\varphi = \tau + \pi$ with $\tau \in GP_nF$ and $\pi \in P_mF$ and again $\varphi \in GP_{n,m}F$. □

Theorem 5.10 *Suppose that Conjecture 3.9 holds for $(n, n-1)$. Let $\varphi \in P_{n,n-2}F$ be defined by (σ, π). Let $\psi \in WF$ with $\dim\psi = 2^n$. Then the following are equivalent.*

(i) *$\varphi_{F(\psi)}$ is isotropic.*

(ii) *$\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$.*

(iii) *There exists $\tau \in P_nF$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and either*
- *ψ is similar to τ, or*
- *there exist $x \in \dot{F}$ and $\rho \in P_{n,n-1}F$ such that ρ is defined by $(\tau, \pi \otimes \langle\langle x \rangle\rangle)$ and ψ is similar to ρ, or*
- *there exists $\chi \in P_{n,n-2}F$ such that χ is defined by (τ, π) and ψ is similar to χ.*

Proof. The equivalence of (i) and (ii) is clear from Proposition 5.1. One readily checks that (iii) implies that $\psi_{F(\pi)}$ is similar $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and we are in (ii). Finally, (i) implies by Proposition 5.6 that $\psi \equiv \pi \otimes \mu \pmod{J_nF}$ for some $\mu \in WF$, $0 \leq \dim\mu \leq 2$. If $\dim\mu \in \{0, 2\}$ then $\psi \in GP_nF$ or $\psi \in GP_{n,n-1}F$ (the latter only if $\pi \otimes \mu \neq 0$ and because we assumed that Conjecture 3.9 holds for $(n, n-1)$). If $\dim\mu = 1$ we have $\psi \in GP_{n,n-2}F$ by Lemma 5.9 together with Proposition 5.8(i). All this together with the fact that $\psi_{F(\pi)}$ is similar to $\sigma_{F(\pi)}$ readily imply (iii) and we leave the details to the reader. □

Corollary 5.11 *Let $\varphi \in P_{n,n-2}F$ be defined by (σ, π). Let $\psi \in WF$ with $\dim\psi = 2^n$. Suppose that $n \leq 4$ or that ψ contains a Pfister neighbor of dimension $2^{n-1}+1$. Then the following are equivalent.*

(i) *$\varphi_{F(\psi)}$ is isotropic.*

(ii) *$\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$.*

(iii) *There exists $\tau \in P_nF$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and either*

- *ψ is similar to τ, or*
- *there exist $x \in \dot{F}$ and $\rho \in P_{n,n-1}F$ such that ρ is defined by $(\tau, \pi \otimes \langle\langle x \rangle\rangle)$ and ψ is similar to ρ, or*
- *there exists $\chi \in P_{n,n-2}F$ such that χ is defined by (τ, π) and ψ is similar to χ.*

Proof. This is an immediate consequence of the previous theorem and Propositions 3.11 and 3.15. □

Corollary 5.12 *Suppose that Conjecture 3.9 holds for $(n, n-1)$. Let $\varphi \in P_{n,n-2}F$ be defined by (σ, π) and suppose that $F(\pi)/F$ is excellent. Let $\psi \in WF$ with $\dim \psi \geq 2$. Then the following are equivalent.*

(i) *$\varphi_{F(\psi)}$ is isotropic.*
(ii) *$\psi_{F(\pi)}$ is similar to a subform of $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$.*
(iii) *There exists $\tau \in P_nF$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and*

- *ψ is similar to a subform of τ, or*
- *there exist $x \in \dot{F}$ and $\rho \in P_{n,n-1}F$ such that ρ is defined by $(\tau, \pi \otimes \langle\langle x \rangle\rangle)$ and ψ is similar to a subform of ρ, or*
- *there exists $\chi \in P_{n,n-2}F$ such that χ is defined by (τ, π) and ψ is similar to a subform of χ.*

In particular, the equivalence of (i), (ii) *and* (iii) *always holds for $n \leq 4$.*

Proof. This is an immediate consequence of Theorem 5.10 and Proposition 5.2. Furthermore, if $n \leq 4$ then Conjecture 3.9 holds by Proposition 3.11, and $F(\pi)/F$ is excellent since π is of fold ≤ 2. □

Corollaries 5.5 and 5.12 give us a fairly complete picture for which forms $\psi \in WF$ a given form $\varphi \in P_{n,m}F$, which is defined by (σ, π), becomes isotropic over $F(\psi)$ in the cases $(n, m) \in \{(2,1), (3,1), (3,2), (4,2)\}$. In a certain sense, we know this in general in the case $(n, m) = (n, 1)$ or $(n, 2)$ by Propositions 5.1 and 5.2. It comes down to characterizing those forms $\tilde{\psi}$ of dimension 2^n for which $\tilde{\psi}_{F(\pi)} \simeq \sigma_{F(\pi)}$. In the cases $(2,1)$ and $(3,2)$ we have a very precise description by Corollary 5.5. In the cases $(3,1)$ and $(4,2)$ we can essentially reduce this problem to the determination of those $\tau \in P_nF$ with $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ or $\tau \perp -\sigma \in W(F(\pi)/F)$. This narrows down the set of forms we have to look at quite considerably.

The following example shows that if $\varphi \in P_{n,m}F$ with $n - m \geq 3$ and if $\psi \in WF$, $\dim \psi = 2^n$, then $\varphi_{F(\psi)}$ being isotropic does generally not imply that ψ is similar to a Pfister form or a twisted Pfister form, something which cannot happen in the cases considered above.

Example 5.13 Let $F = \mathbb{R}(t)$ be the rational function field in one variable t over the reals. Let $m \geq 1$ and $n - m \geq 3$. Let $\sigma \simeq \langle\langle 1, \cdots, 1 \rangle\rangle \in P_nF$ and $\pi \simeq \langle\langle 1, \cdots, 1, -t \rangle\rangle \in P_mF$. We then have

$$\varphi \simeq (\sigma \perp -\pi)_{\mathrm{an}} \simeq \langle \underbrace{1, \cdots, 1}_{2^n - 2^{m-1}}, \underbrace{t, \cdots, t}_{2^{m-1}} \rangle \in P_{n,m}F \ .$$

Let

$$\psi \simeq (\sigma \perp -\langle 1,1,1\rangle \otimes \pi)_{\rm an} \simeq \langle \underbrace{1,\cdots,1}_{2^n-3\cdot 2^{m-1}}, \underbrace{t,\cdots,t}_{3\cdot 2^{m-1}}\rangle \ .$$

One easily sees that φ and ψ are anisotropic (for example by passing to the power series field $\mathbb{R}((t)) \supset F$ and applying Springer's theorem [L 1, Ch. 6, Proposition 1.9], [S, Ch. 6, Corollary 2.6(i)]). Clearly, $\psi_{F(\pi)} \simeq \sigma_{F(\pi)} \simeq \varphi_{F(\pi)}$. Thus, by Proposition 5.1, $\varphi_{F(\psi)}$ is isotropic. We claim that ψ is neither similar to a Pfister form nor to a twisted Pfister form. First, using that σ, $\langle\langle 1,1\rangle\rangle \otimes \pi \in J_{m+2}F$, we note that

$$\psi \equiv \sigma - \langle 1,1,1\rangle \otimes \pi \equiv -\langle 1,1,1\rangle \otimes \pi + \langle\langle 1,1\rangle\rangle \otimes \pi \equiv \pi \not\equiv 0 \pmod{J_{m+2}F} \ .$$

Hence, $\deg\psi = \deg\pi = m$. Clearly, ψ is not similar to a Pfister form. Furthermore, ψ is also not similar to twisted Pfister form. For otherwise, since $\deg\psi = m$, we have $\psi \in GP_{n,m}F$ and by definition, there exist anisotropic forms $\tau \in GP_nF$ and $\rho \in GP_mF$ such that $\psi = \tau + \rho$ in WF. Thus, $\tau + \rho = \sigma - \langle 1,1,1\rangle \otimes \pi$ or $\rho + \langle 1,1,1\rangle \otimes \pi = \sigma - \tau$ in WF and we get $\dim(\rho \perp \langle 1,1,1\rangle \otimes \pi)_{\rm an} = \dim(\sigma \perp -\tau)_{\rm an}$. Now $\dim(\sigma \perp -\tau)_{\rm an} = 0$ or $\geq 2^n$ by Lemma 3.2. We also have $\dim\rho = 2^m$ and $\dim\langle 1,1,1\rangle \otimes \pi = 3\cdot 2^m$. Thus,

$$\begin{aligned} 0 < 2^{m+1} = \dim\langle 1,1,1\rangle \otimes \pi - \dim\rho \leq \dim(\rho \perp \langle 1,1,1\rangle \otimes \pi)_{\rm an} \leq \\ \leq \dim\langle 1,1,1\rangle \otimes \pi + \dim\rho = 2^{m+2} < 2^n \ . \end{aligned}$$

This obviously yields a contradiction. Note, however, that $\psi \in P^w_{n,m}F$. □

6 THE EQUIVALENCE CLASS OF A TWISTED PFISTER FORM

Recall that two forms φ and ψ over F are called equivalent, we write $\varphi \sim \psi$, if $\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are isotropic. Since the function field of an isotropic form is purely transcendental over the ground field and since anisotropic forms stay anisotropic over purely transcendental extensions, the question whether $\varphi \sim \psi$ holds is of interest only in the case of anisotropic forms. Let us denote the equivalence class of a form φ over F with respect to "$\sim$" by $\mathrm{Equiv}(\varphi)$.

We know by Proposition 2.1(vii) that if φ is an anisotropic Pfister form then $\mathrm{Equiv}(\varphi) = \{\psi \in WF \,|\, \psi \text{ is a Pfister neighbor of } \varphi\}$. The equivalence classes of forms of dimension ≤ 5 and certain forms of dimension 6, 7, and 8 have been determined in [W], [H 1], [H 2], [H 4], [Lag]. Furthermore, for forms in $P_{n,n-1}F$ we have the following result (cf. [H 4, Corollary 3.4, Theorem 4.4]).

THEOREM 6.1 *Let $n \geq 2$ and let $\varphi \in P_{n,n-1}F$.*
(i) *Let $\psi \in WF$ with $\dim\psi = 2^n$. Then $\varphi \sim \psi$ iff ψ is similar to φ.*
(ii) *Let $n \leq 3$. Then*

$$\mathrm{Equiv}(\varphi) = \{\psi \in WF \,|\, x\psi \subset \varphi \text{ for some } x \in \dot F \text{ and } \dim\psi > 2^n - 2^{n-2}\} \ .$$

In view of part (ii) of this theorem, the following conjecture seems natural (see also [H 4, Conjecture 4.3]).

CONJECTURE 6.2 *Let $n \geq 2$ and $\varphi \in P_{n,n-1}F$. Then*

$$\mathrm{Equiv}(\varphi) = \{\psi \in WF \,|\, x\psi \subset \varphi \text{ for some } x \in \dot F \text{ and } \dim\psi > 2^n - 2^{n-2}\} \ .$$

Let us right away state what we propose as the corresponding conjecture for forms in $P_{n,n-2}F$ and which we will prove to be correct in the cases $n \le 4$ (see Corollary 6.11).

Conjecture 6.3 *Let $n \ge 3$ and let $\varphi \in P_{n,n-2}F$ be defined by (σ, π). Let $\psi \in WF$. Then the following statements are equivalent.*

(i) $\psi \in \mathrm{Equiv}(\varphi)$.

(ii) *There exists $\chi \in P_{n,n-2}F$ such that*

- *χ is defined by (τ, π) for some $\tau \in P_nF$ with $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$,*
- *$x\psi \subset \chi$ for some $x \in \dot{F}$, and*
- $\dim \psi > 2^n - 2^{n-3}$.

It will be crucial to determine first those $\psi \in WF$ of dimension 2^n such that $\psi \sim \varphi$, and we will start with some more general results.

Theorem 6.4 *Let $1 \le m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Let $\psi \in WF$ be anisotropic with $\dim \psi = 2^n$. Then the following are equivalent.*

(i) $\psi \in \mathrm{Equiv}(\varphi)$ *(i.e., $\psi \sim \varphi$).*

(ii) *$\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$ and $\psi \in P^w_{n,m}F$ with twist π.*

(iii) *$\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$ and $\deg \psi = m$.*

Proof. We clearly may assume that ψ is anisotropic.

(ii)$\Rightarrow$(i). Since $\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$, and since $\varphi \equiv \pi \otimes \eta \pmod{J_nF}$ and $\psi \equiv \pi \otimes \mu \pmod{J_nF}$ with $\dim \eta \equiv \dim \mu \equiv 1 \pmod 2$, it follows directly from Proposition 5.1 and the symmetry of the situation that $\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are both isotropic. Hence, $\varphi \sim \psi$.

(i)$\Rightarrow$(ii). Let now $\varphi \sim \psi$. Then, because $\varphi_{F(\psi)}$ is isotropic, $\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$ by Proposition 5.1 and $\psi \equiv \pi \otimes \mu \pmod{J_nF}$ for some $\mu \in WF$ by Proposition 5.6. Suppose $\dim \mu$ is even so that we have $\deg(\pi \otimes \mu) \ge m+1$. Let $K = F(\pi)$ and let L be the generic splitting field of $\pi \otimes \mu$ as defined in Section 2. Then $(\pi \otimes \mu)_K = 0$ in WK and it follows that the free composite $M = KL$ is purely transcendental over K (cf. Proposition 2.2(iii)). Since φ_K is anisotropic we have that φ_{KL} is also anisotropic and thus, φ_L is anisotropic as well. Since $\varphi \sim \psi$, it follows that ψ_L stays also anisotropic. But $\psi_L \equiv (\pi \otimes \mu)_L \equiv 0 \pmod{J_nL}$ and $\dim \psi = 2^n$. This yields that $\psi \in GP_nL$. Now $\varphi \sim \psi$ also implies that $\varphi_L \sim \psi_L$ and we conclude that φ_L is similar to ψ_L, in particular, $\varphi_L \in GP_nL$ and $\deg \varphi_L = n > m = \deg \varphi$. But [AK, Satz 20] implies that $\deg \varphi_L = \deg \varphi = m$ because L is a generic splitting field of $\pi \otimes \mu$ and $\deg(\pi \otimes \mu) \ge m+1 > m$ since $\dim \mu$ is even. This is clearly a contradiction and we therefore have that $\dim \mu$ is odd.

(ii)$\Leftrightarrow$(iii). The condition that $\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$, which appears in both statements, implies that $\varphi_{F(\psi)}$ is isotropic by Proposition 5.1, and thus we get $\psi \equiv \pi \otimes \mu \pmod{J_nF}$ for some $\mu \in WF$ in both (ii) and (iii). The equivalence of (ii) and (iii) now follows from the easy observation that $\deg \psi = m$ iff $\deg(\pi \otimes \mu) = m$ iff $\dim \mu$ is odd. $\square$

Corollary 6.5 *Let $n \geq 3$. Let $\varphi \in P_{n,n-2}F$ be defined by (σ, π). Let $\psi \in WF$ be anisotropic with $\dim \psi = 2^n$. Then $\varphi \sim \psi$ iff $\psi_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$ and $\psi \equiv x\pi \pmod{J_nF}$ for some $x \in \dot{F}$.*

In particular, if Conjecture 3.9 holds for $(n, n-2)$ or $(n, n-1)$ (which is fulfilled if $n \leq 4$), or if ψ contains a Pfister neighbor of dimension $2^{n-1}+1$, then $\varphi \sim \psi$ iff ψ is similar to some $\chi \in P_{n,n-2}F$ which is defined by (τ, π) such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$.

Proof. This follows from Theorem 6.4 together with Propositions 3.6, 3.11, 3.15, 5.6, 5.8 and Lemma 5.9. We leave the details to the reader. □

Definition 6.6 Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. We define $\mathfrak{E}(\varphi)$ to be the set of all $\psi \in WF$ with $\dim \psi > 2^n - 2^{m-1}$ for which there exist $\tilde{\psi} \in P^w_{n,m}F$ with twist π such that

- $\psi \subset \tilde{\psi}$, and
- $\tilde{\psi}_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$.

In view of Theorem 6.4 and Corollary 3.8, we conjecture the following.

Conjecture 6.7 *Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Then* $\mathrm{Equiv}(\varphi) = \mathfrak{E}(\varphi)$.

Proposition 6.8 *Let $1 \leq m < n$. Let $\varphi \in P^w_{n,m}F$ with twist $\pi \in P_mF$. Then* $\mathfrak{E}(\varphi) \subset \mathrm{Equiv}(\varphi)$.

Proof. Let $\psi \in \mathfrak{E}(\varphi)$. Then there exist $\tilde{\psi}, \mu \in WF$ with $\dim \tilde{\psi} = 2^n$, $\dim \mu$ odd, such that $\tilde{\psi}_{F(\pi)}$ is similar to $\varphi_{F(\pi)}$ and $\tilde{\psi} \equiv \pi \otimes \mu \pmod{J_nF}$. By Theorem 6.4 this implies $\tilde{\psi} \sim \varphi$. Furthermore, $\tilde{\psi}$ has the property that $\psi \subset \tilde{\psi}$, and we also have $\dim \psi > 2^n - 2^{m-1}$. Hence, $\psi \sim \tilde{\psi}$ by Corollary 3.8 and therefore $\psi \sim \tilde{\psi} \sim \varphi$, i.e., $\psi \in \mathrm{Equiv}(\varphi)$. □

The main result of this section is the following.

Theorem 6.9 *Conjecture 6.7 holds for $m \leq 2$.*

Before we prove the theorem, we consider a special case.

Lemma 6.10 *Let $\varphi \in P_{n,m}F$ be defined by (σ, π) and let $\psi \subset \varphi$ with $\dim \psi \leq 2^n - 2^{m-1}$. Then there exists a field extension K/F such that $\varphi_K \in P_{n,m}K$ is defined by (σ_K, π_K) and $\psi_K \subset \sigma_K$. In particular, $\psi_{F(\varphi)}$ is anisotropic. If $m = 1$ such a K can be chosen to be of the form $K = F(\beta)$ for some $\beta \in P_2F$.*

Proof. Let K/F be a field extension. If φ_K, σ_K and π_K all stay anisotropic then one easily concludes that one still has $\varphi_K \in P_{n,m}K$ and that it is defined by (σ_K, π_K). To prove this lemma, we may assume that $\dim \psi = 2^n - 2^{m-1}$. Let $\psi' \in WF$, $\dim \psi' = 2^{m-1}$ such that $\varphi \simeq \psi \perp \psi'$. Then, in WF, we have $\varphi = \sigma - \pi = \psi + \psi'$ or $\sigma \perp -\psi = \psi' \perp \pi$. Note that $\dim \psi' = 2^{m-1} = \frac{1}{2}\dim \pi = 2^m$. By [H 3, Remark 1 and Theorem 4], there exists a field K in the generic splitting tower of $\psi' \perp \pi$ such that $i_W((\psi' \perp \pi)_K) = 2^{m-1}$, i.e., $-\psi'_K \subset \pi_K$, and π_K is anisotropic (see also [HuR, Corollaries 1.9 and 1.12]). In particular, $\dim((\psi' \perp \pi)_K)_{\mathrm{an}} = 2^{m-1}$. By comparing dimensions, we get $\sigma_K \simeq \psi_K \perp ((\psi' \perp \pi)_K)_{\mathrm{an}}$ and hence $\psi_K \subset \sigma_K$.

Note that if $m = 1$ then $\dim(\psi' \perp \pi) = 3$, so $\psi' \perp \pi$ is a Pfister neighbor of some $\beta \in P_2 F$. In our construction, the field K in the splitting tower of $\psi' \perp \pi$ is either F itself if $\psi' \perp \pi$ is already isotropic in which case $F(\beta)/F$ is purely transcendental, or it is $F(\psi' \perp \pi)$ if $\psi' \perp \pi$ is anisotropic. In this case, the field $F(\psi' \perp \pi)$ is equivalent to $F(\beta)$ since $\psi' \perp \pi$ is a Pfister neighbor of β. In any case, the field in the splitting tower which we consider is equivalent to $F(\beta)$ and thus we may as well choose $K = F(\beta)$.

It remains to show that φ_K and σ_K are anisotropic. Since $\dim \psi' = 2^{m-1} < \dim \pi$, it follows from [H 3, Theorem 1] that $\psi'_{F(\pi)}$ is anisotropic. Now $\pi_{F(\pi)} = 0$ and thus $((\psi' \perp \pi)_{F(\pi)})_{\mathrm{an}} \simeq \psi'_{F(\pi)}$ and we have $i_W((\psi' \perp \pi)_{F(\pi)}) = 2^{m-1} = i_W((\psi' \perp \pi)_K)$. By Proposition 2.2(iii), we have that $L = K{\cdot}F(\pi)$ is purely transcendental over $F(\pi)$. Now $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$ is anisotropic. Hence, $\varphi_L \simeq \sigma_L$ stays anisotropic which clearly implies that φ_K and σ_K are anisotropic. By our remark at the beginning, we have that $\varphi_K \in P_{n,m}K$ is defined by (σ_K, π_K). Since $\varphi_K \in P_{n,m}K$ we have that φ_K cannot be similar to a subform of σ_K. Therefore, $\sigma_{K(\varphi)}$ stays anisotropic and thus, $\psi_{K(\varphi)}$ stays also anisotropic. This obviously yields that $\psi_{F(\varphi)}$ is anisotropic. □

We added the additional statement in the case $m = 1$ because we will need this particular fact later on in the proof of Proposition 7.8

Proof of Theorem 6.9. Let $1 \le m \le 2$ and $m < n$. Let $\varphi \in WF$ be anisotropic and $\dim \varphi = 2^n$. Suppose that $\varphi \equiv \pi \otimes \eta \pmod{J_n F}$ for some anisotropic $\pi \in P_m F$ and some $\eta \in WF$ with $\dim \eta$ odd. By Proposition 6.8, it remains to show that $\mathrm{Equiv}(\varphi) \subset \mathfrak{E}(\varphi)$.

So let $\psi \in WF$ with $\psi \sim \varphi$. Clearly, $\dim \psi \ge 2$. Now $\psi \sim \varphi$ implies that $\varphi_{F(\psi)}$ is isotropic. Since $m \le 2$ we have that $F(\pi)/F$ is excellent. Proposition 5.2 implies that then there exists $\tilde{\psi} \in WF$, $\dim \tilde{\psi} = 2^n$, such that $\psi \subset \tilde{\psi}$ and, possibly after scaling, $\tilde{\psi}_{F(\pi)} \simeq \varphi_{F(\pi)}$. By Proposition 5.1, we have that $\varphi_{F(\tilde{\psi})}$ is isotropic. Now $\tilde{\psi}_{F(\psi)}$ is isotropic as $\psi \subset \tilde{\psi}$. We also have that $\psi_{F(\varphi)}$ is isotropic because $\psi \sim \varphi$. Hence, $\tilde{\psi}_{F(\varphi)}$ is isotropic as well (cf. Proposition 2.1(viii)), and therefore $\varphi \sim \tilde{\psi}$. By Theorem 6.4, there exists $\mu \in WF$, $\dim \mu$ odd, such that $\tilde{\psi} \equiv \pi \otimes \mu \pmod{J_n F}$. By Corollary 3.13, there exists a field extension K/F such that $\tilde{\psi}_K \in GP_{n,m}K$. We have already seen that $\psi \sim \varphi \sim \tilde{\psi}$. In particular, $\psi_{F(\tilde{\psi})}$ is isotropic which clearly yields that $\psi_{K(\tilde{\psi})}$ is also isotropic. Now $\psi_K \subset \tilde{\psi}_K \in GP_{n,m}K$. Lemma 6.10 implies that $\dim \psi > 2^n - 2^{m-1}$. This completes the proof. □

COROLLARY 6.11 *Conjecture 6.3 holds for $n \le 4$.*

Proof. This is an immediate consequence of Corollary 6.5 and Theorem 6.9. □

EXAMPLE 6.12 We return to the forms φ and ψ over $F = \mathbb{R}(t)$ which we defined in Example 5.13. We had $\varphi \in P_{n,m}F$ being defined by (σ, π) where $n - m \ge 3$. We showed that $\psi_{F(\pi)} \simeq \varphi_{F(\pi)}$ and by our construction we had that $\psi \equiv -\langle 1,1,1\rangle \otimes \pi \pmod{J_n F}$. Hence, by Theorem 6.4, $\varphi \sim \psi$. However, we also showed that $\psi \notin GP_{n,m}F$. This shows that if $\varphi \in P_{n,m}F$ and $\psi \sim \varphi$ with $\dim \psi = \dim \varphi = 2^n$ then generally this does not imply $\psi \in GP_{n,m}F$ if $n - m \ge 3$. □

We know that two anisotropic Pfister forms are equivalent iff they are isometric. To finish this section, we would like to say something about equivalence of twisted Pfister forms. Since their dimensions are always 2-powers, equivalent twisted Pfister forms must be of the same dimension. We have the following.

Proposition 6.13 *Let $1 \leq m, \ell < n$. Let $\varphi \in P_{n,m}F$ be defined by (σ, π), and let $\psi \in P_{n,\ell}F$ be defined by (τ, ρ). Then $\varphi \sim \psi$ iff $\pi \simeq \rho$ and $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. Furthermore, if this is the case then we have the following.*

(i) *If $m = n-1$ then φ is similar to ψ.*

(ii) *If $m \leq n-2$ then φ is similar to ψ iff $\sigma \simeq \tau$.*

Proof. We have $\deg \varphi = \deg \pi = m$, $\deg \psi = \deg \rho = \ell$ and $\varphi \equiv -\pi \pmod{J_n F}$, $\psi \equiv -\rho \pmod{J_n F}$. If $\varphi \sim \psi$ then, by Theorem 6.4, $m = \ell$ and $\pi \equiv \varphi \equiv \psi \equiv \rho \pmod{J_{m+1}F}$ which readily yields $\pi \simeq \rho$. Also by Theorem 6.4, we have that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)}$ is similar to $\psi_{F(\pi)} \simeq \tau_{F(\pi)}$ (here, we already use $\pi \simeq \rho$), which immediately implies $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. This shows the "only if" part. The converse follows also easily from Theorem 6.4.

Let us now assume that $\varphi \sim \psi$. If $m = n-1$ we know from Theorem 6.1 that φ is similar to ψ. So let us finally assume that $m \leq n-2$. If $\sigma \simeq \tau$ then obviously $\varphi \simeq \psi$ by definition of a twisted Pfister form. Conversely, suppose that $\varphi \simeq a\psi$ for some $a \in \dot{F}$. Then, in WF,

$$\begin{aligned} 0 &= \varphi - a\psi = \sigma - \pi - a(\tau - \pi) \\ &= \sigma - a\tau - \pi \otimes \langle\langle -a \rangle\rangle \ . \end{aligned}$$

Now σ, $\tau \in P_n F$ and $\pi \otimes \langle\langle -a \rangle\rangle \in P_{m+1}F$ with $m+1 < n$. We therefore get $0 \equiv -\pi \otimes \langle\langle -a \rangle\rangle \pmod{I^n F}$ and the Arason-Pfister Hauptsatz implies that $\pi \otimes \langle\langle -a \rangle\rangle = 0$. Hence, $0 = \sigma - a\tau$ or $\sigma \simeq a\tau$ which implies that $\sigma \simeq \tau$ as σ and τ are both n-fold Pfister forms. □

This little result has a nice application. It is of interest to determine $\mathrm{Equiv}(\varphi)$ for a given anisotropic form $\varphi \in WF$, $\dim \varphi \geq 2$. Clearly, if ψ is similar to φ then $\psi \sim \varphi$. More generally, if $a\psi \subset \varphi$ for some $a \in \dot{F}$ and $\dim \psi > \dim \varphi - i_1(\varphi)$, then $\psi_{F(\varphi)}$ is easily seen to be isotropic. Obviously, so is $\varphi_{F(\psi)}$. Hence, $\psi \sim \varphi$. Even more generally, if there exists an anisotropic $\gamma \in WF$ such that $a\varphi \subset \gamma$ and $b\psi \subset \gamma$ for some $a, b \in \dot{F}$ such that $\dim \varphi, \dim \psi > \dim \gamma - i_1(\gamma)$, then by the same reasoning as above, $\varphi \sim \gamma \sim \psi$.

Another situation where we have $\varphi \sim \psi$ (both forms anisotropic) is when $\dim \varphi = \dim \psi \geq 2$ and there exists $a \in \dot{F}$ such that $\varphi \perp a\psi$ is similar to some $\pi \in P_n F$. Clearly, we have $\dim \varphi = \dim \psi = 2^{n-1}$. Then π is isotropic and hence hyperbolic over $F(\varphi)$ and $F(\psi)$. In particular, $\varphi_{F(\varphi)} \simeq -a\psi_{F(\varphi)}$ and $\varphi_{F(\psi)} \simeq -a\psi_{F(\psi)}$. Comparing dimensions and Witt indices, we conclude that $\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are both isotropic, i.e., $\varphi \sim \psi$.

This leads to the following definitions.

Definition 6.14 Let φ, $\psi \in WF$ be anisotropic. Then φ and ψ are *neighbors* if there exists an anisotropic $\gamma \in WF$, $\dim \gamma \geq 2$, such that φ and ψ are similar to subforms of γ and $\dim \varphi, \dim \psi > \dim \gamma - i_1(\gamma)$.

φ and ψ are called *conjugate* if $\dim \varphi = \dim \psi$ and there exists $a \in \dot{F}$ such that $\varphi \perp a\psi \in GP_n F$ for some n.

If in the definition of neighbor the form γ is a similar to a Pfister form, then we have that φ and ψ are both Pfister neighbors of the same Pfister form. So this definition of neighbor is a natural generalization of a Pfister neighbor. Our definition of conjugate forms is slightly more general than the definition of conjugate forms in [K 2, Definition 8.7].

Remark 6.15 Let φ and ψ be anisotropic forms over F.

(i) If φ and ψ are similar, say, $\varphi \simeq a\psi$, then φ and ψ are neighbors. If in addition $\dim\varphi = \dim\psi = 2^n$ then φ and ψ are also conjugate. This is because $\varphi \perp -a\psi$ is isometric to the hyperbolic $(n+1)$-fold Pfister form.

(ii) Suppose that φ and ψ are neighbors and that $\dim\varphi = \dim\psi = 2^n$. If $\dim\gamma > 2^n$ then $\varphi_{F(\gamma)}$ and $\psi_{F(\gamma)}$ are anisotropic, see Proposition 2.1(vi). A form γ as in the definition above with $\dim\gamma > 2^n$ can therefore not exists. So if γ is such a form as in the definition, we must have $\dim\gamma = 2^n$ which immediately implies that φ is similar to ψ. Hence, two anisotropic forms of dimension 2^n are similar iff they are neighbors.

(iii) Suppose that $\dim\varphi = \dim\psi = 2^n$. Then φ and ψ are similar or conjugate iff there exists an $a \in \dot{F}$ such that $\varphi \perp a\psi \in W(F(\varphi)/F) \cap W(F(\psi)/F)$ (cf. [K 2, Theorem 8.8]).

Generally, conjugate forms are not similar. In a forthcoming paper, we will investigate such examples and the relationship between conjugacy and similarity.

The first examples known to us of forms φ and ψ with $\varphi \sim \psi$ but where φ and ψ are neither neighbors nor conjugate were given by twisted Pfister forms.

Proposition 6.16 *Let $n \geq 3$ and $1 \leq m \leq n-2$. Let φ, $\psi \in P_{n,m}F$ be defined by (σ,π) and (τ,π), respectively, and assume that $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$ but $\sigma \not\simeq \tau$. Then $\varphi \sim \psi$ but φ and ψ are neither neighbors nor conjugate.*

Proof. By Proposition 6.13, we know that $\varphi \sim \psi$ and that φ is not similar to ψ. By Remark 6.15(ii), φ and ψ are not neighbors.

Suppose that φ and ψ are conjugate, i.e., $\varphi \perp a\psi \in GP_{n+1}F$ for some $a \in \dot{F}$. Then

$$0 \equiv \varphi + a\psi \equiv \sigma - \pi + a(\tau - \pi) \equiv -\pi \otimes \langle\langle a \rangle\rangle \pmod{I^nF}$$

because $\varphi \perp a\psi \in GP_{n+1}F \subset I^nF$ and $\sigma, \tau \in P_nF \subset I^nF$. Since $\dim(\pi \otimes \langle\langle a \rangle\rangle) = 2^{m+1} < 2^n$, the Arason-Pfister Hauptsatz implies $\pi \otimes \langle\langle a \rangle\rangle = 0$ and thus $\varphi \perp a\psi = \sigma \perp a\tau$ in WF. Comparing dimensions and because $\varphi \perp a\psi \simeq \rho \in GP_{n+1}F$ we get that $\sigma \perp a\tau \simeq \rho \in GP_{n+1}F$. Hence, $\rho_{F(\sigma)}$ becomes isotropic and therefore hyperbolic. Thus, in $WF(\sigma)$,

$$0 = \rho_{F(\sigma)} = \sigma_{F(\sigma)} + a\tau_{F(\sigma)} = a\tau_{F(\sigma)}$$

which yields that σ is similar to a subform of τ. This in turn implies that $\sigma \simeq \tau$, a contradiction. □

In the last section, we will construct examples of forms σ, τ, and π which satisfy the conditions in Proposition 6.16. Let us conclude this section with another example of equivalent forms which are neither neighbors nor conjugate.

Example 6.17 Let $F = \mathbb{R}(t)$ and let φ and ψ be the anisotropic forms in Example 5.13. Then $\varphi \sim \psi$ but φ and ψ are neither neighbors nor conjugate. We leave the details to the reader. □

7 Twisted Pfister forms over the function field of a Pfister form

Let $\varphi \in P_{n,m}F$ be defined by (σ,π), i.e., $\sigma \in P_nF$ and $\pi \in P_mF$ are anisotropic, $\ln(\sigma,\pi) = m-1$, and $\varphi \simeq (\sigma \perp -\pi)_{\mathrm{an}}$. Let $\tau \in P_nF$. We know by Proposition 5.1 that $\sigma_{F(\tau)}$ is isotropic iff $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. So let us from now on assume that $\sigma_{F(\tau)}$ is isotropic, i.e., $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. This also implies by Proposition 5.8 that there exists $\alpha \in P_{m-1}F$ which divides π, σ, and τ.

Izhboldin [I] used twisted Pfister forms in his construction of Pfister forms which yield non-excellent function field extensions. More precisely, he essentially showed that for φ and τ as above and in the particular case where $m=1$ and $\ln(\sigma,\tau)=1$, then $(\varphi_{F(\tau)})_{\mathrm{an}}$ is *not* defined over F. It is our aim to generalize this result. First, let us note the following.

Lemma 7.1 *Let $\varphi \in P_{n,m}F$ be defined by (σ,π) and let $\tau \in P_nF$ with $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$, i.e., $\varphi_{F(\tau)}$ is isotropic. Then*

$$\dim(\varphi_{F(\tau)})_{\mathrm{an}} = \begin{cases} 2^m & \text{if } \sigma \simeq \tau \\ 2^n - 2^m & \text{if } \sigma \not\simeq \tau \end{cases}$$

Proof. Suppose first that $\sigma \simeq \tau$. Then $\sigma_{F(\tau)} = 0$ and $\pi_{F(\tau)}$ is anisotropic because $m<n$, and in $WF(\tau)$ we have $\varphi_{F(\tau)} = \sigma_{F(\tau)} - \pi_{F(\tau)} = -\pi_{F(\tau)}$. It follows immediately that $(\varphi_{F(\tau)})_{\mathrm{an}} \simeq -\pi_{F(\tau)}$ and $\dim(\varphi_{F(\tau)})_{\mathrm{an}} = 2^m$.

Now suppose that $\sigma \not\simeq \tau$. Then $\sigma_{F(\tau)}$ is anisotropic and thus we have, using $\varphi = \sigma - \pi$ in WF and $\varphi_{F(\tau)}$ isotropic, that $2^n > \dim(\varphi_{F(\tau)})_{\mathrm{an}} \geq \dim\sigma - \dim\pi = 2^n - 2^m$. By Proposition 3.6 we therefore have $\dim(\varphi_{F(\tau)})_{\mathrm{an}} = 2^n - 2^m$. □

We now come to the main result of this section

Theorem 7.2 *Let $\varphi \in P_{n,m}F$ be defined by (σ,π) and let $\tau \in P_nF$ with $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$, i.e., $\varphi_{F(\tau)}$ is isotropic. Then the following are equivalent.*

(i) *There exists a Pfister neighbor χ of τ such that $\chi \subset \varphi$.*

(ii) *There exists a Pfister neighbor χ of τ with $\dim\chi = 2^{n-1}+2^{m-1}$ such that $\chi \subset \varphi$.*

(iii) *$(\varphi_{F(\tau)})_{\mathrm{an}}$ is defined over F.*

(iv) $\ln(\sigma,\tau) \geq n-1$.

Proof. (i)⇒(ii). Let $a \in \dot{F}$ and $\chi \subset \varphi$ with $\dim\chi \geq 2^{n-1}+1$ such that $\chi \subset a\tau$. Let $\eta \simeq (\varphi \perp -a\tau)_{\mathrm{an}}$. Then $\dim\eta \leq \dim\varphi + \dim\tau - 2\dim\chi \leq 2^n - 2$. Note that we have $\eta \equiv \varphi - a\tau \equiv \sigma - \pi - a\tau \equiv -\pi \pmod{I^nF}$ as σ, $\tau \in P_nF \subset I^nF$. Thus, we get $\eta_{F(\pi)} \equiv 0 \pmod{I^nF(\pi)}$ and the Arason-Pfister Hauptsatz implies $\eta_{F(\pi)} = 0$ or $\eta_{F(\pi)} \in W(F(\pi)/F)$. Hence, by Proposition 2.1(v), there exists $\mu \in WF$ such that $\eta \simeq \mu \otimes \pi$. Thus, $\dim\pi = 2^m$ divides $\dim\eta$ and therefore $\dim\eta = \dim(\varphi \perp -a\tau)_{\mathrm{an}} \leq 2^n - 2^m$ or $i_W(\varphi \perp -a\tau) \geq 2^{n-1}+2^{m-1}$. In particular, φ and $a\tau$ have a common subform of dimension $2^{n-1}+2^{m-1}$.

(ii)⇒(iii). If $\sigma \simeq \tau$ we have already seen in the proof of Lemma 7.1 that $(\varphi_{F(\tau)})_{\mathrm{an}} \simeq -\pi_{F(\tau)}$ and we are done. Hence, we may assume that $\sigma \not\simeq \tau$ and thus $\dim(\varphi_{F(\tau)})_{\mathrm{an}} = 2^n - 2^m$ by Lemma 7.1. Let $\chi \subset \varphi$ such that $\dim\chi = 2^{n-1}+2^{m-1}$

and $\chi \subset a\tau$ for some $a \in \dot{F}$. Write $\varphi \simeq \chi \perp \tilde{\varphi}$ and $a\tau \simeq \chi \perp \tilde{\tau}$ for suitable $\tilde{\varphi}$, $\tilde{\tau} \in WF$. In $WF(\tau)$, we have

$$\varphi_{F(\tau)} = (\varphi \perp -a\tau)_{F(\tau)} = (\tilde{\varphi} \perp -\tilde{\tau})_{F(\tau)} \ .$$

Now an easy check shows that $\dim(\tilde{\varphi} \perp -\tilde{\tau}) = 2^n - 2^m = \dim(\varphi_{F(\tau)})_{\mathrm{an}}$. Therefore, we must have $(\varphi_{F(\tau)})_{\mathrm{an}} \simeq (\tilde{\varphi} \perp -\tilde{\tau})_{F(\tau)}$ and we see that $(\varphi_{F(\tau)})_{\mathrm{an}}$ is defined over F by $\tilde{\varphi} \perp -\tilde{\tau}$.

(iii)$\Rightarrow$(iv). Let $\eta \in WF$ such that $(\varphi_{F(\tau)})_{\mathrm{an}} \simeq \eta_{F(\tau)}$. By Lemma 7.1 we have $2^m \leq \dim \eta \leq 2^n - 2^m$ and thus $0 < \dim(\varphi \perp -\eta)_{\mathrm{an}} < 2^{n+1}$. Also, by our choice of η, $(\varphi \perp -\eta)_{\mathrm{an}} \in W(F(\tau)/F)$ and by Proposition 2.1(v) there exists $a \in \dot{F}$ such that $(\varphi \perp -\eta)_{\mathrm{an}} \simeq a\tau$. Hence, in WF we get $a\tau = \varphi - \eta = \sigma - \pi - \eta$ or $\sigma \perp -a\tau = \pi \perp \eta$. Now $\dim(\pi \perp \eta) \leq 2^m + 2^n - 2^m = 2^n$ and we get that $i_W(\sigma \perp -a\tau) \geq 2^{n-1}$. By Lemma 3.2, $\ln(\sigma, \tau) \geq n - 1$.

(iv)$\Rightarrow$(i) This is rather obvious in the case where $\sigma \simeq \tau$, i.e., $\ln(\sigma, \tau) = n$.

So let us assume that $\ln(\sigma, \tau) = n - 1$ and let $\rho \simeq (\sigma \perp -\tau)_{\mathrm{an}}$. Then $\dim \rho = 2^n$ and in fact $\rho \in GP_nF$ as $\rho \in I^nF$. Let $\psi \simeq (\varphi \perp -\tau)_{\mathrm{an}}$. Then, in WF,

$$\psi = \varphi - \tau = \sigma - \pi - \tau = \rho - \pi \ .$$

This yields

$$2^n - 2^m = \dim \rho - \dim \pi \leq \dim \psi \leq \dim \rho + \dim \pi = 2^n + 2^m \ .$$

On the other hand, in $WF(\pi)$,

$$\psi_{F(\pi)} = \varphi_{F(\pi)} - \tau_{F(\pi)} = \sigma_{F(\pi)} - \tau_{F(\pi)} - \pi_{F(\pi)} = 0$$

as $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$ and $\pi_{F(\pi)} = 0$. Thus, there exists $\gamma \in WF$ with $\psi \simeq \gamma \otimes \pi$ by Proposition 2.1(v). Now $\dim \gamma$ must be odd for otherwise $\psi \in I^{m+1}F$, but $\psi \equiv \sigma - \pi - \tau \equiv -\pi \not\equiv 0 \pmod{I^{m+1}F}$. Hence, there are two cases. Either $\dim \psi = 2^n - 2^m$ or $\dim \psi = 2^n + 2^m$. If $\dim \psi = 2^n - 2^m$ then by the definition of ψ we get $i_W(\varphi \perp -\tau) = \frac{1}{2}(\dim \varphi + \dim \tau - \dim \psi) = 2^{n-1} + 2^{m-1}$. Thus, φ and τ have a common subform of dimension $2^{n-1} + 2^{m-1}$ and we are done.

So let us finally assume that $\dim \psi = 2^n + 2^m$ so that in fact $\psi \simeq (\varphi \perp -\tau)_{\mathrm{an}} \simeq \rho \perp -\pi$. Now π divides ψ and we have that π also divides ρ. But $\rho \in GP_nF$. Hence, there exist $\delta \in P_{n-m}F$ and $x \in \dot{F}$ such that $\rho \simeq x\pi \otimes \delta$. Thus,

$$x\psi \simeq \pi \otimes (\delta \perp \langle -x \rangle) \subset \pi \otimes \delta \otimes \langle\langle -x \rangle\rangle \in P_{n+1}F \ .$$

This shows that ψ is a Pfister neighbor of $\beta \simeq \pi \otimes \delta \otimes \langle\langle -x \rangle\rangle \in P_{n+1}F$. Since ψ is anisotropic, β is anisotropic as well. Also, $\psi_{F(\tau)} = \varphi_{F(\tau)} - \tau_{F(\tau)} = \varphi_{F(\tau)}$ in $WF(\tau)$. Comparing dimensions, we conclude that $\psi_{F(\tau)}$ is isotropic and that therefore also $\beta_{F(\tau)}$ is isotropic and hence hyperbolic, which in turn implies that $\beta \simeq \tau \otimes \langle\langle t \rangle\rangle$ for some $t \in \dot{F}$. Since $x\psi \subset \beta$ we get $\psi \subset x\beta \simeq \tau \otimes \langle x, xt \rangle$. Now $\psi \perp \tau = \varphi$ in WF and by comparing dimensions we see that $\psi \perp \tau$ is isotropic. In particular, there exists $y \in D(\psi) \perp D(-\tau)$. Since $y \in D(\psi) \subset D(\tau \otimes \langle x, xt \rangle)$ and since $\tau \in P_nF$, we may assume by Lemma 3.1 that for suitable $z \in \dot{F}$ we have $\tau \otimes \langle x, xt \rangle \simeq \tau \otimes \langle y, z \rangle$.

But $-y \in D(\tau) = G(\tau)$. If we write $\tau \otimes \langle x, xt\rangle \simeq \psi \perp \mu$ for suitable $\mu \in WF$ with $\dim \mu = 2^{n+1} - \dim \psi = 2^n - 2^m$, we obtain

$$\psi \perp \mu \simeq \tau \otimes \langle x, xt\rangle \simeq y\tau \perp z\tau \simeq -\tau \perp z\tau$$

and thus, in WF,

$$\mu = z\tau - \tau - \psi = z\tau - \varphi$$

(here we use $\psi = \varphi - \tau$.) In particular, $(z\tau \perp -\varphi)_{\mathrm{an}} \simeq \mu$ (recall that $\psi \perp \mu$ is similar to the anisotropic Pfister form β and that therefore μ is also anisotropic), which implies that $i_W(z\tau \perp -\varphi) = \frac{1}{2}(\dim \tau + \dim \varphi - \dim \mu) = 2^{n-1} + 2^{m-1}$, i.e., φ and $z\tau$ have a common subform of dimension $2^{n-1} + 2^{m-1}$ which is obviously a Pfister neighbor of τ. □

Remark 7.3 If φ contains a Pfister neighbor of τ of dimension $> 2^{n-1} + 2^{m-1}$ then $\sigma \simeq \tau$. For in this case, there exists $a \in \dot{F}$ such that $i_W(\varphi \perp -a\tau) > 2^{n-1} + 2^{m-1}$ or $\dim(\varphi \perp -a\tau)_{\mathrm{an}} < 2^n - 2^m$. But in WF we have $\varphi \perp -a\tau = \sigma \perp -a\tau \perp -\pi$ and we must necessarily have $\dim(\sigma \perp -a\tau)_{\mathrm{an}} < 2^n$ which, by Lemma 3.1, implies $\ln(\sigma, \tau) = n$, in other words $\sigma \simeq \tau$.

Conversely, if $\sigma \simeq \tau$ then the largest Pfister neighbor of τ contained in φ has dimension $2^n - 2^{m-1}$. That φ contains such a Pfister neighbor of this dimension follows readily from Remark 3.5(ii) (using the notation there, one may take $\alpha \otimes \sigma_1'$). On the other hand φ does not contain *any* Pfister neighbor of dimension $> 2^n - 2^{m-1}$. For suppose otherwise. Let $\chi \subset \varphi$ be such a Pfister neighbor of some $\tilde{\chi} \in P_nF$ with $\dim \chi > 2^n - 2^{m-1}$. Then $\tilde{\chi} \sim \chi \sim \varphi$, the first equivalence because χ is a Pfister neighbor of $\tilde{\chi}$, the second one by Corollary 3.8. But this is absurd since φ is clearly not a Pfister neighbor of $\tilde{\chi}$.

Corollary 7.4 *Let $\varphi \in P_{n,n-1}F$ and $\tau \in P_nF$. Then $(\varphi_{F(\tau)})_{\mathrm{an}}$ is defined over F.*

Proof. If $\varphi_{F(\tau)}$ stays anisotropic then there is nothing to show. So let us assume that $\varphi_{F(\tau)}$ is isotropic. By Proposition 5.1, we have that $\varphi_{F(\pi)} \simeq \sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. It follows immediately from Proposition 5.8(ii) that $\ln(\sigma, \tau) \geq n - 1$. The desired result follows now from Theorem 7.2. □

Statements (i) resp. (ii) of Theorem 7.2 essentially say that the obstruction to $(\varphi_{F(\tau)})_{\mathrm{an}}$ being defined over F is the non-existence of a Pfister neighbor of τ as a subform of φ, and by Corollary 7.4 this can only happen if $n \geq 3$ and $(n, m) \neq (n, n-1)$. This is not at all obvious as the case of the function field of a 2-fold Pfister form η shows. There are many examples of fields F with anisotropic forms $\psi \in WF$ and $\eta \in P_2F$ such that $\psi_{F(\eta)}$ is isotropic but ψ does not contain a Pfister neighbor of η (for such examples we refer to [LVG], [HLVG], [HVG]). However, since $F(\tau)/F$ is excellent we have, by definition of excellence, that $(\psi_{F(\eta)})_{\mathrm{an}}$ is defined over F. Conversely, if $\tau \in P_nF \geq 3$ such that $F(\tau)/F$ is not excellent then there might still be many forms ψ which contain Pfister neighbors of τ but where $(\psi_{F(\tau)})_{\mathrm{an}}$ is not defined over F. For example, let φ be such that $(\varphi_{F(\tau)})_{\mathrm{an}}$ is not defined over F and put $\psi \simeq \tau \perp \varphi$. Then $(\psi_{F(\tau)})_{\mathrm{an}} \simeq (\varphi_{F(\tau)})_{\mathrm{an}}$ is not defined over F, but ψ contains τ itself as a subform.

Twisted Pfister forms also yield new non-trivial examples of $F(\tau)$-minimal forms where $\tau \in P_nF$, $n \geq 3$. Recall that for a field extension K/F we say that φ is

K-minimal if φ is anisotropic, φ_K is isotropic, and if $\eta \subset \varphi$ with $\dim \eta < \dim \varphi$ then η_K is anisotropic. We are interested in the case where $K = F(\tau)$ for some anisotropic $\tau \in P_n F$. If $n = 1$ the K-minimal forms are exactly the scalar multiples of τ, cf. Proposition 2.1(iii). For $n = 2$, one can show that K-minimal forms are always of odd dimension ≥ 3. One can even construct a field F with some $\tau \in P_2 F$ such that to each odd integer $m \geq 3$ there exists a K-minimal form of dimension m, cf. [HVG].

Not much is known about K-minimal forms in the case $K = F(\tau)$ with $\tau \in P_n F$, $n \geq 3$. The following is known.

Theorem 7.5 ([H 3, Theorem 3], [H 2, Corollary 4.2].) *Let $\tau \in P_n F$ be anisotropic and $K = F(\tau)$.*

(i) *The K-minimal forms of dimension $\leq 2^{n-1}+1$ are exactly the Pfister neighbors of τ of dimension $2^{n-1}+1$.*

(ii) *If $n \leq 3$ then the K-minimal forms of dimension $\leq 2^{n-1}+2$ are exactly the Pfister neighbors of τ of dimension $2^{n-1}+1$. In particular, if $\varphi \in WF$ is anisotropic with $\dim \varphi \leq 2^{n-1}+2$, then φ_K is isotropic iff φ contains a Pfister neighbor of τ.*

In view of this result, Lemma 7.1 and Theorem 7.2, we get the following.

Proposition 7.6 *Let $n \geq 3$ and $1 \leq m \leq n-2$. Let $\varphi \in P_{n,m}F$ be defined by (σ, π) and let $\tau \in P_n F$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and $\ln(\sigma, \tau) \leq n-2$. Then each $F(\tau)$-minimal form χ contained in φ has dimension $2^{n-1}+2 \leq \dim \chi \leq 2^n - 2^{m-1}+1$. Moreover, if $(n,m) = (3,1)$ then $7 \leq \dim \chi \leq 8$.*

Proof. By Lemma 7.1, $i_W(\varphi_{F(\tau)}) = 2^{m-1}$. Thus, any subform of φ of dimension $2^n - 2^{m-1}+1$ becomes isotropic over $F(\tau)$. Hence, if $\chi \subset \varphi$ is $F(\tau)$-minimal we must necessarily have $\dim \chi \leq 2^n - 2^{m-1}+1$. We know by Theorem 7.2 that φ does not contain any Pfister neighbor of τ. Therefore, by Theorem 7.5(i), we must have $\dim \chi \geq 2^{n-1}+2$, and if $n = 3$ then Theorem 7.5(ii) even implies that $\dim \chi \geq 7$. □

In fact, Izhboldin [I] showed that with φ, σ, τ as in Proposition 7.6, if $m = 1$ and if $\ln(\sigma, \tau) = 1$ then φ itself is $F(\tau)$-minimal. This leads us to conjecture the following.

Conjecture 7.7 *Let $n \geq 3$ and $1 \leq m \leq n-2$. Let $\varphi \in P_{n,m}F$ be defined by (σ, π) and let $\tau \in P_n F$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and $\ln(\sigma, \tau) = m$. Let $\chi \subset \varphi$. Then χ is $F(\tau)$-minimal iff $\dim \chi = 2^n - 2^{m-1}+1$.*

Note that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ implies that $\ln(\sigma, \tau) \geq m$, cf. Proposition 5.8. In our conjecture, we require that the linkage of σ and τ is at the lower end, i.e., $\ln(\sigma, \tau) = m$. This will be needed in the proof of the conjecture in the case $m = 1$ and it is for this reason that we imposed this condition in the conjecture.

Proposition 7.8 (Izhboldin [I].) *Let $n \geq 3$. Let $\varphi \in P_{n,1}F$ be defined by (σ, π) and let $\tau \in P_n F$ such that $\tau_{F(\pi)} \simeq \sigma_{F(\pi)}$ and $\ln(\sigma, \tau) = 1$. Then φ is $F(\tau)$-minimal.*

Proof. $\varphi_{F(\tau)}$ is isotropic by Proposition 5.1. To prove that φ is $F(\tau)$-minimal it suffices to show that if $\eta \subset \varphi$ and $\dim \eta = 2^n - 1$, then $\eta_{F(\tau)}$ stays anisotropic.

By Lemma 6.10, there exists $\beta \in P_2F$ such that for $K = F(\beta)$ we have that σ_K is anisotropic and $\eta_K \subset \sigma_K$. Suppose $\eta_{F(\tau)}$ is isotropic. Then $\eta_{K(\tau)}$ is isotropic and hence, $\sigma_{K(\tau)}$ is isotropic and therefore hyperbolic. This implies that τ_K is similar and thus isometric to σ_K. Let $\psi \simeq (\sigma \perp -\tau)_{\mathrm{an}}$. Clearly, $\psi \in W(K/F)$ and hence there exists $\gamma \in WF$ such that $\psi \simeq \beta \otimes \gamma$. Now $\mathrm{ln}(\sigma, \tau) = 1$ and thus $\dim \psi = 2^{n+1} - 4$. Hence, $\dim \gamma = 2^{n-1} - 1$ is odd and one readily concludes that $\psi \equiv \beta \otimes \gamma \equiv \beta \not\equiv 0 \pmod{I^3F}$ as $\beta \in P_2F$ is anisotropic (if β were isotropic, the anisotropic form ψ would stay anisotropic over $K = F(\beta)$). But $\psi = \sigma - \tau \in I^nF$ with $n \geq 3$, a contradiction. □

Remark 7.9 The reason why this proof works so smoothly in the case $m = 1$ is that the field K from Lemma 6.10 is of a very nice form which just fits the situation. One would hope that with the field K from Lemma 6.10 one could give a similar proof also for $m > 1$. Consider the situation in Conjecture 7.7. To show that the conjecture is true it suffices to show that if $\eta \subset \varphi$ with $\dim \eta = 2^n - 2^{m-1}$ then $\eta_{F(\tau)}$ is anisotropic. One might want to proceed as in the proof above. Let K be as in Lemma 6.10 such that $\eta_K \subset \sigma_K$ and φ_K, σ_K, π_K stay anisotropic. The problem is to show that $\tau_K \not\simeq \sigma_K$. This worked for $m = 1$ because one can choose $K = F(\beta)$ for some $\beta \in P_2F$. If $m > 1$ then our construction in the proof of Lemma 6.10 generally leads to a field K for which it is not so clear why $\tau_K \not\simeq \sigma_K$ should hold.

8 Constructions of twisted Pfister forms

In this section we explicitly construct examples of $\varphi \in P_{n,m}F$ defined by (σ, π), and $\tau \in P_nF$ such that $\mathrm{ln}(\sigma, \tau) = k$ for some $m \leq k \leq n-2$ such that $\sigma_{F(\pi)} \simeq \tau_{F(\pi)}$. By Proposition 5.8 we then have $\mathrm{ln}(\tau, \pi) = m - 1$ and thus, we get a form $\psi \in P_{n,m}F$ defined by (τ, π) simply by putting $\psi \simeq (\tau \perp -\pi)_{\mathrm{an}}$. By Proposition 6.16 this shows the existence of $\varphi, \psi \in P_{n,m}F$ such that $\varphi \sim \psi$ but φ and ψ are neither neighbors nor conjugate, and by Theorem 7.2 it also shows the existence of $\varphi \in P_{n,m}F$ and $\tau \in P_nF$ such that $(\varphi_{F(\tau)})_{\mathrm{an}}$ is not defined over F. Note that by the symmetry of the situation we also have that $(\psi_{F(\sigma)})_{\mathrm{an}}$ is not defined over F. Hence, $F(\tau)/F$ and $F(\sigma)/F$ are both non-excellent field extensions.

In the first example, we will achieve this over a purely transcendental extension of the rationals $\mathbb{Q}$, and in the second example we will actually generalize Izhboldin's approach in [I].

Example 8.1 Let $1 \leq m \leq k \leq n-2$. To simplify notations, let $\ell = n-2-k$ so that $k+\ell+2 = n$. Let $F = \mathbb{Q}(x_1, \cdots, x_k, y_0, \cdots, y_\ell)$ be the rational function field in the $k+\ell+1 = n-1$ variables x_i and y_j over the rationals $\mathbb{Q}$. Let $p_0, \cdots, p_\ell$ be distinct prime numbers with $p_i \equiv 7 \pmod 8$. We now define Pfister forms $\sigma, \tau \in P_nF$ and $\pi \in P_mF$ as follows:

$$\begin{array}{rcl} \sigma & \simeq & \langle\langle 1, x_1, \cdots, x_k, y_0, \cdots, y_\ell \rangle\rangle \ ; \\ \tau & \simeq & \langle\langle 2, x_1, \cdots, x_k, p_0y_0, \cdots, p_\ell y_\ell \rangle\rangle \ ; \\ \pi & \simeq & \langle\langle x_1, \cdots, x_{m-1}, -x_m \rangle\rangle \ . \end{array}$$

One easily sees that σ, τ, and π are anisotropic (for instance by passing to the iterated power series field in the variables x_i, y_j, and then repeatedly applying Springer's theorem [L 1, Ch. 6, Proposition 1.9], [S, Ch. 6, Corollary 2.6(i)]).

Claim 1: $\ln(\sigma,\tau)=k$.
Proof.

$$\sigma\perp-\tau\simeq\langle\langle x_1,\cdots,x_k\rangle\rangle\otimes\underbrace{(\langle\langle 1,y_0,\cdots,y_\ell\rangle\rangle\perp-\langle\langle 2,p_0y_0,\cdots,p_\ell y_\ell\rangle\rangle)}_{\gamma}\ .$$

For $\emptyset\neq I\subset\{0,\cdots,\ell\}$ we define $Y_I=\prod_{i\in I}y_i$ and $P_I=\prod_{i\in I}p_i$. We find that

$$\gamma\simeq\langle 1,1,-1,-2\rangle\perp\mathop{\perp}_{\emptyset\neq I\subset\{0,\cdots,\ell\}}Y_I\langle 1,1,-P_I,-2P_I\rangle\ .$$

By Springer's theorem, we get

$$i_W(\gamma)=i_W(\langle 1,1,-1,-2\rangle)+\sum_{\emptyset\neq I\subset\{0,\cdots,\ell\}}i_W(\langle 1,1,-P_I,-2P_I\rangle)\ ,$$

where the Witt indices of the forms on the right hand side is computed over $\mathbb{Q}$. Now $i_W(\langle 1,1,-1,-2\rangle)=1$ as $\langle 1,1,-1,-2\rangle\simeq\mathbb{H}\perp\langle 1,-2\rangle$ and $\langle 1,-2\rangle$ is anisotropic over $\mathbb{Q}$. By passing to the local field $\mathbb{Q}_{p_i}$ for some $i\in I$, we get for the Legendre symbols $\left(\frac{-2}{p_i}\right)\neq 1$ and $\left(\frac{-1}{p_i}\right)\neq 1$ as $p_i\equiv 7\pmod 8$. Hence, $\langle 1,1\rangle$ and $\langle 1,2\rangle$ are anisotropic over $\mathbb{Q}_{p_i}$ and thus also $\langle 1,1,-P_I,-2P_I\rangle$ because p_i divides P_I exactly to the first power (note that all the p_j's in the product P_I are distinct !). Hence, $i_W(\langle 1,1,-P_I,-2P_I\rangle)=0$ and we have $i_W(\gamma)=1$. Again by Springer's theorem, we readily conclude that

$$i_W(\sigma\perp-\tau)=i_W(\langle\langle x_1,\cdots,x_k\rangle\rangle\otimes\gamma)=2^ki_W(\gamma)=2^k\ .$$

Hence, $\ln(\sigma,\tau)=k$.

Claim 2: $\ln(\sigma,\pi)=\ln(\tau,\pi)=m-1$. *In particular,* $\varphi\simeq(\sigma\perp-\pi)_{\mathrm{an}}$, $\psi\simeq(\tau\perp-\pi)_{\mathrm{an}}\in P_{n,m}F$.
Proof. This can be shown in a similar way as before.

$$\sigma\perp-\pi\simeq\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle\otimes(\langle\langle 1,x_m,\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle\perp-\langle\langle -x_m\rangle\rangle)\ ,$$

and by Springer's theorem we obtain that

$$\langle\langle 1,x_m,\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle\perp-\langle\langle -x_m\rangle\rangle\simeq\mathbb{H}\perp\langle\langle 1,x_m,\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle'\perp\langle x_m\rangle$$

has Witt index 1 as $\langle\langle 1,x_m,\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle'\perp\langle x_m\rangle$ is anisotropic (here, ρ' denotes the pure part of a Pfister form ρ), and that therefore

$$i_W(\sigma\perp-\pi)=\dim\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle=2^{m-1}$$

which in turn implies that $\ln(\sigma,\pi)=m-1$. A similar argument shows that $\ln(\tau,\pi)=m-1$ and we omit the details. It is now obvious that $\varphi\simeq(\sigma\perp-\pi)_{\mathrm{an}}$ and $\psi\simeq(\tau\perp-\pi)_{\mathrm{an}}$ have dimension 2^n and are in $P_{n,m}F$.

Claim 3: $\sigma_{F(\pi)}\simeq\tau_{F(\pi)}$.
Proof. Let $K=F(\pi)$. We have

$$0=\pi_K=\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle_K\perp-x_m\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle_K$$

and therefore $\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle_K \simeq x_m\langle\langle x_1,\cdots,x_{m-1}\rangle\rangle_K$. Hence,

$$\langle\langle x_1,\cdots,x_{m-1},x_m\rangle\rangle_K \simeq \langle\langle x_1,\cdots,x_{m-1},1\rangle\rangle_K \ .$$

Note also that $\langle\langle 1,2\rangle\rangle \simeq \langle\langle 1,1\rangle\rangle$ and that $\langle\langle 1,1\rangle\rangle \simeq a\langle\langle 1,1\rangle\rangle$ for any positive $a\in\dot{\mathbb{Q}}$. In particular, $\langle\langle 1,1,y_i\rangle\rangle \simeq \langle\langle 1,1,p_iy_i\rangle\rangle$. All this together yields

$$\begin{array}{rcl}\sigma_K & \simeq & \langle\langle 1,x_1,\cdots,x_{m-1},x_m,x_{m+1},\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle_K\\ & \simeq & \langle\langle 1,x_1,\cdots,x_{m-1},1,x_{m+1},\cdots,x_k,y_0,\cdots,y_\ell\rangle\rangle_K\\ & \simeq & \langle\langle 1,x_1,\cdots,x_{m-1},1,x_{m+1},\cdots,x_k,p_0y_0,\cdots,p_\ell y_\ell\rangle\rangle_K\\ & \simeq & \langle\langle 2,x_1,\cdots,x_{m-1},1,x_{m+1},\cdots,x_k,p_0y_0,\cdots,p_\ell y_\ell\rangle\rangle_K\\ & \simeq & \langle\langle 2,x_1,\cdots,x_{m-1},x_m,x_{m+1},\cdots,x_k,p_0y_0,\cdots,p_\ell y_\ell\rangle\rangle_K\\ & \simeq & \tau_K \ . \ \square\end{array}$$

This completes our first example. Similar examples can be constructed also if one replaces $\mathbb{Q}$ by any global field $\mathbb{K}$ of characteristic $\neq 2$ and with the iterated power series field $F=\mathbb{K}((x_1))\cdots((x_k))((y_0))\cdots((y_\ell))$. For the u-invariant of such a field we get $u(F)=2^{k+\ell}u(\mathbb{K})=4\cdot 2^{n-1}=2^{n+1}$. Thus, we can construct examples of (non-formally real) fields F with $u(F)=2^{n+1}$ and $\tau\in P_nF$ such that $F(\tau)/F$ is not excellent. If F is non-formally real and $u(F)<2^n+2^{n-1}$ then examples of the type we constructed above cannot exist. For in order for examples of this type to exist one needs two n-fold Pfister forms σ, τ such that $\mathrm{ln}(\sigma,\tau)\leq n-2$ or $\dim(\sigma\perp -\tau)_{\mathrm{an}}\geq 2^n+2^{n-1}$. So the question is: Are there (non-formally real) fields F with $u(F)<2^n+2^{n-1}$ such that there exists $\tau\in P_nF$ with $F(\tau)/F$ not excellent? In [H 5] it was shown that if F is linked (which implies that $u(F)\in\{0,1,2,4,8\}$) then $F(\tau)/F$ is excellent for all Pfister forms τ over F (here, F may be formally real or non-formally real). Among the many other results in [H 5] let us only mention that if τ is a Pfister form over a field F and if the Hasse number of F, $\tilde{u}(F)$, is ≤ 6 or if $\dim\tau\geq 2\tilde{u}(F)$, then $F(\tau)/F$ is excellent. This is of course mainly of interest in the case where F is formally real. For if F is non-formally real then there are no anisotropic forms of dimension $>\tilde{u}(F)$.

COROLLARY 8.2 *To each $n\geq 3$ there exists a field F such that there are anisotropic n-fold Pfister forms ρ, σ over F with $F(\rho)/F$ excellent and $F(\sigma)/F$ not excellent.*

Proof. We only show this for $n=3$ to keep the notations simple. Let $F=\mathbb{Q}((x))((y))$. The previous example shows that for $\sigma\simeq\langle\langle 1,x,y\rangle\rangle$ we have that $F(\sigma)/F$ is not excellent. Let $\rho\simeq\langle\langle 1,1,1\rangle\rangle$. Since ρ is defined over $\mathbb{Q}$, it is not hard to see that the field $E=F(\rho)=\mathbb{Q}((x))((y))(\rho)$ is contained in $L=\mathbb{Q}(\rho)((x))((y))$. Let ψ be an anisotropic form over F. By Springer's theorem, we can write $\psi\simeq\psi_0\perp x\psi_1\perp y\psi_2\perp xy\psi_3$ where the ψ_i are forms over $\mathbb{Q}$ which are uniquely determined up to isometry over $\mathbb{Q}$. Let $K=\mathbb{Q}(\rho)\subset E$. It is known that function fields of Pfister forms over global fields are always excellent (cf. [ELW 2], [H 5], see also the remarks preceding this corollary). Hence, there are forms μ_i defined over $\mathbb{Q}$ such that $(\mu_i)_K\simeq((\psi_i)_K)_{\mathrm{an}}$. In WE we obviously have

$$(\psi_0)_E\perp x(\psi_1)_E\perp y(\psi_2)_E\perp xy(\psi_3)_E=(\mu_0)_E\perp x(\mu_1)_E\perp y(\mu_2)_E\perp xy(\mu_3)_E.$$

The right hand side is defined over F by $\mu_0\perp x\mu_1\perp y\mu_2\perp xy\mu_3$. To show excellence, it remains to show that $(\mu_0)_E\perp x(\mu_1)_E\perp y(\mu_2)_E\perp xy(\mu_3)_E$ is anisotropic. Indeed, this form is anisotropic over the bigger field L. This is because $L = K((x))((y))$, the $(\mu_i)_K$ are anisotropic and by Springer's theorem we have that $(\mu_0)_L \perp x(\mu_1)_L \perp y(\mu_2)_L \perp xy(\mu_3)_L$ is anisotropic. Thus, we have shown that $(\psi_{F(\rho)})_{\mathrm{an}}$ is defined over F by $\mu_0 \perp x\mu_1 \perp y\mu_2 \perp xy\mu_3$, which in turn proves the excellence of $F(\rho)/F$. □

The next example generalizes Izhboldin's construction in [I] which was carried out only in the case $m = \ln(\sigma, \tau) = 1$. However, we will show that his arguments can be applied, after some minor modifications, to the more general situation where m can be any positive integer $\leq n-2$ and $\ln(\sigma, \tau) = k$ can be any integer with $m \leq k \leq n-2$.

EXAMPLE 8.3 Let $n \geq 3$ and let $1 \leq m \leq k \leq n-2$. Let $\tau \in P_n F$ be anisotropic. We will construct a unirational field extension E/F such that there exist anisotropic forms $\pi \in P_m E$ and $\sigma \in P_n E$ with $\ln(\sigma, \pi) = \ln(\tau_E, \pi) = m-1$ and $\ln(\sigma, \tau_E) = k$ (note that τ will stay anisotropic over any unirational field extension), and furthermore $\sigma_{E(\pi)} \simeq \tau_{E(\pi)}$. We then get a form $\varphi \simeq (\sigma \perp -\pi)_{\mathrm{an}} \in P_{n,m}E$, and by Theorem 7.2 we have that $(\varphi_{E(\tau)})_{\mathrm{an}}$ is not defined over E. In particular, $E(\tau)/E$ is not excellent. Our construction will involve various field extensions of F. Their relations among each other are shown in a diagram below.

As for the construction of E, let this time $\ell = n - k \geq 2$ and write

$$\tau \simeq \langle\langle a_1, \cdots, a_k, b_1, \cdots, b_\ell \rangle\rangle$$

for suitable $a_i, b_j \in \dot{F}$. Let $F_0 = F(y_1, \cdots, y_\ell)$ and $F_1 = F_0(x) = F(x, y_1, \cdots, y_\ell)$ be rational function fields in the variables $x, y_1, \cdots, y_\ell$ over F. Let

$$\begin{array}{rcl} \sigma & \simeq & \langle\langle a_1, \cdots, a_k, y_1, \cdots, y_\ell \rangle\rangle \in P_n F_0 \\ \pi & \simeq & \langle\langle x, a_2, \cdots, a_m \rangle\rangle \in P_m F_1 \end{array}$$

After passing to the iterated power series field in the variables $x, y_1, \cdots, y_\ell$ and by repeatedly applying Springer's theorem, one readily checks that τ_{F_1}, σ_{F_1}, π_{F_1} are anisotropic and that $\ln(\sigma_{F_1}, \tau_{F_1}) = k$ and $\ln(\sigma_{F_1}, \pi_{F_1}) = \ln(\tau_{F_1}, \pi_{F_1}) = m-1$. We leave the details to the reader. The aim is to construct E/F_1 such that $\sigma_{E(\pi)} \simeq \tau_{E(\pi)}$, such that this form and π_E stay anisotropic and such that $\ln(\sigma_E, \tau_E) = k$. Note that we will have $m \geq \ln(\sigma_E, \pi_E) \geq \ln(\sigma_{F_1}, \pi_{F_1}) = m-1$. Now $\ln(\sigma_E, \pi_E) = m$ implies that π_E divides σ_E and thus $\sigma_{E(\pi)} = 0$, a contradiction to its anisotropy. Hence, we will still have $\ln(\sigma_E, \pi_E) = m-1$ and similarly $\ln(\tau_E, \pi_E) = m-1$.

To get this field E, we first define the following forms over F_1 which again are easily seen to be anisotropic:

$$\begin{array}{rcl} \tilde{\tau} & \simeq & \langle\langle x, a_2, \cdots, a_k, b_1, \cdots, b_\ell \rangle\rangle \\ \tilde{\sigma} & \simeq & \langle\langle x, a_2, \cdots, a_k, y_1, \cdots, y_\ell \rangle\rangle . \end{array}$$

Let E be the generic splitting field of the anisotropic form defined by $(\sigma - \tau) - (\tilde{\sigma} - \tilde{\tau})$ in WF_1. Then, in WE, $(\sigma - \tau)_E - (\tilde{\sigma} - \tilde{\tau})_E = 0$ or $(\sigma - \tau)_E = (\tilde{\sigma} - \tilde{\tau})_E$. As π divides both $\tilde{\sigma}$ and $\tilde{\tau}$, we get that $\tilde{\sigma}_{E(\pi)} = \tilde{\tau}_{E(\pi)} = 0$, hence, $(\sigma - \tau)_{E(\pi)} = 0$, i.e., $\sigma_{E(\pi)} \simeq \tau_{E(\pi)}$.

We first show that $\sigma_{E(\pi)} \simeq \tau_{E(\pi)}$ is anisotropic. Let $F_2 = F_1(\sqrt{-x})$. Then F_2/F_0 is purely transcendental and thus σ_{F_2} and τ_{F_2} stay anisotropic and we still have $\ln(\sigma_{F_2}, \tau_{F_2}) = \ln(\sigma_{F_0}, \tau_{F_0}) = k$. Furthermore, $\langle\langle x \rangle\rangle_{F_2} = \langle\langle -1 \rangle\rangle_{F_2} = 0$ and hence

$\pi_{F_2} = \tilde{\sigma}_{F_2} = \tilde{\tau}_{F_2} = 0$ in WF_2. Let K be the generic splitting field over F_2 of $(\sigma \perp -\tau)_{F_2}$. Clearly, $(\sigma - \tau)_K = 0$ in WK, i.e., $\sigma_K \simeq \tau_K$. We claim that $\sigma_K \simeq \tau_K$ is anisotropic. Now $\ln(\sigma_{F_2}, \tau_{F_2}) = k < n$ and thus $\sigma_{F_2} \not\simeq \tau_{F_2}$, i.e., $(\sigma \perp -\tau)_{F_2} \neq 0$. Clearly, $\deg(\sigma \perp -\tau)_{F_2} \geq n$. Suppose that $\sigma_K \simeq \tau_K = 0$. Then by [AK, Satz 20] it follows that $\deg(\sigma \perp -\tau)_{F_2} = n$ and $(\sigma \perp -\tau)_{F_2} \equiv \sigma_{F_2} \pmod{J_{n+1}F_2}$. Hence, $-\tau_{F_2} \equiv 0 \pmod{J_{n+1}F_2}$ and the Arason-Pfister Hauptsatz implies that $\tau_{F_2} = 0$, a contradiction.

Obviously, $\tilde{\sigma}_{K(\pi)} = \tilde{\tau}_{K(\pi)} = 0$. Thus, $(\sigma - \tau)_{K(\pi)} - (\tilde{\sigma} - \tilde{\tau})_{K(\pi)} = 0$. Since E/F_1 is the generic splitting field of $(\sigma - \tau) - (\tilde{\sigma} - \tilde{\tau}) \in WF_1$, we have by Proposition 2.2(iii) that $E\cdot K(\pi)/K(\pi)$ is purely transcendental. But $K(\pi)/K$ is purely transcendental as well because $\pi_K = 0$. Hence, $E\cdot K(\pi)/K$ is purely transcendental and therefore $\sigma_{E\cdot K(\pi)} \simeq \tau_{E\cdot K(\pi)}$ is anisotropic because $\sigma_K \simeq \tau_K$ is anisotropic. Since $E(\pi) \subset E\cdot K(\pi)$ we have that $\sigma_{E(\pi)} \simeq \tau_{E(\pi)}$ is anisotropic.

Let now $F_3 = F_1(\sqrt{a_1 X})$. Again, we clearly have that F_3/F_0 is purely transcendental. Furthermore, $\sigma_{F_3} \simeq \tilde{\sigma}_{F_3}$ and $\tau_{F_3} \simeq \tilde{\tau}_{F_3}$ as $a_1 = X$ in $\dot{F}_3/\dot{F}_3^2$. Hence, $(\sigma - \tau)_{F_3} - (\tilde{\sigma} - \tilde{\tau})_{F_3} = 0$ in WF_3 and we have that $E\cdot F_3/F_3$ is purely transcendental by the same reason as before. Hence, $E\cdot F_3/F_0$ is purely transcendental as well and thus, since $F_0 \subset E \subset E\cdot F_3$, we conclude that E/F_0 is unirational. Therefore, $\ln(\sigma_E, \tau_E) = \ln(\sigma_{F_0}, \tau_{F_0}) = k$ as desired. Obviously, E/F is also unirational as F_0/F is purely transcendental. This completes Izhboldin's construction. $\square$

$$\begin{tikzcd}
 & E\cdot K(\pi) \arrow[dd, dash] \arrow[dr, dash] & \\
 & & K(\pi) \arrow[dd, dash] \\
 & E(\pi) \arrow[dd, dash] & \\
E\cdot F_3 \arrow[dd, dash] \arrow[dr, dash] & & K \arrow[dd, dash] \\
 & E \arrow[dd, dash] & \\
F_3 = F_1(\sqrt{a_1 x}) \arrow[dr, dash] & & F_2 = F_1(\sqrt{-x}) \arrow[dl, dash] \\
 & F_1 = F_0(x) \arrow[d, dash] & \\
 & F_0 = F(y_1, \cdots, y_\ell) \arrow[d, dash] & \\
 & F &
\end{tikzcd}$$

COROLLARY 8.4 (Izhboldin [I].) *Let $\tau \in P_nF$ be anisotropic, $n \geq 3$. Then there exists a unirational field extension E/F such that $E(\tau)/E$ is not excellent.*

References

[AK] Arason, J.Kr.; Knebusch, M.: *Über die Grade quadratischer Formen.* Math. Ann. 234 (1978), 167–192.

[AP] Arason, J.Kr.; Pfister, A.: *Beweis des Krullschen Durschnittsatzes für den Wittring.* Invent. Math. 12 (1971), 173–176.

[EL] Elman, R.; Lam, T.Y.: *Pfister forms and K-theory of fields.* J. Algebra 23 (1972), 181–213.

[ELW 1] Elman, R.; Lam, T.Y.; Wadsworth, A.R.: *Amenable fields and Pfister extensions.* Proc. of Quadratic Forms Conference (ed. G. Orzech). Queen's Papers in Pure and Applied Mathematics No. 46 (1977), 445–491.

[ELW 2] Elman, R.; Lam, T.Y.; Wadsworth, A.R.: *Function fields of Pfister forms.* Invent. Math. 51 (1979), 61–75.

[F 1] Fitzgerald, R.W.: *Witt kernels of function field extensions.* Pacific J. Math. 109 (1983), 89–106.

[F 2] Fitzgerald, R.W.: *Quadratic forms of height two.* Trans. Amer. Math. Soc. 283 (1984), 339–351.

[H 1] Hoffmann, D.W.: *Isotropy of* 5*-dimensional quadratic forms isotropic over the function field of a quadric.* K-Theory and Algebraic Geometry: Connections with Quadratic Forms and Division Algebras, Proceedings of the 1992 Santa Barbara Summer Research Institute (eds. B. Jacob, A. Rosenberg). Proc. Symp. Pure Math. 58.2 (1995), 217–225.

[H 2] Hoffmann, D.W.: *On* 6*-dimensional quadratic forms isotropic over the function field of a quadric.* Comm. Alg. 22 (1994), 1999–2014.

[H 3] Hoffmann, D.W.: *Isotropy of quadratic forms over the function field of a quadric.* Math. Z. 220 (1995), 461–476.

[H 4] Hoffmann, D.W.: *On quadratic forms of height* 2 *and a theorem of Wadsworth.* To appear in: Trans. Amer. Math. Soc.

[H 5] Hoffmann, D.W.: *Amenability and excellence of function fields of Pfister forms.* Preprint 1995.

[H 6] Hoffmann, D.W.: *Splitting of quadratic forms, I: Forms of small dimension.* Preprint 1995.

[H 7] Hoffmann, D.W.: *On the dimensions of anisotropic quadratic forms with trivial invariants.* SFB 343, Universität Bielefeld, Preprint no. 88 (1995).

[HLVG] Hoffmann, D.W.; Lewis, D.W.; Van Geel, J.: *Minimal forms for function fields of conics.* K-Theory and Algebraic Geometry: Connections with Quadratic Forms and Division Algebras, Proceedings of the 1992 Santa Barbara Summer Research Institute (eds. B. Jacob, A. Rosenberg). Proc. Symp. Pure Math. 58.2 (1995), 227–237.

[HVG] HOFFMANN, D.W.; VAN GEEL, J.: *Minimal forms with respect to function fields of conics*. Manuscripta Math. 86 (1995), 23–48.

[HuR] HURRELBRINK, J.; REHMANN, U.: *Splitting patterns of quadratic forms*. Math. Nachr. 176 (1995), 111–127.

[I] IZHBOLDIN, O.T.: *On the nonexcellence of field extensions* $F(\pi)/F$. Doc. Math. J. DMV 1 (1996), 127–136 (cf. also Preprint 1995).

[Ka] KAHN, B.: *Formes quadratiques de hauteur et de degré* 2. To appear in: Indag. Math.

[K 1] KNEBUSCH, M.: *Generic splitting of quadratic forms* I. Proc. London Math. Soc. 33 (1976), 65–93.

[K 2] KNEBUSCH, M.: *Generic splitting of quadratic forms* II. Proc. London Math. Soc. 34 (1977), 1–31.

[Lag] LAGHRIBI, A.: *Isotropie de certaines formes quadratiques de dimension* 7 *et* 8 *sur le corps des fonctions d'une quadrique*. To appear in: Duke Math. J.

[L 1] LAM, T.Y.: The Algebraic Theory of Quadratic Forms. Reading, Massachusetts: Benjamin 1973 (revised printing 1980).

[L 2] LAM, T.Y.: *Ten lectures on quadratic forms over fields*. Proc. of Quadratic Forms Conference (ed. G. Orzech). Queen's Papers in Pure and Applied Mathematics No. 46 (1977), 1–102.

[LVG] LEWIS, D.W.; VAN GEEL, J.: *Quadratic forms isotropic over the function field of a conic*. Indag. Math. 5 (1994), 325–339.

[M] 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.

[P] PFISTER, A.: *Quadratische Formen in beliebigen Körpern*. Invent. Math. 1 (1966), 116–132.

[S] SCHARLAU, W.: Quadratic and Hermitian Forms. Berlin, Heidelberg, New York, Tokyo: Springer 1985.

[W] WADSWORTH, A.R.: *Similarity of quadratic forms and isomorphisms of their function fields*. Trans. Amer. Math. Soc. 208 (1975), 352–358.

Detlev W. Hoffmann
Laboratoire de Mathématiques
Faculté des Sciences
Université de Franche-Comté
25030 Besançon Cedex
France
detlev@math.univ-fcomte.fr

The Polytope of All Triangulations of a Point Configuration

Jesús A. de Loera[1], Serkan Hoşten[1],
Francisco Santos[2], Bernd Sturmfels[3]

Received: December 9, 1995
Revised: March 4, 1996

Communicated by Günter M. Ziegler

Abstract. We study the convex hull $P_{\mathcal{A}}$ of the 0-1 incidence vectors of all triangulations of a point configuration $\mathcal{A}$. This was called the universal polytope in [4]. The affine span of $P_{\mathcal{A}}$ is described in terms of the co-circuits of the oriented matroid of $\mathcal{A}$. Its intersection with the positive orthant is a quasi-integral polytope $Q_{\mathcal{A}}$ whose integral hull equals $P_{\mathcal{A}}$. We present the smallest example where $Q_{\mathcal{A}}$ and $P_{\mathcal{A}}$ differ. The duality theory for regular triangulations in [5] is extended to cover all triangulations. We discuss potential applications to enumeration and optimization problems regarding all triangulations.

1991 Mathematics Subject Classification: Primary 52B55; Secondary 90C27.

1 Introduction

We are interested in the set of all triangulations of a configuration $\mathcal{A} = \{a_1, \ldots, a_n\} \subset \mathbf{R}^d$. The subset of *regular* triangulations is well-understood thanks to its bijection with the vertices of the *secondary polytope* (see [7, Chapter 7] and [16, Lecture 9]). But non-regular triangulations remain a mystery: for instance, it is still unknown whether any two triangulations of $\mathcal{A}$ can be connected by a sequence of bistellar flips [12]. Non-regular triangulations are abundant: if $\mathcal{A}$ is the vertex set of the cyclic polytope $C_{4n-4}(4n)$, there are at least 2^n triangulations (Proposition 5.10 in this article), while the number of regular triangulations is $O(n^4)$.

[1]Supported in part by the National Science Foundation
[2]Supported in part by the Spanish Dirección General de Investigacion Científica y Técnica
[3]Supported in part by the David and Lucille Packard Foundation and the National Science Foundation

One approach in understanding non-regular triangulations is to replace the secondary polytope by a larger polytope $P_{\mathcal{A}}$ whose vertices are in bijection with all triangulations of $\mathcal{A}$. The polytope $P_{\mathcal{A}}$ is isomorphic to the *universal polytope* introduced by Billera, Filliman and Sturmfels [4]. They expressed the secondary polytope as a projection of $P_{\mathcal{A}}$ and they showed $dim(P_{\mathcal{A}}) = \binom{n-1}{d+1}$ when $\mathcal{A}$ is in general position. We shall now fix some notation, and define the polytope $P_{\mathcal{A}}$.

Throughout this paper $\mathcal{A} \subset \mathbf{R}^d$ will denote a d-dimensional configuration of n possibly repeated points. By a *k-simplex* we mean a sub-configuration of $\mathcal{A}$ consisting of $k+1$ affinely independent points. A *triangulation* of $\mathcal{A}$ is a collection T of d-simplices whose convex hulls cover $conv(\mathcal{A})$ and intersect properly: for any σ and τ in T we have $conv(\sigma \cap \tau) = conv(\sigma) \cap conv(\tau)$. Let $\Delta(\mathcal{A})$ denote the collection of d-simplices in $\mathcal{A}$. We define $P_{\mathcal{A}}$ as the convex hull in $\mathbf{R}^{\Delta(\mathcal{A})}$ of the set of incidence vectors of all triangulations of $\mathcal{A}$. For a triangulation T the *incidence vector* v_T has coordinates $(v_T)_\sigma = 1$ if $\sigma \in T$ and $(v_T)_\sigma = 0$ if $\sigma \notin T$. We also consider the polytope $Q_{\mathcal{A}} = \mathit{aff}(P_{\mathcal{A}}) \cap \mathbf{R}_+^{\Delta(\mathcal{A})}$, which is the linear programming relaxation of $P_{\mathcal{A}}$. We denote by $M(\mathcal{A})$ the oriented matroid of affine dependencies of the point configuration $\mathcal{A}$.

We first present linear equations defining the affine hull $\mathit{aff}(P_{\mathcal{A}})$ of $P_{\mathcal{A}}$. These equations involve the cocircuits (see [6, Chapter 1] or [16, Lecture 6]) of $M(\mathcal{A})$: for any τ which is a $(d-1)$-simplex of $\mathcal{A}$, let H_τ be the hyperplane that contains τ and let H_τ^+ and H_τ^- denote the two open half-spaces defined by H_τ. We recall that the cocircuits of the oriented matroid $M(\mathcal{A})$ are the resulting partitions $(\mathcal{A} \cap H_\tau^+, \mathcal{A} \cap H_\tau, \mathcal{A} \cap H_\tau^-)$ of $\mathcal{A}$. Consider the following linear form:

$$Co_\tau \quad := \sum_{\sigma=\tau\cup\{a\},\ a\in\mathcal{A}\cap H_\tau^+} x_\sigma \quad - \sum_{\sigma=\tau\cup\{a\},\ a\in\mathcal{A}\cap H_\tau^-} x_\sigma \tag{1}$$

We call Co_τ the *cocircuit form* associated with the $(d-1)$-simplex τ. If $conv(\tau) \cap int(conv(\mathcal{A})) \neq \emptyset$, we say that τ is an *interior* $(d-1)$-simplex. In this case neither of the two sums in (1) is void. Moreover, every triangulation T of $\mathcal{A}$ contains either no d-simplex containing τ or exactly two, one in the first sum and one in the second. Thus Co_τ vanishes at the incidence vector v_T of every triangulation of $\mathcal{A}$, and hence, on $\mathit{aff}(P_{\mathcal{A}})$. We call the equations $Co_\tau = 0$, for interior $(d-1)$-simplices τ, the *interior cocircuit equations.* We summarize our main results:

Theorem 1.1 *Let $\mathcal{A}$ be a point configuration with the above conventions.*

(i) *The affine span of $P_{\mathcal{A}}$ in $\mathbf{R}^{\Delta(\mathcal{A})}$ is defined by the linear equations $Co_\tau = 0$ for every interior $(d-1)$-simplex τ, together with one non-homogeneous linear equation valid on $P_{\mathcal{A}}$.*

(ii) *$P_{\mathcal{A}}$ coincides with the integral hull of $Q_{\mathcal{A}}$; i.e., the lattice points in $Q_{\mathcal{A}}$ are precisely the incidence vectors of triangulations of $\mathcal{A}$.*

(iii) *Two triangulations T_1 and T_2 of $\mathcal{A}$ are neighbors in the edge graph of $P_{\mathcal{A}}$ if and only if they are neighbors in the edge graph of $Q_{\mathcal{A}}$.*

(iv) *For the case of the n-gon and configurations with at most $d+3$ points, we have $Q_{\mathcal{A}} = P_{\mathcal{A}}$. This is not true in general for $n \geq d+4 \geq 6$.*

The following three types of non-homogeneous equations may be used to complete the description of *aff*($P_{\mathcal{A}}$) in part (i) of Theorem 1.1: If τ is a non-interior $(d-1)$-simplex and $conv(\tau)$ is a facet of $conv(\mathcal{A})$, then the cocircuit form Co_τ has constant value equal to ± 1 on $P_{\mathcal{A}}$. This produces new valid equations for *aff*($P_{\mathcal{A}}$) which we call *boundary cocircuit equations*. They can be expressed in the form

$$\sum_{\sigma=\tau\cup\{a\},\ a\in\mathcal{A}\setminus\tau} x_\sigma \quad = \quad 1. \tag{2}$$

Another set of valid equations for *aff*($P_{\mathcal{A}}$) can be obtained as follows: let $p \in conv(\mathcal{A})$ be a point not lying in the convex hull of any $(d-1)$-simplex of $\mathcal{A}$. Every triangulation of $\mathcal{A}$ satisfies the equation:

$$\sum_{\sigma\in\Delta(\mathcal{A}),\ p\in conv(\sigma)} x_\sigma \;=\; 1. \tag{3}$$

Recall that the *chamber complex* of $\mathcal{A}$ is the common refinement of all triangulations of $\mathcal{A}$ (see [1],[5]). We call the equations of type (3) *chamber equations*, because the simplices in the sum only depend on the chamber in which p lies. Note that the boundary cocircuit equations (2) are a particular case of chamber equations.

Finally, if we denote by $vol(\cdot)$ the standard volume form on $\mathbf{R}^d$, the following *volume equation* is satisfied by every triangulation of $\mathcal{A}$:

$$\sum_{\sigma\in\Delta(\mathcal{A})} vol(conv(\sigma))\, x_\sigma \quad = \quad vol(conv(\mathcal{A})). \tag{4}$$

Remark 1.2 The (interior and boundary) cocircuit equations depend only on the oriented matroid $M(\mathcal{A})$ of affine dependencies of $\mathcal{A}$. This holds neither for the volume equation nor for chamber equations: for example, all configurations consisting on the six vertices of a convex planar hexagon have the same oriented matroid, while the number of chambers can be 24 or 25, depending on the coordinates of the vertices.

Clearly, if $\mathcal{A}$ has no simplicial facets, then there are no boundary cocircuit equations. However, every configuration $\mathcal{A}$ has *some* chamber equations which can be obtained from the oriented matroid $M(\mathcal{A})$. (Such chambers arise from *lexicographic extensions*; see [6, Figure 7.2.2, page 296].) Any of them, together with the interior cocircuit equations, will provide a description of *aff*($P_{\mathcal{A}}$) in terms of $M(\mathcal{A})$. Part (ii) of Theorem 1.1 implies that this yields a description of $P_{\mathcal{A}}$ itself in terms of $M(\mathcal{A})$.

In Section 2 we examine the affine span of $P_{\mathcal{A}}$ and we prove part (i) of Theorem 1.1. A surprising consequence (Corollary 2.3) is that *aff*($P_{\mathcal{A}}$) is spanned by the *regular* triangulations only. This implies the formula $dim(P_{\mathcal{A}}) = \binom{n-1}{d+1}$ when $\mathcal{A}$ is in general position. Section 3 contains the proof of parts (ii) and (iii) in Theorem 1.1. As a consequence of part (iii) we obtain a combinatorial characterization of the edges of $P_{\mathcal{A}}$ (Theorem 3.3). Section 4 contains the proof of part (iv). We also discuss computational issues regarding the enumeration of triangulations and the optimization of linear cost functions over $P_{\mathcal{A}}$. In Section 5 we present a duality theory relating (non-regular) triangulations of $\mathcal{A}$ with (virtual) chambers in the Gale transform of $\mathcal{A}$.

2 Equations defining the affine span of $P_{\mathcal{A}}$

We introduce now some basic definitions and properties concerning regular triangulations. For a more detailed description and the relevant background the reader may consult [4],[5],[7, Chapter 7],[9] and [16].

A regular triangulation of $\mathcal{A}$ is a triangulation which is obtained by projecting the lower envelope of a $(d+1)$-dimensional simplicial polytope onto $conv(\mathcal{A})$. In other words, a triangulation is regular if it supports a piecewise convex linear functional.

In [7, Chapter 7] the collection of regular triangulations of a point configuration is identified with the vertex set of a polytope $\Sigma(\mathcal{A})$ of dimension $n-d-1$ embedded in $\mathbf{R}^n$. This polytope, called the *secondary polytope*, is a projection of $P_{\mathcal{A}}$ [4]. The projection map $\pi : \mathbf{R}^{\Delta(\mathcal{A})} \to \mathbf{R}^{\mathcal{A}}$ is given by $\pi(e_\sigma) = vol(\sigma)\sum_{a\in\sigma} e_a$, where e_σ and e_a denote the standard basis vectors.

A characterization of the edges of $\Sigma(\mathcal{A})$ is given in [7, Chapter 7]. This uses the notions of *circuits* and *bistellar flips*. We only define bistellar flips in the general position case: a *circuit* of the point configuration $\mathcal{A}$ is a minimal affinely dependent set. If $\mathcal{A}$ is in general position circuits are subsets of cardinality $d+2$. The unique (up to scaling factor) affine dependency equation satisfied by a circuit Z splits it into two subsets Z^+ and Z^- consisting of the points which have positive and negative coefficients respectively. Any circuit Z has exactly two triangulations $t(Z)^+ = \{Z \setminus \{a\}, \quad a \in Z^+\}$ and $t(Z)^- = \{Z\setminus\{a\}, \quad a \in Z^-\}$. If a triangulation T of $\mathcal{A}$ contains one of the two triangulations of a circuit Z (say $t(Z)^+$), then $T' = T \setminus t(Z)^+ \cup\ t(Z)^-$ is again a triangulation of $\mathcal{A}$. The operation that passes from T to T' (or vice versa) is called a *bistellar flip.* Two regular triangulations are neighbors in the 1-skeleton of the secondary polytope $\Sigma(\mathcal{A})$ if and only if they differ by a bistellar flip. This implies that any two regular triangulations can be transformed to one another by a finite sequence of bistellar flips. It is unknown whether this property is true for non-regular triangulations.

Our next goal is to prove part (i) of Theorem 1.1. We first state a lemma about the behavior of triangulations under the matroidal operations of *deletion* $\mathcal{A} \mapsto \mathcal{A}\backslash a_i$ and *contraction* $\mathcal{A} \mapsto \mathcal{A}/a_i$, where a_i is a vertex of $\mathcal{A}$. We regard $\mathcal{A}/a_i$ as a configuration of typically $n-1$ points (maybe less, if a_i was a repeated point) in affine $(d-1)$-space. The convex hull of $\mathcal{A}/a_i$ is the *vertex figure* of $conv(\mathcal{A})$ at a_i.

Lemma 2.1 *Let $a_1 \in \mathcal{A}$ be a vertex of $conv(\mathcal{A})$. Every triangulation of $\mathcal{A}\backslash a_1$ can be extended to a triangulation of $\mathcal{A}$. Every* regular *triangulation of $\mathcal{A}/a_1$ can be extended to a triangulation of $\mathcal{A}$. The latter fails for non-regular triangulations.*

Proof: Every (regular or non-regular) triangulation T' of $\mathcal{A}\backslash a_1$ can be extended to a triangulation T of $\mathcal{A}$ by the *placing* operation described on page 444 in [9]. In this situation T is regular if and only if T' is regular.

The extension property of regular triangulations for contractions follows from the identification of the secondary fan of $\mathcal{A}$ with the chamber complex of $\mathcal{B}$, where $\mathcal{B}$ is a Gale transform of $\mathcal{A}$ (see [5], in particular Lemma 3.2 and the paragraph after Lemma 3.4). The reasoning is this: let T' be a regular triangulation of $\mathcal{A}/a_1$. It corresponds to a chamber $C_{T'}$ of $\mathcal{B}\backslash b_1$, where $b_1 \in \mathcal{B}$ is the point corresponding to a_1 in the Gale transform. The chamber $C_{T'}$ may split into smaller chambers when passing to the chamber complex of $\mathcal{B}$ and any such smaller chamber C_T corresponds to a triangulation T of $\mathcal{A}$ with $T/a_1 = T'$.

To see that regularity is necessary in the previous paragraph, let $\mathcal{A} \subset \mathbf{R}^3$ be the configuration $a_1 = (20,0,1)$, $a_2 = (1,20,0)$, $a_3 = (0,1,20)$, $a_4 = (10,0,0)$, $a_5 = (0,10,0)$, $a_6 = (0,0,10)$ and $a_7 = (-10,-10,-10)$. Let T' be the triangulation of $\mathcal{A}/a_7$ given by the triangles $\{1,2,5\}$, $\{1,3,4\}$, $\{1,4,5\}$, $\{2,3,6\}$, $\{2,5,6\}$, $\{3,4,6\}$ and $\{4,5,6\}$. There is no triangulation T of $\mathcal{A}$ such that $T/a_7 = T'$. The non-convex polyhedron with vertices $a_1,\dots,a_6$ and triangular faces determined by the above list, together with the triangle $\{1,2,3\}$, is the *Schönhardt polyhedron* [13] which cannot be triangulated without adding a new point. $\square$

Part (i) of Theorem 1.1 is a consequence of the following theorem and the existence of non-homogeneous forms vanishing on $P_{\mathcal{A}}$.

THEOREM 2.2 *Let $h = \sum_{\sigma\in\Delta(\mathcal{A})} c_\sigma x_\sigma$ ($c_\sigma \in \mathbf{R}$) be any homogeneous linear form on $\mathbf{R}^{\Delta(\mathcal{A})}$. The following properties are equivalent:*

(i) *h is a linear combination of the interior cocircuit forms Co_τ* (1).

(ii) *h vanishes on (the incidence vector of) every triangulation of $\mathcal{A}$.*

(iii) *h vanishes on (the incidence vector of) every regular triangulation of $\mathcal{A}$.*

Proof: (i)$\Rightarrow$(ii)$\Rightarrow$(iii) are obvious. We prove (iii)$\Rightarrow$(i).

Let $P_{\mathcal{A}}^{reg}$ denote the convex hull of all points v_T where T is a regular triangulation of $\mathcal{A}$. Thus $P_{\mathcal{A}}^{reg} \subset P_{\mathcal{A}}$. Let $h = \sum c_\sigma x_\sigma$ be any linear form which vanishes on $P_{\mathcal{A}}^{reg}$. We shall prove that h is a linear combination of the interior cocircuit forms using a double induction on $n = |\mathcal{A}|$ and $d = dim(\mathcal{A})$. Assume that the statement is true for any configuration of smaller cardinality or smaller dimension.

Let a_1 be a vertex of $conv(\mathcal{A})$. Let us suppose that $\mathcal{A}\backslash a_1$ still spans $\mathbf{R}^d$. Otherwise $P_{\mathcal{A}}^{reg}$ and $P_{\mathcal{A}\backslash a_1}^{reg}$ are affinely isomorphic and the theorem follows by induction. The interior cocircuit forms Co_τ vanish on $P_{\mathcal{A}}^{reg}$. If $a_1 \notin \tau$ then Co_τ involves at most one d-simplex of the form $\sigma = \{a_1\}\cup\tau$. Subtracting appropriate multiples of those Co_τ from h we get another linear form h_1 in which the variables x_σ corresponding to these simplices do not appear. That is,

$$h_1 \quad = \sum_{\substack{\sigma: a_1\in\sigma \\ conv(\sigma\backslash a_1)\subset\, boundary(conv(\mathcal{A}))}} c_\sigma x_\sigma \quad + \sum_{\sigma: a_1\notin\sigma} c'_\sigma x_\sigma .$$

The second sum $h_2 = \sum_{\sigma: a_1\notin\sigma} c'_\sigma x_\sigma$ is a linear form vanishing on $P_{\mathcal{A}\backslash a_1}^{reg}$. Indeed, let T' be any regular triangulation of $\mathcal{A}\backslash a_1$. Pick a regular triangulation T of $\mathcal{A}$ that extends T' as in Lemma 2.1. Since a_1 is a vertex of $conv(\mathcal{A})$, the triangulation T cannot contain a simplex σ of the form $\{a_1\}\cup\tau$ where $conv(\tau)$ is in the boundary of $conv(\mathcal{A})$. This fact together with $h_1(v_T) = 0$ implies $h_2(v_T) = 0$, and consequently $h_2(v_{T'}) = 0$.

Every cocircuit form Co_τ of $\mathcal{A}\backslash a_1$ is either a cocircuit form of $\mathcal{A}$ as well or can be extended to a cocircuit form of $\mathcal{A}$ by adding a single variable $x_{\{a_1\}\cup\tau}$ with the appropriate sign. By induction hypothesis, h_2 is a linear combination of the cocircuit forms of $\mathcal{A}\backslash a_1$. We extend this presentation to a linear combination of cocircuit forms of $\mathcal{A}$, which vanishes on $P_{\mathcal{A}}^{reg}$. We subtract it from h_1 to get a new form h_3 which vanishes on $P_{\mathcal{A}}^{reg}$ and involves only d-simplices σ of the form $a_1 \cup \tau$:

$$h_3 \quad = \sum_{\sigma : a_1 \in \sigma} c''_{\sigma} x_{\sigma}.$$

The assignment $\tau \mapsto \{a_1\} \cup \tau$ defines a bijection between the simplices of $\mathcal{A}/a_1$ and the simplices of $\mathcal{A}$ containing a_1. Therefore we can interpret h_3 as a linear form on $P_{\mathcal{A}/a_1}^{reg}$. Lemma 2.1 guarantees that h_3 vanishes on $P_{\mathcal{A}/a_1}^{reg}$. By the induction hypothesis, h_3 is a linear combination of cocircuit forms Co_τ of $\mathcal{A}/a_1$. We replace each variable x_τ in this linear combination by the corresponding variable $x_{\{a_1\}\cup\tau}$. This transforms cocircuit forms of $\mathcal{A}/a_1$ into cocircuit forms $Co_{\{a_1\}\cup\tau}$ of $\mathcal{A}$. Therefore h_3 is a linear combination of cocircuit forms of $\mathcal{A}$. This proves Theorem 2.2 and Theorem 1.1 (i).
□

Corollary 2.3 *The linear subspace of $\mathbf{R}^{\Delta(\mathcal{A})}$ parallel to aff$(P_{\mathcal{A}})$ is spanned by all vectors $v_T - v_{T'}$ where T and T' are regular triangulations of $\mathcal{A}$ differing by a bistellar flip.*

Proof: It follows from Theorem 2.2 that the linear subspace in question is spanned by all vectors $v_T - v_{T'}$ where T and T' are regular triangulations. Since every pair of regular triangulations is connected by a sequence of bistellar flips, the corollary follows. □

Theorem 2.2 implies $dim(P_{\mathcal{A}}) = |\Delta(\mathcal{A})| - R - 1$, where R is the rank of the interior cocircuit forms. If the points of $\mathcal{A}$ are in general position then $|\Delta(\mathcal{A})| = \binom{n}{d+1}$. In this case the vector $v_T - v_{T'}$, where T and T' are regular triangulations of $\mathcal{A}$ differing by a bistellar flip, equals $v_{t(Z)^+} - v_{t(Z)^-}$, where $t(Z)^+$ and $t(Z)^-$ are the two triangulations of the circuit Z on which the bistellar flip is supported.

Theorem 2.4 *Let $\mathcal{A} \subset \mathbf{R}^d$ be a configuration of n points in general position. Let $a_1 \in \mathcal{A}$ and R be the rank of the interior cocircuit forms.*

(i) *The (interior and boundary) cocircuit forms Co_τ for which $a_1 \notin \tau$ form a basis for the space of linear forms vanishing on the linear space parallel to $P_{\mathcal{A}}$. Thus, $R + 1 = \binom{n-1}{d}$.*

(ii) *The vectors $v_{t(Z)^+} - v_{t(Z)^-}$, for the circuits Z containing a_1, form a basis of the linear space parallel to $P_{\mathcal{A}}$. Thus, $dim(P_{\mathcal{A}}) = \binom{n-1}{d+1}$.*

Proof: The cocircuit forms Co_τ for the $(d-1)$-simplices τ not containing a_1 are linearly independent, because each simplex σ containing a_1 appears in exactly one of them. Thus, $R + 1 \geq \binom{n-1}{d}$. Likewise the vectors $v_{t(Z)^+} - v_{t(Z)^-}$ for the circuits Z containing a_1 are linearly independent, because each simplex σ not containing a_1 appears in exactly one of them. Thus $dim(P_{\mathcal{A}}) \geq \binom{n-1}{d+1}$. This together with the formula $dim(P_{\mathcal{A}}) = \binom{n}{d+1} - R - 1$ finishes the proof. □

Proposition 2.5 *If $\mathcal{A}$ is a configuration in general position, then aff$(P_{\mathcal{A}})$ is defined by the chamber equations* (3).

Proposition 2.5 is generally false for configurations in special position, because a collection of d-simplices may uniquely cover all open chambers without being a triangulation. For example, the configuration consisting of the four vertices of a quadrilateral plus the intersection of its diagonals has 3 triangulations ($dim(P_{\mathcal{A}}) = 2$) while the chamber equations have 7 non-negative integer solutions whose convex hull is a 4-dimensional polytope. For the vertex set of the 3-cube we have calculated that $dim(P_{\mathcal{A}}) = 29$ but the chamber equations define an affine space of dimension 35. Proposition 2.5 is implied by Theorem 2.2 and the following lemma, which expresses interior cocircuit forms as differences of chambers.

Lemma 2.6 *Let $\mathcal{A}$ be a configuration in general position in $\mathbf{R}^d$. Let C_1 and C_2 be two neighboring maximal chambers and τ the unique $(d-1)$-simplex containing their common facet. Then*

$$\sum_{\sigma : C_1 \subset conv(\sigma)} x_\sigma \quad - \sum_{\sigma : C_2 \subset conv(\sigma)} x_\sigma \quad = \quad Co_\tau .$$

Proof: Let H be the hyperplane defined by τ, with half-spaces $H^+ \supset C_1$ and $H^- \supset C_2$. If σ is any d-simplex which contains C_1, then either σ contains C_2 as well or $\sigma = a \cup \tau$ where $a \in H^+$ (similarly for C_2). □

If $\mathcal{A}$ is in special position then more than one $(d-1)$-simplex may contain the common facet of C_1 and C_2. Call Ω the collection of them. In this case, with similar arguments one can prove that the formula in Lemma 2.6 has to be corrected by substituting $\sum_{\tau \in \Omega} Co_\tau$ for Co_τ.

Remark 2.7 Let M be the incidence matrix of the chambers and the d-simplices of $\mathcal{A}$. If $\mathcal{A}$ is in general position then Proposition 2.5 implies that

$$aff(P_{\mathcal{A}}) \quad = \quad \{x \in \mathbf{R}^{\Delta(\mathcal{A})} : M \cdot x = \mathbf{1}\}.$$

Row and column bases of M have been studied in [1] and [2] and a formula is given for *rank*(M) in [2]. The formulae in Theorem 2.4 are a special case of that formula.

3 The relation between $P_{\mathcal{A}}$ and $Q_{\mathcal{A}}$

Part (i) of Theorem 1.1 implies that the linear programming relaxation $Q_{\mathcal{A}}$ of $P_{\mathcal{A}}$ is defined by the interior cocircuit equations $Co_\tau = 0$ plus an extra non-homogeneous equation satisfied on $P_{\mathcal{A}}$, and the inequalities $x_\sigma \geq 0$ for each simplex σ of $\mathcal{A}$. Clearly $P_{\mathcal{A}} \subset Q_{\mathcal{A}}$. We shall examine the relationship between these two polytopes.

We call *support* of a point $v \in \mathbf{R}^{\Delta(\mathcal{A})}$ (and denote it *supp*(v)) the collection of d-simplices σ for which $v_\sigma \neq 0$.

Lemma 3.1 (i) *$Q_{\mathcal{A}}$ is a subpolytope of the unit cube conv$(x : x \in \{0,1\}^{\Delta(\mathcal{A})})$.*

(ii) *Every vertex of $P_{\mathcal{A}}$ is also a vertex of $Q_{\mathcal{A}}$.*

(iii) *If v is any point in $Q_{\mathcal{A}}$ then $supp(v)$ covers $conv(\mathcal{A})$, i.e.,*

$$\bigcup_{\sigma:\sigma\in supp(v)} conv(\sigma) \quad = \quad conv(\mathcal{A}).$$

(iv) *A vertex v of $Q_{\mathcal{A}}$ is a vertex of $P_{\mathcal{A}}$ if and only if its support contains a triangulation.*

Proof: The chamber equations (3) are valid for $Q_{\mathcal{A}}$, and they imply part (iii). Also, since every x_σ appears in at least one of them, the non-negativity constraints imply $x_\sigma \leq 1$ for all σ. This proves part (i) which, in turn, implies part (ii). The only-if-direction of (iv) is obvious. For the if-direction, suppose first that $supp(v)$ is the support of a triangulation T. Then the chamber equations imply that v is the incidence vector of T, hence is a vertex of $P_{\mathcal{A}}$. If $supp(v)$ strictly contains a triangulation T then it cannot be a vertex of $Q_{\mathcal{A}}$ because $v_\varepsilon := \frac{v-\varepsilon v_T}{1-\varepsilon}$ is still a point in $Q_{\mathcal{A}}$, for a sufficiently small positive ε. But then $v = (1-\varepsilon)v_\varepsilon + \varepsilon v_T$ is not a vertex of $Q_{\mathcal{A}}$. □

Theorem 3.2 *Every integral point of $Q_{\mathcal{A}}$ is the incidence vector of a triangulation of $\mathcal{A}$; i.e., $P_{\mathcal{A}}$ is the integral hull of $Q_{\mathcal{A}}$.*

Proof: Let v be an integral point of $Q_{\mathcal{A}}$. By Lemma 3.1 (iii) we only need to prove that any two simplices in $supp(v)$ intersect properly. Suppose this is not the case for two simplices σ_1 and σ_2 in $supp(v)$, i.e.:

$$conv(\sigma_1 \cap \sigma_2) \neq conv(\sigma_1) \cap conv(\sigma_2).$$

Take a point a in $(conv(\sigma_1) \cap conv(\sigma_2)) \backslash conv(\sigma_1 \cap \sigma_2)$. Then the minimal face (subset) F of σ_1 with $a \in conv(F)$ is not a face of σ_2. For each simplex σ of $supp(v)$ having F as a face, consider the convex polyhedral cone

$$c(\sigma) \;:=\; a + pos(conv(\sigma) - a) \;=\; \{\, \lambda p + (1-\lambda)a \;:\; p \in conv(\sigma),\, \lambda \geq 0 \,\}.$$

Note that the facets of $c(\sigma)$ are in 1-to-1 correspondence with the facets of σ which contain F. We claim that $conv(\mathcal{A})$ is contained in the union of such cones. Suppose a point b of $conv(\mathcal{A})$ lies outside their union. Then b "sees" a facet of some cone $c(\sigma)$, where $\sigma \in supp(v)$. Let τ be the corresponding facet of σ, which contains F. By the choice of τ, there is no d-simplex in $supp(v)$ having τ as a facet and lying in the half-space containing b. This violates the interior cocircuit equation $Co_\tau(v) = 0$, since $v \geq 0$. Therefore an open neighborhood of a in $conv(\mathcal{A})$ is covered by those simplices in $supp(v)$ which have F as a face. The interior of one of these simplices intersects the interior of $conv(\sigma_2)$. This violates the chamber equations for v. □

We next prove that the edges of $P_{\mathcal{A}}$ are also edges of $Q_{\mathcal{A}}$. A different proof for this theorem can be derived from more general results about $0-1$ polytopes due to Matsui and Tamura [10]. Here we present a self-contained proof in the context of triangulations.

Theorem 3.3 *Let T_1 and T_2 be two distinct triangulations of $\mathcal{A}$. The following statements are equivalent:*

(i) *v_{T_1} and v_{T_2} are not neighbors in $Q_{\mathcal{A}}$.*

(ii) *v_{T_1} and v_{T_2} are not neighbors in $P_{\mathcal{A}}$.*

(iii) *There exist two triangulations T_3 and T_4 of $\mathcal{A}$ (different from T_1 and T_2) such that $v_{T_1} + v_{T_2} = v_{T_3} + v_{T_4}$.*

(iv) *There exist partitions $T_1 = R_1 \cup L_1$ and $T_2 = R_2 \cup L_2$ such that $L_1 \cup R_2$ and $R_1 \cup L_2$ are two other triangulations of $\mathcal{A}$, different from T_1 and T_2.*

Proof: (iv)⇒(iii)⇒(ii)⇒(i) are obvious. We only need to show (i)⇒(iv). We define a graph G whose nodes are the d-simplices of T_1. Two d-simplices of T_1 are adjacent in G if and only if they share a common facet and this facet is not a facet of any d-simplex of T_2. The graph G has the following property: If H is a connected component of G then the (topological) boundary of $\cup_{\sigma \in H} conv(\sigma)$ is the union of $(d-1)$-simplices which are faces of both T_1 and T_2.

Next we will construct L_1, L_2, R_1 and R_2. Let σ_0 be a simplex of T_1 which is not in T_2. Let L_1 be the collection of d-simplices in the same connected component of G as σ_0 and let $R_1 = T_1 \backslash L_1$. Moreover, let $|R| = \cup_{\sigma \in R_1} conv(\sigma)$ and $|L| = \cup_{\sigma \in L_1} conv(\sigma)$. Finally, let L_2 (resp. R_2) be the collection of d-simplices of T_2 whose convex hull intersect the interior of $|L|$ (resp. of $|R|$). By the property of G mentioned above, no simplex of T_2 can intersect the interiors of both $|L|$ and $|R|$. Thus T_2 is the disjoint union of R_2 and L_2. Also, by the same property, the simplices of $L_1 \cup R_2$ (same for $R_1 \cup L_2$) intersect properly. Moreover, they cover $conv(\mathcal{A})$, because their union covers $conv(\mathcal{A})$ twice. We conclude that the disjoint unions $L_1 \cup R_2$ and $L_2 \cup R_1$ are triangulations of $\mathcal{A}$. Clearly $L_1 \neq L_2$, because $\sigma_0 \in L_1 \backslash L_2$. Let us assume that $R_1 = R_2$ and prove that then v_{T_1} and v_{T_2} are neighbors in $Q_{\mathcal{A}}$. This will finish the proof of the theorem.

Let v be a point in the minimal face F of $Q_{\mathcal{A}}$ containing v_{T_1} and v_{T_2}. This face is defined setting all coordinates not appearing in v_{T_1} or v_{T_2} equal to zero. Thus, $supp(v) \subset T_1 \cup T_2$. The entry of v corresponding to any d-simplex in $R_1 = R_2$ equals 1, because of the chamber equations. On the other hand, for any two d-simplices σ_1 and σ_2 adjacent in G, the interior cocircuit equations imply $v_{\sigma_1} = v_{\sigma_2}$, since these two simplices are the only ones in $T_1 \cup T_2$ having $\tau = \sigma_1 \cap \sigma_2$ as a facet. Thus, the entries of v corresponding to simplices in L_1 have a constant value ε. With this, the chamber equations imply that the entries corresponding to simplices in L_2 have a constant value $1-\varepsilon$. Thus, $v = \varepsilon v_{T_1} + (1-\varepsilon) v_{T_2}$. This implies that F is a segment, i.e., that v_{T_1} and v_{T_2} are neighbors. □

The above theorem implies that any two integral vertices of $Q_{\mathcal{A}}$ (triangulations) are connected by a path of integral vertices. Note that it is still conceivable that we could have two triangulations of $\mathcal{A}$ which are not connected by bistellar flips.

4 Examples and applications

Our linear programming relaxation $Q_{\mathcal{A}}$ is generally not a lattice polytope. Therefore $P_{\mathcal{A}}$ is strictly contained in $Q_{\mathcal{A}}$. In this section we will exhibit some cases where $P_{\mathcal{A}} = Q_{\mathcal{A}}$ and the smallest configuration for which $P_{\mathcal{A}} \neq Q_{\mathcal{A}}$.

Theorem 4.1 *Let $\mathcal{A} \subset \mathbf{R}^d$ be a configuration of n points. The equality $P_\mathcal{A} = Q_\mathcal{A}$ holds in the following cases:*

(i) *$d = 2$ and all points lie on the boundary of a convex polygon.*

(ii) *$d = 1$.*

(iii) *$n \leq d + 3$.*

Proof: (i) Let v be a vertex of $Q_\mathcal{A}$. Let S be a subset of $supp(v)$ where all triangles in S intersect properly and cover a convex sub-polygon of $conv(\mathcal{A})$. Suppose that S is maximal with these two properties. Let e be an edge of the sub-polygon covered by S. Then S is contained in one of the two half-planes defined by e. By maximality of S and the interior cocircuit equations, e must be a segment on the boundary of $conv(\mathcal{A})$. This proves that S covers $conv(\mathcal{A})$ and hence is a triangulation. Lemma 3.1 (iv) implies that v is a vertex of $P_\mathcal{A}$.
(ii) The proof is a minor variation of case (i).
(iii) Let $\mathcal{S} = \{T_1, \ldots, T_k\}$ be the collection of all triangulations of $\mathcal{A}$. By Corollary 5.9 below, every triangulation T_i of $\mathcal{A}$ contains a simplex σ_i which is not contained in any other triangulation. Therefore, setting the coordinate of σ_i equal to zero defines a facet of $P_\mathcal{A}$ that contains every triangulation but T_i. Thus, $P_\mathcal{A}$ is a $(k-1)$-simplex. The fact that all facets of $P_\mathcal{A}$ are defined by setting coordinates equal to zero implies that $P_\mathcal{A} = Q_\mathcal{A}$. □

Example 4.2 *A fractional vertex of $Q_\mathcal{A}$.*
For any $\mathcal{A}$ with $P_\mathcal{A} \neq Q_\mathcal{A}$ we must have $n \geq d+4 \geq 6$. A minimal example is provided by the vertices $1, \ldots, 5$ of a regular pentagon plus its center 0. This configuration is in general position and has 20 triangles and 16 triangulations. Consider the vector $v \in \mathbf{R}^{20}$ with coordinates $v_{\{123\}} = v_{\{234\}} = v_{\{345\}} = v_{\{145\}} = v_{\{125\}} = v_{\{013\}} = v_{\{024\}} = v_{\{035\}} = v_{\{014\}} = v_{\{025\}} = 1/2$ and all other coordinates zero. It satisfies the interior and boundary cocircuit equations. Therefore v lies in $Q_\mathcal{A}$. Since $supp(v)$ does not contain any triangulation, $P_\mathcal{A} \neq Q_\mathcal{A}$. This fractional point is the only vertex of $Q_\mathcal{A}$ which is not in $P_\mathcal{A}$.

Remark 4.3 The property $P_\mathcal{A} = Q_\mathcal{A}$ is neither sufficient nor necessary for a configuration $\mathcal{A}$ to have all triangulations regular. In Example 4.2 all triangulations are regular but $P_A \neq Q_A$. For the canonical example of the planar configuration which has non-regular triangulations (six points which form two triangles with parallel edges) we still have $P_\mathcal{A} = Q_\mathcal{A}$.

Let C_n be the vertex set of a planar n-gon. The following proposition gives an irredundant inequality presentation of the $\binom{n-1}{3}$-dimensional polytope $P_{C_n} = Q_{C_n}$.

Proposition 4.4 *For $n \geq 5$ the facets of P_{C_n} are defined by $x_\sigma = 0$ where σ is a triangle with at most one of its edges lying on the boundary of C_n.*

Proof: We call a triangle *external* if it has two edges on the boundary of C_n. We first show that, for any external triangle σ, $x_\sigma = 0$ does not define a facet of P_{C_n}. Without loss of generality we can assume $\sigma = \{1, 2, n\}$. Suppose that $x_\sigma = 0$ defines a facet of

P_{C_n}. Then $\binom{n-1}{3}$ affinely independent triangulations not involving σ together with a triangulation T that contains σ will affinely span $\mathit{aff}(P_{C_n})$. The triangulation T contains a triangle $\{i, 2, n\}$ for some $i \neq 1$. Since $n \geq 5$ there exists another triangulation S which contains $\{j, 2, n\}$ where $j \neq 1, i$. But then S cannot be expressed as an affine combination of the above set of triangulations since $\{j, 2, n\}$ cannot appear in any of them. This contradiction shows that an external triangle does not define a facet of Q_{C_n}.

For $n = 5$ a direct investigation shows that the five non-external triangles correspond to facets. For $n > 5$, suppose σ is a non-external triangle. Then there exists a vertex i such that $C_{n-1} := C_n \backslash i$ contains σ as a non-external triangle. Without loss of generality we can assume $i = 1$. Let τ be the external triangle $\{1, 2, n\}$. By induction, $x_\sigma = 0$ defines a facet of C_{n-1}. In other words, there are $\binom{n-2}{3}$ affinely independent triangulations of C_{n-1} which do not contain σ. This set of affinely independent triangulations can be extended to triangulations of C_n by adding τ. Now we need to produce $\binom{n-2}{2}$ additional affinely independent triangulations which do not contain σ: for each j and k such that $2 < j < k < n$, we can construct a triangulation which contains the triangles $\{1, 2, j\}, \{1, j, k\}$ and $\{1, k, n\}$ (and hence not τ), and which does not contain σ. Similarly, for each $2 < l < n$ we can construct triangulations which contain $\{1, 2, l\}$ and $\{1, l, n\}$. These additional $\binom{n-2}{2}$ triangulations together with the previous ones form a set of $\binom{n-1}{3}$ affinely independent points in P_{C_n}, none of them containing σ. □

Remark 4.5 The argument that $x_\sigma = 0$ does not define a facet of Q_{C_n} whenever σ is an external triangle can be generalized. The "external" simplices of a point configuration $\mathcal{A}$ in general position are the following: if any vertex a_i of $\mathit{conv}(\mathcal{A})$ is deleted, then the point configuration $\mathcal{A} \backslash a_i$ will again be in general position. Each facet of $\mathit{conv}(\mathcal{A} \backslash a_i)$ that is visible from a_i together with a_i will form a simplex σ which will not define a facet for $Q_\mathcal{A}$. The argument is identical to the one above.

We close this section with remarks about using $Q_\mathcal{A}$ to enumerate the triangulations of $\mathcal{A}$ and to solve optimization problems over $P_\mathcal{A}$. If $\mathcal{A}$ is in general position then, by Remark 2.7

$$\begin{aligned} Q_\mathcal{A} &= \{x \in \mathbf{R}_+^{\Delta(\mathcal{A})} : M \cdot x = \mathbf{1}\} , \\ P_\mathcal{A} &= \mathit{conv}\{x \in \{0,1\}^{\Delta(\mathcal{A})} : M \cdot x = \mathbf{1}\}. \end{aligned}$$

This means that $P_\mathcal{A}$ is a *set partitioning polytope*. Even if $\mathcal{A}$ is not in general position, by introducing extra variables for the interior $(d-1)$-simplices of $\mathcal{A}$, $P_\mathcal{A}$ can be realized as a set partitioning polytope. This has important implications for enumeration and optimization purposes. For example, Trubin [15] shows that if P is a set partitioning polytope and Q its linear relaxation then P is quasi-integral, i.e., every edge of P is also an edge of Q. Balas and Padberg [3] and Matsui and Tamura [10] have given a characterization of adjacency between vertices of P. This leads to an algorithm for optimizing linear functionals over P using Q. Starting from an integral vertex of Q the algorithm finds the optimal solution visiting only integral vertices of Q in a fashion similar to the simplex method in linear programming. The same adjacency characterization can be used to enumerate the vertices of P as well.

Unfortunately, no implementation of the Balas-Padberg procedure is known to us. Moreover enumerating the triangulations of $\mathcal{A}$ using the existing vertex enumeration packages has two major drawbacks. First of all, an inequality presentation of $P_{\mathcal{A}}$ is either hard to get or it might involve too many constraints. Secondly if one uses $Q_{\mathcal{A}}$ instead of $P_{\mathcal{A}}$ then there is an excessive number of fractional vertices which have to be enumerated too. Our experiments using several vertex enumeration packages allowed us only to compute small examples. However the situation is more promising for optimization problems over $P_{\mathcal{A}}$ thanks to the efficient implementations of *branch-and-bound* and *branch-and-cut* algorithms for integer programming. Finding triangulations using minimum or maximum number of simplices, or finding a minimum/maximum cost triangulation fall into the category of these optimization problems. In order to illustrate the sizes of the problems one can attempt to solve and the efficiency of using linear and integer programming techniques over $P_{\mathcal{A}}$ we give the following example. Triangulating the d-dimensional cube with minimal number of simplices is important for their use in algorithms for the computation of fixed points of continuous maps [14]. This is equivalent to minimizing the functional $\sum x_\sigma$ over $P_{\mathcal{A}}$ where $\mathcal{A}$ is the vertex set of the d-cube. In the case of the 4-cube ($d = 4, n = 16$) the system defining $Q_{\mathcal{A}}$ has 1257 equations and 3008 variables. Using CPLEX 3.0 on a SPARC10 workstation, it takes 150 seconds to verify that the 4-cube's minimal triangulation has 16 simplices. The size of the corresponding linear programs for the d-cube with $d \geq 5$ gets too big. In this case one should formulate a smaller linear program which exploits the symmetries of the d-cube. This was done successfully for $d = 5, 6$ in [8].

5 Duality

In this section we extend the duality theory in [5] to cover all triangulations. Two applications are included: a short proof of Carl Lee's result that every triangulation of a configuration of $d + 3$ points is regular [9] and an exponential lower bound for the number of triangulations of the cyclic polytope $C_{4n-4}(4n)$.

Following [5], we now consider a configuration $\mathcal{A} = \{a_1, a_2, \ldots, a_n\}$ of *vectors* spanning $\mathbf{R}^{d+1}$. Thus $|\Delta(\mathcal{A})| \leq \binom{n}{d+1}$ and equality holds if $\mathcal{A}$ is in general position. If there exists an affine hyperplane H ($0 \notin H$) which intersects $pos(a_i)$ for every $a_i \in \mathcal{A}$, then we can consider $\mathcal{A}$ as a d-dimensional point configuration in H (identified with $\mathbf{R}^d$). If this is the case $\mathcal{A}$ is said to be *acyclic*, and we are in the setting of the previous sections. For the non-acyclic case see Remark 5.2 below.

Let $\mathcal{B} = \{b_1, b_2, \ldots, b_n\}$ be a spanning subset of $\mathbf{R}^{n-d-1}$ which is a *Gale transform* of $\mathcal{A}$. This means that $\sum_{i=1}^n a_i \otimes b_i = 0$ in $\mathbf{R}^{d+1} \otimes \mathbf{R}^{n-d-1}$. In particular, $M(\mathcal{B})$ is the oriented matroid dual to $M(\mathcal{A})$.

Let $\mathcal{T}(\mathcal{A}) \subset \{0,1\}^{\Delta(\mathcal{A})}$ be the set of incidence vectors of all triangulations of $\mathcal{A}$ and $\mathcal{T}_{reg}(\mathcal{A})$ the subset corresponding to regular triangulations. Similarly let $\Gamma(\mathcal{A}) \subset \{0,1\}^{\Delta(\mathcal{A})}$ be the set of all chambers of $\mathcal{A}$. By the results in [5],

$$\mathcal{T}_{reg}(\mathcal{A}) = \Gamma(\mathcal{B}) \qquad \text{and} \qquad \mathcal{T}_{reg}(\mathcal{B}) = \Gamma(\mathcal{A}). \tag{5}$$

We identify each cocircuit form Co_τ of $\mathcal{A}$ with its coefficient vector in $\{0, +1, -1\}^{\Delta(\mathcal{A})}$. Let $Co(\mathcal{A})$ denote the collection of all *cocircuit vectors* Co_τ, where τ runs over all linearly independent d-subsets of $\mathcal{A}$, and let $Co_{int}(\mathcal{A})$ be the subset of interior cocircuit vectors. Recall that Co_τ is *interior* if and only if both $+1$ and -1 appear among the coordinates of Co_τ. Dually, let ρ be any spanning $(d + 2)$-subset of $\mathcal{A}$. Then ρ

contains a unique signed circuit $Z = (Z_+, Z_-)$ of $\mathcal{A}$. We define the *circuit vector*

$$Ci_\rho \quad := \quad \sum_{i \in Z_-} e_{\rho\backslash i} \quad - \quad \sum_{j \in Z_+} e_{\rho\backslash j}.$$

(The $e_{\rho\backslash i}$ are standard basis vectors in $\mathbf{R}^{\Delta(\mathcal{A})}$.) We say that Ci_ρ is an *interior* circuit vector if $Z_+ \neq \emptyset$ and $Z_- \neq \emptyset$. Let $Ci(\mathcal{A})$ denote the set of all circuit vectors and $Ci_{int}(\mathcal{A})$ the subset of interior circuit vectors. $\mathcal{A}$ is acyclic if and only if $Ci(\mathcal{A}) = Ci_{int}(\mathcal{A})$. We fix the standard inner product $\langle \cdot\,, \cdot \rangle$ on $\mathbf{R}^{\Delta(\mathcal{A})}$. Here is our first duality theorem.

THEOREM 5.1 *Let $\mathcal{A} \subset \mathbf{R}^{d+1}$ and $\mathcal{B} \subset \mathbf{R}^{n-d-1}$ be Gale transforms of each other.*

(i) *Circuit and cocircuit vectors satisfy $Ci(\mathcal{A}) = Co(\mathcal{B})$ and $Ci(\mathcal{B}) = Co(\mathcal{A})$.*

(ii) *If $\mathcal{A}$ is in general position then the subspaces spanned by $Ci(\mathcal{A})$ and $Co(\mathcal{A})$ are orthogonal complements in $\mathbf{R}^{\Delta(\mathcal{A})}$.*

Proof: Recall the following two facts from oriented matroid duality: A $(d+2)$-subset of $\mathcal{A}$ is spanning if and only if the complementary $(n-d-2)$-subset of $\mathcal{B}$ is linearly independent. The signed circuits of $\mathcal{A}$ are the signed cocircuits of $\mathcal{B}$ and vice versa. These two facts imply assertion (i).
(ii) Let Ci_ρ be a circuit vector and let Co_τ be a cocircuit vector. If $supp(Ci_\rho) \cap supp(Co_\tau) = \emptyset$ then $\langle Co_\tau, Ci_\rho \rangle = 0$. Otherwise $supp(Ci_\rho) \cap supp(Co_\tau) = \{\sigma_1, \sigma_2\}$ where $\sigma_1 = \{\tau, i\} = \rho\backslash j$ and $\sigma_2 = \{\tau, j\} = \rho\backslash i$. If a_i and a_j are in different half-spaces defined by τ then $\{\tau, i\}$ and $\{\tau, j\}$ have the same sign in Ci_ρ since they appear in the same triangulation of ρ. If a_i and a_j are in the same half-space then $\{\tau, i\}$ and $\{\tau, j\}$ have opposite signs in Ci_ρ since they do not belong to the same triangulation of ρ. This shows

$$\langle Co_\tau, Ci_\rho \rangle \;=\; 0 \qquad \text{for all} \quad Co_\tau \in Co(\mathcal{A})\,,\; Ci_\rho \in Ci(\mathcal{A}). \tag{6}$$

If $\mathcal{A}$ is acyclic, by the proof of Theorem 2.4 we have $dim(Co(\mathcal{A})) = \binom{n-1}{d}$ and $dim(Ci(\mathcal{A})) = \binom{n-1}{d+1}$. We conclude that $Ci(\mathcal{A})$ and $Co(\mathcal{A})$ span orthogonal complements. If $\mathcal{A}$ is in general position but not acyclic, then $\mathcal{B}$ is acyclic and the result follows from (i). □

The orthogonality relation (6) need not hold for configurations $\mathcal{A}$ in special position. For example, let $\mathcal{A} = \{a_1, a_2, \ldots, a_6\}$ be the vertex set of the regular octahedron where a_5 and a_6 are not connected by an edge. If we choose $\tau = \{1, 2, 3\}$ and $\rho = \{1, 2, 3, 4, 5\}$, then $\langle Co_\tau, Ci_\rho \rangle \neq 0$.

REMARK 5.2 The main results of Section 2 and 3 can be extended to vector configurations. The dimension formula given for point configurations in general position should be corrected to $dim(P_{\mathcal{A}}) = \binom{n-1}{d+1} - 1$ whenever $\mathcal{A}$ is a non-acyclic configuration in general position. This can be proved using duality.

PROPOSITION 5.3 *For every vector configuration $\mathcal{A}$, the orthogonal complement of $Ci(\mathcal{A})$ is contained in $Co(\mathcal{A})$ and vice versa.*

Proposition 5.3 can be deduced from the following theorem.

Theorem 5.4 *For a vector* $X \in \{0,1\}^{\Delta(\mathcal{A})}$ *the following are equivalent:*

(a) $\forall T \in \mathcal{T}_{reg}(\mathcal{A}) \quad \langle X, T\rangle = 1$ *and* $\forall Ci_\rho \in Ci_{int}(\mathcal{A}) \quad \langle X, Ci_\rho\rangle = 0.$

(b) $\forall T \in \mathcal{T}(\mathcal{A}) \quad \langle X, T\rangle = 1$ *and* $\forall Ci_\rho \in Ci_{int}(\mathcal{A}) \quad \langle X, Ci_\rho\rangle = 0$

(c) X *is the incidence vector of a triangulation of* $\mathcal{B}$*, that is, the set of all* $(n-d-1)$*-subsets* σ *satisfying* $X_{\{1,\ldots,n\}\setminus\sigma} = 1$ *defines a triangulation of* $\mathcal{B}$*.*

Proof: The equivalence of (a) and (b) follows from $\mathit{aff}(\mathcal{T}_{reg}(\mathcal{A})) = \mathit{aff}(\mathcal{T}(\mathcal{A}))$ which is a consequence of Corollary 2.4. Using Theorem 5.1 (i) and equation (5) we see that (a) is equivalent to

$\forall C \in \Gamma(\mathcal{B}) \quad \langle X, C\rangle = 1$ and $\forall Co_\tau \in Co_{int}(\mathcal{B}) \quad \langle X, Co_\tau\rangle = 0.$

The 0-1-solutions X to this are the triangulations of $\mathcal{B}$ by Theorem 1.1. □

Corollary 5.5 *If* $\mathcal{A}$ *is in general position then for a vector* $X \in \{0,1\}^{\Delta(\mathcal{A})}$ *the following are equivalent:*

(a) $\langle X, T\rangle = 1$ *for all* $T \in \mathcal{T}_{reg}(\mathcal{A})$.

(b) $\langle X, T\rangle = 1$ *for all* $T \in \mathcal{T}(\mathcal{A})$.

(c) X *is the incidence vector of a triangulation of* $\mathcal{B}$*.*

The chambers of $\mathcal{A}$ constitute a (generally proper) subset of the vectors $X \subset \{0,1\}^{\Delta(\mathcal{A})}$ characterized by the three equivalent conditions in Theorem 5.4. We propose the following interpretation for the remaining solutions:

Definition 5.6 *The solutions* $X \in \{0,1\}^{\Delta(\mathcal{A})}$ *to the system* (a) *in Theorem 5.4, viewed as collections of* $(d+1)$*-subsets in* $\mathcal{A}$*, are called the* virtual chambers *of* $\mathcal{A}$*. Writing* $\Gamma_{virt}(\cdot)$ *for the set of virtual chambers, we have*

$$\mathcal{T}(\mathcal{A}) = \Gamma_{virt}(\mathcal{B}) \qquad \textit{and} \qquad \mathcal{T}(\mathcal{B}) = \Gamma_{virt}(\mathcal{A}). \tag{7}$$

Remark 5.7 There are two kinds of virtual chambers in $\Gamma_{virt}(\mathcal{B}) \setminus \Gamma(\mathcal{B})$: the first kind of these can become real chambers in a different realization of $M(\mathcal{B})$. These correspond to the triangulations of $\mathcal{A}$ which are regular in some other realization of $M(\mathcal{A})$. The second kind, the "truly" virtual chambers, will never show up as real chambers and thus they correspond to non-regular triangulations which never become regular. An example of the second kind can be found in Proposition 9.6.4 in [6].

We now present two applications of our duality results.

Proposition 5.8 (Carl Lee, [9]) *If* $\mathcal{A}$ *is a vector configuration in* $\mathbf{R}^{d+1}$ *with* $|\mathcal{A}| \leq d+3$ *then every triangulation of* $\mathcal{A}$ *is regular.*

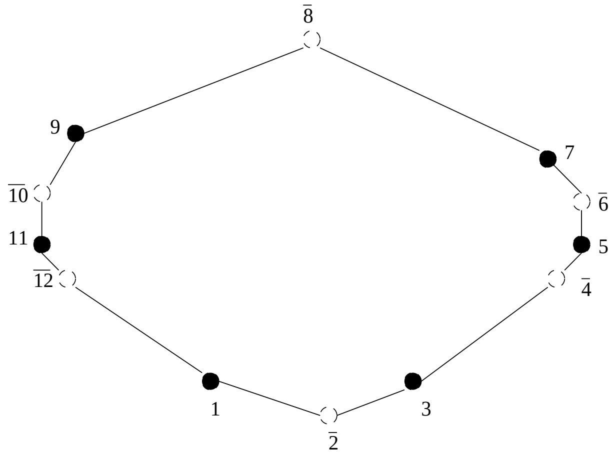

Figure 1: A Gale diagram of $C_{12}(8)$

Proof: If $|\mathcal{A}| = d+1$, there is a trivial triangulation which is regular. The case $|\mathcal{A}| = d+2$ is still very easy, since the linear transform $\mathcal{B}$ of $\mathcal{A}$ is one-dimensional. Let us assume that $|\mathcal{A}| = d+3$. Then $\mathcal{B}$ is two-dimensional and its simplices are pairs of independent vectors. Let C be a virtual chamber of $\mathcal{B}$ and let $\{b_1, b_2\}$ be a simplex in the support of C. We have the following:
(a) if $b' \in \mathcal{B}$ is such that $pos(b_1, b_2) \subset pos(b_1, b')$ then $\{b_1, b'\}$ is in the support of C.
(b) if $b' \in \mathcal{B}$ lies in $pos(b_1, b_2)$, then exactly one of $\{b_1, b'\}$ and $\{b', b_2\}$ is in the support of C.

Both properties follow from $\forall Ci_\rho \in Ci_{int}(\mathcal{B})$, $\langle C, Ci_\rho \rangle = 0$. Moreover, since $\forall T \in \mathcal{T}(\mathcal{B})$, $\langle C, T \rangle = 1$, a simplex $\{b_i, b_j\}$ where $pos(b_i, b_j) \not\subseteq pos(b_1, b_2)$ cannot be in C. From these we conclude that there is a unique minimal simplex lying in C. This implies that C is a real chamber. □

Corollary 5.9 *If $\mathcal{A}$ is a vector configuration in $\mathbf{R}^{d+1}$ with $|\mathcal{A}| \leq d+3$, then every triangulation of $\mathcal{A}$ contains a simplex which is not used in any other triangulation of $\mathcal{A}$.*

Proof: For $|\mathcal{A}| = d+1$ and for $|\mathcal{A}| = d+2$ the statement is again trivial. For the case $|\mathcal{A}| = d+3$, the minimal simplex of C in the proof of the Theorem 5.8 is such a simplex. □

The triangulations of all cyclic polytopes $C_d(n)$ are known to be connected by bistellar flips [11]. We close the paper with a result about the abundance of non-regular triangulations of $C_d(n)$.

Proposition 5.10 *If $\mathcal{A}$ is the vertex set of $C_{4n-4}(4n)$, then $\mathcal{A}$ has $O(n^4)$ regular triangulations and has at least 2^n triangulations.*

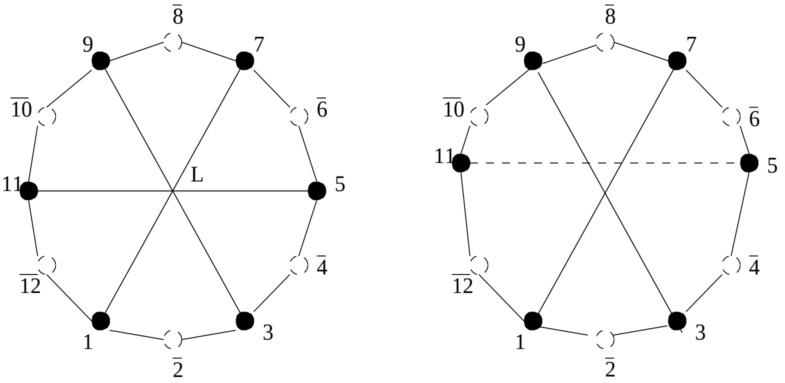

Figure 2: A perturbation that creates a chamber

Proof: The fact that $\mathcal{A}$ has $O(n^4)$ regular triangulations was established in [4], Theorem 5.7 (note that their "d" corresponds to the "$d+1$" in our notation). A Gale transform $\mathcal{B} = \{b_1, b_2, \dots, b_{4n}\}$ of $\mathcal{A}$ can be depicted as a $(4n)$-gon under taking the antipodals of $\{b_2, b_4, \dots, b_{4n}\}$ (see Figure 1 and also compare with Figure 1 in [5]). That configuration can be assumed to be a regular $4n$-gon, in particular, the sub-configuration $\{b_1, b_3, \dots, b_{4n-1}\}$ is a regular $2n$-gon. Hence, the cones $pos(b_i, b_{i+2n}), 1 \leq i \leq 2n-1$ and i odd, intersect in a half-line L. Now by perturbing the vertices of this $2n$-gon, we can create 2^n different chambers which correspond to 2^n distinct triangulations of $\mathcal{A}$ (see Figure 2). These chambers are virtual chambers of $\mathcal{A}$, because the small perturbation does not change the oriented matroid. This shows that $C_{4n-4}(4n)$ has at least 2^n non-regular triangulations. All of them become regular for some point configuration combinatorially equivalent to $\mathcal{A}$. ◻

Acknowledgments: We thank Victor Reiner for helpful remarks and questions. We are also grateful to David Avis for his help on using his vertex enumeration package RS.

References

[1] T.V. Alekseyevskaya, I.M. Gel'fand and A. Zelevinsky *Arrangements of real hyperplanes and the associated partition function*, Soviet Math. Doklady 36, (1988), 589-593.

[2] T.V. Alekseyevskaya *Combinatorial bases in systems of simplices and chambers*, to appear on FPSAC 94, International Conference on Formal Power Series and Algebraic Combinatorics, 1994.

[3] E. Balas and M. Padberg *On the set-covering problem: II. An algorithm for set partitioning,* Operations Research, 23 (1975), 1152-1161.

[4] L. Billera, P. Filliman and B. Sturmfels *Constructions and complexity of secondary polytopes,* Advances in Math. 83 (1990), 155-179.

[5] L. Billera, I.M. Gel'fand and B. Sturmfels *Duality and minors of secondary polyhedra*, J. Combinatorial Theory, Ser. B 57 (1993), 258-268.

[6] A. Björner, M. Las Vergnas, B. Sturmfels, N. White and G. Ziegler *Oriented Matroids*, Cambridge University Press, 1992.

[7] I.M. Gel'fand, M. Kapranov, A. Zelevinsky *Multidimensional Determinants, Discriminants and Resultants*, Birkhäuser, Boston, 1994.

[8] R.B. Hughes *Minimum-cardinality triangulations of the d-cube for $d = 5$ and $d = 6$*, Discrete Mathematics 118 (1993), 75-118.

[9] C.W. Lee *Regular triangulations of convex polytopes,* in: Applied Geometry and Discrete Mathematics-The Victor Klee Festschrift, (P. Gritzmann and B. Sturmfels eds.) Dimacs Series in Discrete Math. and Theoretical Comp. Science 4 (1991), 443-456.

[10] T. Matsui and S. Tamura *Adjacency on combinatorial polyhedra*, Discrete Applied Math. 56 (1995), 311-321.

[11] J. Rambau *Triangulations of cyclic polytopes and higher Bruhat orders,* Report No. 496 (1996), TU Berlin,
`ftp.math.tu-berlin.de/pub/Preprints/combi/Report-496-1996.ps.Z`.

[12] J. Rambau and G. Ziegler *Projections of polytopes and the generalized Baues conjecture*, Discrete Comput. Geometry, to appear.

[13] E. Schönhardt *Über die Zerlegung von Dreieckspolyedern in Tetraeder*, Mathematische Annalen, 98 (1928), 309-312.

[14] M.J. Todd *The Computation of Fixed Points and Applications*, Springer-Verlag, Berlin, 1976.

[15] V.A. Trubin *On a method of solution of integer linear programming problems of a special kind*, Soviet Math. Doklady, 10 (1969), 1544-1546.

[16] G. Ziegler *Lectures on Polytopes*, Springer-Verlag, New York, 1994.

Jesús A. de Loera
The Geometry Center and
School of Mathematics
University of Minnesota
Minneapolis, MN 55455, USA
deloera@geom.umn.edu

Serkan Hoşten
School of Operations Research
Cornell University
Ithaca, NY 14853, USA
hosten@orie.cornell.edu

Fransisco Santos
Departamento de Matemáticas
Estadística y Computación
Universidad de Cantabria
Santander, E-39071, Spain
santos@matsun1.unican.es

Bernd Sturmfels
Department of Mathematics
University of California
Berkeley, CA 94720, USA
bernd@math.berkeley.edu

On the Construction of the Kan Loop Group

Friedhelm Waldhausen

Received: December 19, 1995

Communicated by Ulf Rehmann

Abstract. A re-make, not of the construction, but of its description.

By an *ordered graph* will be meant a triple of sets (P, N, E) together with a pair of structure maps, $N \leftarrow E \rightarrow P$. These data are to be thought of as 'positive vertices', 'negative vertices', 'edges', and 'incidence relations', respectively.

An ordered graph is a sort of ordered simplicial complex. It can be made into a simplicial set by adding degenerate simplices. The details of this step can be neatly described by means of an auxiliary category C_Γ associated to the ordered graph Γ. The set of objects of C_Γ is the disjoint union $N \dot\cup P$; the set of non-identity morphisms is the set E, and the source and target functions on E are given by the two maps $E \rightarrow N$ and $E \rightarrow P$, respectively. The category is a little unusual insofar as no two morphisms in it can be composable unless at least one of them is an identity morphism.

The *nerve* construction produces a simplicial set $N(C_\Gamma)$ now: an m-simplex is a functor $[m] \rightarrow C_\Gamma$ (where $[m]$ denotes the ordered set $(0 < 1 < \cdots < m)$, regarded as a category). The set of m-simplices is thus a disjoint union $N \dot\cup E \dot\cup \ldots \dot\cup E \dot\cup P$, with one entry "$E$" for each surjective monotone map $[m] \rightarrow [1]$. The simplicial set $N(C_\Gamma)$ is 1-*dimensional* in the sense that every non-degenerate simplex has dimension ≤ 1. Instead of $N(C_\Gamma)$ we will henceforth write $N(\Gamma)$ for this simplicial set.

The geometric realization $|N(\Gamma)|$ is a *CW* complex of dimension ≤ 1. The 0-cells of $|N(\Gamma)|$ are indexed by the set $N \dot\cup P$ (disjoint union), and the 1-cells are indexed by the set E.

We will suppose now that the ordered graph Γ is *connected* (equivalently, that the *CW*-complex $|N(\Gamma)|$ is) and *pointed* (i.e., equipped with the choice of an element $x \in P$). We may then speak of the *fundamental group* $\pi_1(\Gamma, x)$. It can be described as the fundamental group of the *CW*-complex $|N(\Gamma)|$ based at $|x|$ or else, in somewhat more combinatorial terms, as the *edge path group* of Γ based at x.

We may also speak of the *universal covering* of Γ (with respect to the chosen basepoint x). This is an ordered graph $\widetilde{\Gamma}$. It comes equipped with an action of $\pi_1(\Gamma, x)$, and with a map $\widetilde{\Gamma} \rightarrow \Gamma$; and these two pieces of data are such that they make $\widetilde{\Gamma}$ into a principal $\pi_1(\Gamma, x)$-bundle over Γ (by definition, this means that the action is free, and that the quotient by the action is identified to Γ by the given map). The construction of all this is as follows, by covering space theory. An element of $\widetilde{N}$ (a 'negative vertex' of $\widetilde{\Gamma}$) consists of a pair of data in Γ, namely (i) a 'negative vertex' v of Γ and (ii) a homotopy class of paths connecting v to the chosen basepoint x. The map $\widetilde{N} \rightarrow N$ is defined as the forgetful map which forgets the path; and the action of $\pi_1(\Gamma, x)$ on $\widetilde{N}$ is given by composing the path with the loop in question. The other data are given similarly.

As we have implicitly used before, the ordered graphs are the objects of a category in an evident way: a map in this category is a triple of maps of sets, $P \to P'$, $N \to N'$, $E \to E'$, so that these maps are compatible to the structure maps of the two ordered graphs in question. It makes sense, consequently, to speak of a *simplicial ordered graph*, a simplicial object in the category of ordered graphs. We note that a simplicial ordered graph will give rise to a bisimplicial set, by *nerve*, and hence to a CW-complex, by geometric realization (this particular geometric realization uses 'prisms'; an equivalent construction, up to canonical isomorphism, would be to pass to the diagonal simplicial set first and then take the geometric realization of that diagonal simplicial set). We are in a position now to describe our basic construction. The construction is implicit in Kan's paper [1], but it was not made explicit there.

CONSTRUCTION. Let $X\colon \Delta^{\mathrm{op}} \to (\mathrm{sets})$, $[n] \mapsto X_n$, be a simplicial set. There is an associated simplicial ordered graph. It has $P_n = X_n$, $N_n = X_0$, $E_n = X_{n+1}$, and the maps $E_n \to P_n$ and $E_n \to N_n$ are given by the 'last face' map and 'last vertex' map, respectively. This simplicial ordered graph will be denoted ΓX.

(Here are some more details. The simplicial set $P_\bullet$ is defined to be isomorphic to X itself, while $N_\bullet$ is defined as the set X_0 considered as a simplicial set in a trivial way. Concerning $E_\bullet$, if $\alpha\colon [n] \to [n']$ is a monotone map then $\alpha^* : E_{n'} \to E_n$ is defined to be the map $X_{n'+1} \to X_{n+1}$ induced from $\alpha \cup \{\infty\} : \ [n] \cup \{\infty\} \to [n'] \cup \{\infty\}$. The map $E_n \to P_n$ is defined to be the map $X_{n+1} \to X_n$ induced from the injective map $[n] \to [n{+}1]$ which misses $n{+}1$, and the map $E_n \to N_n$ is defined to be the map $X_{n+1} \to X_0$ induced from the map $[0] \to [n{+}1]$ taking 0 to $n{+}1$.)

Considering X as a simplicial ordered graph in a trivial way (no edges, no negative vertices) we have a natural inclusion $X \to \Gamma X$. The following will be shown later.

LEMMA. *The map $X \to \Gamma X$ is a weak homotopy equivalence.*

We will suppose now that the simplicial set X is connected and that it is equipped with a basepoint (that is, the choice of an element in X_0). Then $\Gamma_n X$, the ordered graph in degree n of the simplicial ordered graph ΓX, will also be connected (a proof of this fact will be given below) and it will be equipped with a basepoint x_n (namely the degenerate in degree n of the chosen element in X_0). ΓX can thus be considered as a simplicial object of *pointed* ordered graphs, and we can therefore define a simplicial group $G = G(X)$,

$$[n] \ \mapsto \ G_n := \pi_1(\Gamma_n X, x_n) \ .$$

THEOREM. *The simplicial group G is a loop group for X.*

PROOF. In view of the lemma it will suffice to show that there is a principal G-bundle over ΓX with weakly contractible total space ('weakly contractible' means that the map to the one-point-space is a weak homotopy equivalence). For by pulling back such a bundle along the map $X \to \Gamma X$ we can obtain a principal G-bundle over X, and the total space of the latter bundle will again be weakly contractible. (This is so since, for example, a map of principal bundles of simplicial sets is also a map of Kan fibrations [2, Satz 9.5] and the geometric realization of a Kan fibration is a Serre fibration [3]. So the Whitehead theorem applies.)

The universal covering of a pointed ordered graph, as described above, is *functorial*. Hence we have a simplicial ordered graph $[n] \mapsto \widetilde{\Gamma}_n X$, it is obtained from $[n] \mapsto \Gamma_n X$ (that is, from ΓX) by taking the universal covering degreewise. This simplicial ordered graph is weakly contractible in every degree; hence (by [4, Appendix A] for example) it is also weakly contractible globally.

The desired principal bundle is now obtained by observing that the simplicial group G acts on $\widetilde{\Gamma} X$, that the action is free, and that the quotient of $\widetilde{\Gamma} X$ by the action is just ΓX again. $\square$

Proof of Lemma. We give two proofs, both fairly self-contained. The short proof is in an appendix; here is the pedestrian one.

The category of ordered graphs, as well as the category of simplicial sets, is a *functor category*, namely the category of $\nwarrow\!\nearrow$-shaped, respectively of Δ^{op}-shaped, diagrams in the category of sets; and colimits in such a functor category are computed 'pointwise'. It results that the functor $X \mapsto \Gamma X$ commutes with colimits and, what is more to the point here, that the functor $X \mapsto N(\Gamma X)$ (and therefore also $X \mapsto |N(\Gamma X)|$) does so, too. We can apply this fact in two ways. First, by direct limit, we can reduce to proving the lemma for those simplicial sets which are *finite*; that is, there are only finitely many non-degenerate simplices. Next, a finite simplicial set can be obtained from a 'smaller' one by the attaching of a simplicial set *standard k-simplex*, for some k; by induction and the gluing lemma we can therefore reduce to proving the lemma for just the latter kind of simplicial set. In other words, we are reduced now to showing that $|N(\Gamma\Delta^k)|$ is contractible.

To show this, we will work out the cell structure of the *CW*-complex $|N(\Gamma\Delta^k)|$ explicitly. The cells in this complex are of three kinds. First, there are the cells coming from the *positive vertices*; these contribute the copy of $|\Delta^k|$ coming from the inclusion $\Delta^k \to N(\Gamma\Delta^k)$. Next, there are the cells coming from the *negative vertices*; these cells are all 0-dimensional, and there is one such for every vertex of Δ^k.

And, finally, there are the cells coming from the *non-degenerate edges*; of these there is a 'basic' edge for every negative vertex. Namely suppose that the negative vertex corresponds to the l-th vertex of Δ^k. Let $\mathrm{front}_l(\Delta^k)$ denote the copy of Δ^l inside Δ^k whose vertices are the vertex l and its predecessors. Then the *last degenerate* of the generating simplex of $\mathrm{front}_l(\Delta^k)$ gives an l-dimensional edge of the simplicial ordered graph, and this edge is non-degenerate. Conversely, every non-degenerate edge is either of this kind or is a face of one such. Indeed, suppose the edge corresponds to a simplex y of Δ^k and suppose that l is the highest vertex of Δ^k occurring in y. If any vertex $< l$ occurs twice in y, or if the vertex l occurs more than twice, then the edge associated to y is degenerate—contrary to assumption. If, on the other hand, some vertex $< l$ does not occur at all, or if the vertex l occurs only once rather than twice, then the edge associated to y is a proper face of one of higher dimension.

Returning to the 'basic' edge, we note that the associated cell has dimension $l+1$. Its closure is the image of a copy of $|\Delta^l| \times |\Delta^1|$ which is mapped in such a way that all of $|\Delta^l| \times 0$ is identified to a point (corresponding to the negative vertex in question), while $|\Delta^l| \times 1$ is identified to the geometric realization of $\mathrm{front}_l(\Delta^k)$. By induction, there are no identifications over faces of Δ^k which are not of this kind. It results that $|N(\Gamma\Delta^k)|$ is the union of the cones on $|\Delta^0|, |\Delta^1|, \ldots, |\Delta^k|$, each glued along its base to the appropriate subsimplex in $|\Delta^k|$. This complex is indeed contractible. $\square$

APPENDIX (*on generators and relations*).

If the groups $G_n = \pi_1(\Gamma_n X, x_n)$ are expressed as *edge path groups*, one obtains a sort of description of the simplicial group G in terms of the structure of X. This description occurs as a definition of the loop group in [1, section 12]. Another definition of the loop group is given in [1, sections 7 and 9] in terms of *generators and relations*. The equivalence of the two definitions can be explained by combinatorial group theory. Namely, in a connected graph one can choose a *maximal tree*. The fundamental group of the graph can then be identified to the free group freely generated by the edges of the graph *not* in that maximal tree; equivalently, the fundamental group can be identified to the group generated by *all* the edges of the graph, where, however, the edges of the chosen maximal tree are also introduced as relations.

To make this description effective, one needs to know what a maximal tree in the ordered graph $\Gamma_n X$ will look like. The answer is as follows. If the simplicial set X is *reduced* (that is, if X_0, the set of 0-simplices, has only one element) then there is a maximal tree in $\Gamma_n X$ which is such that it contains exactly those edges where the corresponding simplex of X is a *last degenerate*. In the general case of a connected, but not necessarily reduced X, one has to choose a *maximal tree* in X first (a sub-simplicial-set which contains all of X_0 and whose geometric realization is a simply-connected CW-complex of dimension ≤ 1); the pieces in $\Gamma_n X$ coming from this sub-simplicial-set are then, additionally, in the maximal tree in $\Gamma_n X$.

We will justify this description of the maximal tree now (for much of the following, cf. [1, Lemma 9.1] and [1, section 14] in particular). We begin by explaining why, for connected X and for every n, the graph $\Gamma_n X$ is connected. First, every positive vertex of $\Gamma_n X$ can be connected to some negative vertex. Indeed, if the positive vertex corresponds to $x \in X_n$ then the last degenerate of x gives an edge in $\Gamma_n X$ which will connect this positive vertex to a negative vertex (namely the one associated with the 'last vertex' of that last degenerate or, what amounts to the same thing, the 'last vertex' of x itself). Next, all the negative vertices of $\Gamma_n X$ come from $\Gamma_0 X$, by degeneracy, hence it will suffice to show that they can be connected to each other inside $\Gamma_0 X$. It will, in fact, suffice to show this in the special case of two negative vertices where the associated 0-simplices of X are *adjacent* (in making this reduction we are using the assumed fact that X is connected). We are thus in the special case now where the two 0-simplices of X are the faces of some $y \in X_1$. We see that in this case the two negative vertices can be connected to each other by an edge path of length 2 in $\Gamma_0 X$; the two edges in the path are provided by the simplex y on the one hand and by the 1-dimensional degenerate of the last face of y on the other.

Next, suppose that the simplicial set X is a *tree*. We want to show that, in this case, the ordered graph $\Gamma_n X$ is a tree, too, for every n. Now the nerve $N(\Gamma_n X)$ is 1-dimensional, and connected; so it will be a tree if (and only if) it is *acyclic*. To prove the latter, since the functor $X \mapsto N(\Gamma_n X)$ commutes with colimits, we can further reduce, by direct limit and (inductively) the gluing lemma, to dealing with just the two cases where $X = \Delta^0$ or $X = \Delta^1$. We will write P_n, E_n, N_n, respectively, for the sets of positive vertices, edges, and negative vertices of $\Gamma_n X$. In the case $X = \Delta^0$, each of these sets has exactly one element, so $N(\Gamma_n \Delta^0)$ is isomorphic to Δ^1. In the case $X = \Delta^1$, the set P_n has $n+2$ elements which we denote $p_0, p_1, \dots, p_{n+1}$ (where p_{n+1} stands for the map $[n] \to [1]$ with image consisting of only $0 \in [1]$ and where, otherwise, p_i stands for the monotone map $[n] \to [1]$ having the property that

$i \in [n]$ is the smallest element whose image is $1 \in [1]$); the set E_n has $n{+}3$ elements, $e_0, e_1, \ldots, e_{n+2}$, and the set N_n has two elements, n_0 and n_1. The map $E_n \to P_n$ takes e_i to p_i for all $i \le n{+}1$, and, in addition, it takes e_{n+2} to p_{n+1}. The map $E_n \to N_n$ takes the element e_{n+2} into n_0 and it takes all other elements of E_n into n_1. We see that $N(\Gamma_n \Delta^1)$ is a one-point-union of $n{+}1$ copies of Δ^1, together with one extra copy of Δ^1 hanging on to one of the whiskers. It is a tree indeed.

Let X be a connected simplicial set now. Choose a maximal tree T in X. Let P', E', N' denote, respectively, the sets of positive vertices, edges, and negative vertices of $\Gamma_n T$. Let P'' denote the subset of X_n which is complementary to the subset T_n. Let E'' be defined as the subset of X_{n+1} given by the image of P'' under the 'last degeneracy' map. One of the structure maps of $\Gamma_n X$ restricts to a map $E'' \to N'$ (all the negative vertices of $\Gamma_n X$ are contained in N' since T contains all the 0-simplices of X), and the other structure map restricts to a map $E'' \to P''$. The latter map is given by the 'last face' map, and is actually inverse to the above map $P'' \to E''$; in particular it is an isomorphism. In view of this fact, and using the fact established above, that the ordered graph

$$N' \,,\ E' \,,\ P' \ , \quad N' \leftarrow E' \rightarrow P'$$

is indeed a tree, we can now conclude that the sets, and maps,

$$N' \,,\ E' \cup E'' \,,\ P' \cup P'' \quad , \quad N' \leftarrow E' \cup E'' \rightarrow P' \cup P''$$

do form a tree, too. The isomorphisms $N' \approx X_0$ and $P' \cup P'' \approx X_n$ show that this tree contains all the vertices of $\Gamma_n X$. It is therefore a maximal tree. ◻

Appendix (*another view at the lemma*).

The geometric realization $|\Gamma X|$ may be identified to the double mapping cylinder of the following diagram (the terms involved have been defined in connection with the definition of ΓX),

$$|P_{\bullet}| \ \longleftarrow \ |E_{\bullet}| \ \longrightarrow \ |N_{\bullet}| \,.$$

As a consequence, the assertion of the lemma, that the inclusion

$$|X| \,\approx\, |P_{\bullet}| \ \longrightarrow \ |\Gamma X|$$

is a homotopy equivalence, will therefore result once one knows that the map

$$E_{\bullet} \ \longrightarrow \ N_{\bullet}$$

is a (weak) homotopy equivalence. But this is a well known fact: $E_{\bullet}$ is obtained from the simplicial set X by *shifting*, it is a sort of path space on X, and it is homotopy equivalent to the *subspace of constant paths*; that is, the set X_0 regarded as a simplicial set in a trivial way. The latter statement is in fact true with the strongest possible interpretation of homotopy equivalence, namely *simplicial homotopy equivalence*. An account can be found in [4, proposition 1.5]; another in [5, lemma 1.5.1].

References.

[1] Daniel M. Kan, A combinatorial definition of homotopy groups, *Ann. of Math.* **67** (1958), 282–312

[2] Klaus Lamotke, Semisimpliziale algebraische Topologie, *Grundlehren der mathematischen Wissenschaften*, **147**, Springer Verlag (1968)

[3] Daniel G. Quillen, The geometric realization of a Kan fibration is a Serre fibration, *Proc. AMS* **19** (1968), 1499–1500

[4] Graeme Segal, Categories and cohomology theories, *Topology* **13** (1974), 293–312

[5] Friedhelm Waldhausen, Algebraic K-theory of spaces, Algebraic and Geometric Topology, *Springer Lecture Notes* **1126** (1983), 318–419

Friedhelm Waldhausen
Fakultät für Mathematik
Universität Bielefeld
Postfach 33501
D-33501 Bielefeld
Germany

fw@mathematik.uni-bielefeld.de

On the Nonexcellence of Field Extensions $F(\pi)/F$

O. T. Izhboldin[1]

Received: December 30, 1995
Revised: February 21, 1996

Communicated by Alexander S. Merkurjev

Abstract. For any $n \geq 3$, we construct a field F and an n-fold Pfister form φ such that the field extension $F(\varphi)/F$ is not excellent. We prove that $F(\varphi)/F$ is universally excellent if and only if φ is a Pfister neighbor of dimension ≤ 4.

Keywords and Phrases: Quadratic forms, Pfister forms, excellent field extensions.

1991 Mathematics Subject Classification: Primary 11E04; Secondary 11E81, 12F20.

Let F be a field of characteristic different from 2 and φ be a non-degenerate quadratic form on an F-vector space V, by which V gets the structure of a non-degenerate quadratic space. Choosing an orthogonal basis of V we can write φ in the form $a_1x_1^2 + \cdots + a_dx_d^2$. In this case we use the notation $\varphi = \langle a_1, \ldots, a_d\rangle$.

A quadratic form or space φ is called *isotropic* if $\varphi(v) = 0$ for some nonzero vector $v \in V$. We say that φ is *anisotropic* otherwise. Up to isometry, there is exactly one non-degenerated isotropic 2-dimensional quadratic space, namely the *hyperbolic plane* $\mathbb{H}$ equipped with the form $\langle 1, -1\rangle$. A non-degenerate quadratic space is called *hyperbolic* if it is isometric to the orthogonal sum of hyperbolic planes $m\mathbb{H} = \mathbb{H} \perp \cdots \perp \mathbb{H}$.

According to Witt's main theorem any non-degenerate quadratic space V can be decomposed in the orthogonal sum $V = V_{an} \perp V_h$, where V_{an} is anisotropic and $V_h \cong m\mathbb{H}$ is a hyperbolic space. (We will use $\cong$ to denote isometry of quadratic forms or spaces.) Moreover the quadratic space V_{an} is uniquely determined up to isometry. The restriction $\varphi|_{V_{an}}$ is called the *anisotropic part* (or anisotropic kernel) of φ and is denoted by φ_{an}. The number $m = \frac{1}{2}\dim V_h$ is called the Witt index of φ.

For any quadratic space V and any field extension L/F one can provide $V_L = V \otimes_F L$ with a structure of a quadratic space. The corresponding quadratic form we shall denote by φ_L. We say that a quadratic form φ over L is defined over F if there is a quadratic form ξ over F such that $\varphi \cong \xi_L$.

[1]This work was partially supported by grant GAP300 of The International Science Foundation.

It is an important problem to study the behavior of the anisotropic part of forms over F under a field extension L/F. It occurs sometimes that any anisotropic form over F is still anisotropic over L (for example if L/F is of odd degree). In this case for any quadratic form φ over F the anisotropic part $(\varphi_L)_{an}$ of φ over L coincides with $(\varphi_{an})_L$ and hence is defined over F.

However, very often φ becomes isotropic over L. In this case we do not know if the anisotropic part of φ over L is defined over F.

A field extension L/F is called *excellent* if for any quadratic form φ over F the anisotropic part $(\varphi_L)_{an}$ of φ over L is defined over F (i.e., there is a form ξ over F such that $(\varphi_L)_{an} \cong \xi_L$).

It is well known that any quadratic extension is excellent. Since any anisotropic quadratic form ψ over F is still anisotropic over the field of rational functions $F(t)$, every purely transcendental field extension is excellent.

Among all field extensions the fields $F(\varphi)$ of rational functions on the quadric hyper-surface defined by the equation $\varphi = 0$ are of special interest in the theory of quadratic forms. One of the important problems is to find a condition on φ so that the field extension $F(\varphi)/F$ is excellent.

We say that $F(\varphi)/F$ is *universally excellent* if for any extension K/F the extension $K(\varphi)/K$ is excellent.

If φ is isotropic then $F(\varphi)/F$ is purely transcendental, and it follows from Springer's theorem that $F(\varphi)/F$ is excellent and moreover is universally excellent. Thus it is sufficient to consider only the case of anisotropic forms φ.

In [Kn] Knebusch has proved that if φ is an anisotropic form such that $F(\varphi)/F$ is excellent then φ is a *Pfister neighbor*. This means that there is a quadratic form $\pi = \langle 1, -a_1\rangle \otimes \cdots \otimes \langle 1, -a_n\rangle$ (called *n-fold Pfister form*) such that φ is similar to a subform of π and $\dim(\varphi) > \frac{1}{2}\dim(\pi)$. This result gives rise to the natural question whether the field extension $F(\varphi)/F$ is excellent for any Pfister neighbor φ. This problem can be easily reduced to the case of an n-fold Pfister forms φ.

If $n = 1$ then $F(\varphi)/F$ is obviously excellent since $F(\varphi)/F$ is a quadratic extension. Arason [ELW1, Appendix II] has proved that, for $n = 2$, $F(\varphi)/F$ is always excellent (see also [R], [LVG]). Thus the answer to our question is yes for n-fold Pfister forms with $n \le 2$. It was an open problem whether $F(\varphi)/F$ is excellent for any field F and any n-fold Pfister form φ over F (with $n \ge 3$).

In [ELW2] some special cases of this problem were considered: for an n-fold Pfister form φ with $n \ge 3$, the excellence of the field extension $F(\varphi)/F$ was proved for all fields with $\tilde{u}(F) \le 4$. In [H2] Hoffmann considered another special case of the problem. An extension L/F is called *d-excellent* if for any quadratic form ψ of dimension $\le d$ the anisotropic part $(\psi_L)_{an}$ of ψ over L is defined over F. Hoffmann has proved that the extension $F(\varphi)/F$ is 6-excellent for any Pfister neighbor φ.

In this paper we prove that for any $n \ge 3$ there is a field F and an n-fold Pfister form φ such that the field extension $F(\varphi)/F$ is not excellent. Moreover Theorem 1.1 of our paper says that $F(\varphi)/F$ is universally excellent if and only if φ is a Pfister neighbor of an n-fold Pfister form with $n \le 2$, (i.e., either $\dim\varphi \le 3$ or φ is a 4-dimensional form with $\det(\varphi) = 1$). In §3 we use the main construction of the paper to study "splitting pairs" φ, ψ of quadratic forms. More precisely, we construct a "non standard pair" φ, ψ such that φ is isotropic over the function field $F(\psi)$ of the quadric ψ.

Remark. Some results of this paper were developed further by D. Hoffmann in [H4].

1. Main Theorem

We will use the following notation throughout the paper: by $\varphi \perp \psi$, $\varphi \cong \psi$, and $[\varphi]$ we denote respectively orthogonal sum of forms, isometry of forms, and the class of φ in the Witt ring $W(F)$ of the field F. The maximal ideal of $W(F)$ generated by the classes of even dimensional forms is denoted by $I(F)$. We write $\varphi \sim \psi$ if φ is similar to ψ, i.e., $k\varphi = \psi$ for some $k \in F^*$. The anisotropic part of φ is denoted by φ_{an} and $i_W(\varphi)$ denotes the Witt index of φ. We denote by $\langle\langle a_1, \dots, a_n \rangle\rangle$ the n-fold Pfister form

$$\langle 1, -a_1 \rangle \otimes \cdots \otimes \langle 1, -a_n \rangle$$

and by $P_n(F)$ the set of all n-fold Pfister forms. The set of all forms similar to n-fold Pfister forms we denote by $GP_n(F)$. For any field extension L/F we put $\varphi_L = \varphi \otimes L$, $W(L/F) = \ker(W(F) \to W(L))$.

Main Theorem 1.1. *Let φ be an anisotropic form over F. Then the following conditions are equivalent.*

(i) *The field extension $F(\varphi)/F$ is universally excellent, i.e., for any field extension E/F the extension $E(\varphi)/E$ is excellent.*
(ii) *Either* $\dim(\varphi) \leq 3$ *or* $\varphi \in GP_2(F)$.

Proof of (ii) $\Rightarrow$ (i). The case $\dim(\varphi) = 2$ is obvious. If $\dim(\varphi) = 3$ or $\varphi \in GP_2(F)$ the excellence of the extension $E(\varphi)/E$ was proved by Arason (see the introduction).

Proof of (i) $\Rightarrow$ (ii). Since $E(\varphi)/E$ is excellent for any extension E/F, we see that $F(\varphi)/F$ is excellent. It was shown in [Kn, 7.13] that for $F(\varphi)/F$ to be excellent it is necessary that φ is a Pfister neighbor. Let φ be a Pfister neighbor of the n-fold Pfister form π. Since $F(\varphi)$ and $F(\pi)$ are F-equivalent, we can replace φ by π, i.e., we can suppose that $\varphi = \pi$ is an n-fold Pfister form. Thus it is sufficient to prove the following proposition.

Proposition 1.2. *Let π be anisotropic n-fold Pfister form over the field F. If $n \geq 3$ then there is a field extension E/F such that $E(\pi)/E$ is not excellent.*

2. Proof of Proposition 1.2

Lemma 2.1. *Let π and τ be anisotropic n-fold Pfister forms over the field F. Then there is a field extension K/F such that the following conditions hold.*

a) $\pi_K = \tau_K$,
b) π_K *and* τ_K *are anisotropic.*

Proof. Let φ be a Pfister neighbor of τ of dimension $2^{n-1} + 1$. It follows from [H3, Theorem 4] that there exists a field extension K/F such that π_K is anisotropic and $\varphi_K \subset \pi_K$. Hence φ_K is a Pfister neighbor of π_K. Since φ_K is a Pfister neighbor of τ_K, we have $\pi_K = \tau_K$. $\square$

Lemma 2.2. *Let τ and π be anisotropic n-fold Pfister forms over F. Suppose that there is $a \in F^*$ such that $\tau_{F(\sqrt{a})}$ and $\pi_{F(\sqrt{a})}$ are isotropic. Then there is an extension E/F and $x \in E^*$ such that the following conditions hold.*

1) $\pi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})}$,
2) $\pi_{E(\sqrt{x})}$ *and* $\tau_{E(\sqrt{x})}$ *are anisotropic,*
3) E/F *is unirational.*

Remark: We say that E/F is unirational, if there is a purely transcendental finitely generated field extension K/F such that $F \subset E \subset K$.

Proof. Since τ is an n-fold Pfister form and $\tau_{F(\sqrt{a})}$ is isotropic, we can write τ in the form $\tau = \langle\langle a, b_1, \dots, b_{n-1}\rangle\rangle$. Similarly, we can write π in the form $\pi = \langle\langle a, c_1, \dots, c_{n-1}\rangle\rangle$. Let $\widetilde{F} = F(A, B_1, \dots, B_{n-1}, C_1, \dots, C_{n-1})$ be the rational function field in $2n-1$ variables over $\widetilde{F}$.

Put $\widetilde{\tau} = \langle\langle A, B_1, \dots, B_{n-1}\rangle\rangle$ and $\widetilde{\pi} = \langle\langle A, C_1, \dots, C_{n-1}\rangle\rangle$. Let $\gamma = \tau \perp -\pi$ and $\widetilde{\gamma} = \widetilde{\tau} \perp -\widetilde{\pi}$. Let $E/\widetilde{F}$ be the universal field extension such that $\gamma_E = \widetilde{\gamma}_E$, i.e., $E = \widetilde{F}_h$, where $\widetilde{F} = \widetilde{F}_0, \widetilde{F}_1, \dots, \widetilde{F}_h$ is a generic splitting tower of the quadratic form $\gamma \perp -\widetilde{\gamma}$.

It is well known that the following universal property of E holds: For any field extension $K/\widetilde{F}$ the condition $\gamma_K = \widetilde{\gamma}_K$ implies that EK/K is purely transcendental.

Now we prove that conditions 1)–3) of the lemma hold for $x = A$.

1) We have $[\tau_{E(\sqrt{A})}] - [\pi_{E(\sqrt{A})}] = [\gamma_{E(\sqrt{A})}] = [\widetilde{\gamma}_{E(\sqrt{A})}] = [\widetilde{\tau}_{E(\sqrt{A})}] - [\widetilde{\pi}_{E(\sqrt{A})}] = 0$. Hence $[\tau_{E(\sqrt{A})}] = [\pi_{E(\sqrt{A})}]$.

2) Let K/F be as in Lemma 2.1, i.e., τ_K, π_K are anisotropic and $\tau_K = \pi_K$. We have $[\gamma_K] = [\tau_K] - [\pi_K] = 0$

Let $\widetilde{K} = K(A, B_1, \dots, B_{n-1}, C_1, \dots, C_{n-1})$ be the rational function field in $2n-1$ variables over K. We have $[\gamma_{\widetilde{K}(\sqrt{A})}] = [\tau_{\widetilde{K}(\sqrt{A})}] - [\pi_{\widetilde{K}(\sqrt{A})}] = 0$ and $[\widetilde{\gamma}_{\widetilde{K}(\sqrt{A})}] = [\widetilde{\tau}_{\widetilde{K}(\sqrt{A})}] - [\widetilde{\pi}_{\widetilde{K}(\sqrt{A})}] = 0$. Therefore $[\gamma_{\widetilde{K}(\sqrt{A})}] = [\widetilde{\gamma}_{\widetilde{K}(\sqrt{A})}]$. Using the universal property of $E/\widetilde{F}$ we see that $E\widetilde{K}(\sqrt{A})/\widetilde{K}(\sqrt{A})$ is purely transcendental.

It is clear that $\widetilde{K}(\sqrt{A})/K$ is purely transcendental. Therefore $E\widetilde{K}(\sqrt{A})/K$ is purely transcendental. Hence $\tau_{E\widetilde{K}(\sqrt{A})}$ and $\pi_{E\widetilde{K}(\sqrt{A})}$ are anisotropic. Therefore $\tau_{E(\sqrt{A})}$ and $\pi_{E(\sqrt{A})}$ are anisotropic.

3) Let $L = \widetilde{F}(\sqrt{A/a}, \sqrt{B_1/b_1}, \dots, \sqrt{B_{n-1}/b_{n-1}}, \sqrt{C_1/c_1}, \dots, \sqrt{C_{n-1}/c_{n-1}})$. It is clear that $\pi_L = \widetilde{\pi}_L$ and $\tau_L = \widetilde{\tau}_L$. Therefore $\gamma_L = \widetilde{\gamma}_L$. Using the universal property of $E/\widetilde{F}$ we see that EL/L is purely transcendental. It is clear that L/F is purely transcendental. Hence EL/F is purely transcendental. Since $E \subset EL$ we see that E/F is unirational. $\square$

Lemma 2.3. *Let F be a field and π be anisotropic n-fold Pfister form over F. Then there are a unirational extension E/F, an n-fold Pfister form τ over E, and $x \in E^*$ such that the following conditions hold.*

1) $\pi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})}$,
2) $\pi_{E(\sqrt{x})}$ *and* $\tau_{E(\sqrt{x})}$ *are anisotropic,*
3) $\dim(\pi_E \perp -\tau_E)_{an} = 2^{n+1} - 4$.

Proof. Write π in the form $\pi = \langle\langle a, b_1, b_2, \ldots, b_{n-1}\rangle\rangle$. Let $\widetilde{F} = F(T_1, \ldots, T_{n-1})$ be the rational function field in $n-1$ variables over F. Let $\tau = \langle\langle a, T_1, \ldots, T_{n-1}\rangle\rangle$. Obviously

$$(\pi_{\widetilde{F}} \perp -\tau)_{an} = \langle\langle a\rangle\rangle\langle\langle b_1, \ldots, b_{n-1}\rangle\rangle'_{\widetilde{F}} \perp -\langle\langle a\rangle\rangle\langle\langle T_1, \ldots, T_{n-1}\rangle\rangle'.$$

Therefore $\dim(\pi_{\widetilde{F}} \perp -\tau)_{an} = 2^{n+1} - 4$.

The quadratic forms $\pi_{\widetilde{F}(\sqrt{a})}$ and $\tau_{\widetilde{F}(\sqrt{a})}$ are hyperbolic, i.e., all the conditions of Lemma 2.2 hold for $\widetilde{F}$, π, τ. Hence there is a unirational extension $E/\widetilde{F}$ such that

1) $\pi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})}$,
2) $\pi_{E(\sqrt{x})}$ and $\tau_{E(\sqrt{x})}$ are anisotropic,

Since $E/\widetilde{F}$ is unirational, we have $\dim(\pi_E \perp -\tau_E)_{an} = \dim(\pi_{\widetilde{F}} \perp -\tau)_{an} = 2^{n+1} - 4$. Finally E/F is unirational since $E/\widetilde{F}$ is unirational and $\widetilde{F}/F$ is purely transcendental. $\square$

Lemma 2.4. *Let E be a field, $n \geq 3$, $x \in E^*$. Let $\pi, \tau \in P_n(E)$ be such that*

1) $\pi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})}$.
2) $\pi_{E(\sqrt{x})}$ *and* $\tau_{E(\sqrt{x})}$ *are anisotropic.*
3) $\dim(\pi \perp -\tau)_{an} = 2^{n+1} - 4$.

Let $\psi = \tau' \perp \langle x\rangle$ where τ' is such that $\tau = \tau' \perp \langle 1\rangle$.

Then

a) *ψ is anisotropic.*
b) *$\psi_{E(\pi)}$ is isotropic.*
c) *There is no quadratic form γ over E such that $(\psi_{E(\pi)})_{an} = \gamma_{E(\pi)}$.*
d) *For any subform $\xi \subsetneq \psi$ the form $\xi_{F(\pi)}$ is anisotropic, i.e., ψ is a minimal $F(\pi)$-form.*

Proof. a) Obviously $\psi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})}$. By assumption we see that $\tau_{E(\sqrt{x})}$ is anisotropic. Hence $\psi_{E(\sqrt{x})}$ is anisotropic. Therefore ψ is anisotropic too.

b) Suppose that $\psi_{E(\pi)}$ is anisotropic. Since $\psi_{E(\sqrt{x})} = \tau_{E(\sqrt{x})} = \pi_{E(\sqrt{x})}$ we have $[\psi_{E(\pi)(\sqrt{x})}] = [\pi_{E(\pi)(\sqrt{x})}] = 0$. Since $\psi_{E(\pi)}$ is anisotropic and $\psi_{E(\pi)(\sqrt{x})}$ is hyperbolic, we conclude that $\psi_{E(\pi)} = \langle\langle x\rangle\rangle\xi$ where ξ is a quadratic form over $E(\pi)$. Since $\dim(\xi) = 2^{n-1}$ is even, we have $\xi \in I(E(\pi))$. Therefore $\psi_{E(\pi)} = \langle\langle x\rangle\rangle\xi \in I^2(E(\pi))$. Hence $\psi \in I^2(E)$. Therefore $[\langle\langle x\rangle\rangle] = [\tau] - [\psi] \in I^2(E)$, a contradiction.

c) Suppose that $(\psi_{E(\pi)})_{an} = \gamma_{E(\pi)}$ where γ is a quadratic form over E. It is clear that $\dim(\gamma) \leq 2^n - 2$. We have $(\psi \perp -\gamma)_{an} \in W(E(\pi)/E)$. Since π is a Pfister form we conclude that $(\psi \perp -\gamma)_{an} = \pi\mu$, with μ a quadratic form over E.

Since $2 = 2^n - (2^n - 2) \leq \dim(\psi \perp -\gamma)_{an} = 2^n + (2^n - 2) = 2^{n+1} - 2$ and $\dim(\pi) = 2^n$ divides $\dim(\pi\mu)$ we conclude that $\dim(\mu) = 1$. Writing μ in the form $\mu = \langle k\rangle$ we have $(\psi \perp -\gamma)_{an} = k\pi$. Hence $[k\pi] = [\psi] - [\gamma]$. Therefore

$$[\tau \perp -k\pi] = [\tau] - [k\pi] = ([\psi] + [\langle\langle x\rangle\rangle]) - ([\psi] - [\gamma]) = [\langle\langle x\rangle\rangle \perp \gamma].$$

Hence τ and $k\pi$ contain a common subform of dimension

$$\frac{1}{2}(\dim(\tau) + \dim(k\pi) - \dim(\langle\langle x\rangle\rangle \perp \gamma)) \geq \frac{1}{2}(2^n + 2^n - 2^n) = 2^{n-1} \geq 2^{3-1} = 4 > 3.$$

Therefore there is a 3-dimensional form ρ such that $\rho \subset \tau$, $\rho \subset k\pi$. Let $a, b \in E$ be such that $\rho \sim \langle 1, -a, -b\rangle$. Let $\varepsilon = \langle\langle a, b\rangle\rangle$. Obviously $\tau_{E(\varepsilon)}$ and $\pi_{E(\varepsilon)}$ are isotropic. Since τ, π, and ε are anisotropic Pfister forms, we conclude that $\varepsilon \subset \tau$ and $\varepsilon \subset \pi$. Therefore $\dim(\pi \perp -\tau)_{an} \leq \dim(\pi) + \dim(\tau) - 2\dim(\varepsilon) = 2^n + 2^n - 2 \cdot 4 = 2^{n+1} - 8$, a contradiction.

d) We can suppose that ξ is a $(2^n - 1)$-dimensional subform of ψ. let $k \in E^*$ be such that $\xi \perp \langle -k\rangle = \psi$. Set $\widetilde{\xi} = \xi \perp \langle -xk\rangle$. We have

$$[\tau] - [\widetilde{\xi}] = [\tau] - ([\xi] - [\langle xk\rangle]) = ([\psi] + [\langle\langle x\rangle\rangle]) - ([\psi] + [\langle k\rangle] - [\langle xk\rangle]) = [\langle\langle x, k\rangle\rangle].$$

Let $\rho = \langle\langle x, k\rangle\rangle$. We have $[\tau_{E(\rho)}] = [\widetilde{\xi}_{E(\rho)}]$. Comparing dimensions we see that $\tau_{E(\rho)} = \widetilde{\xi}_{E(\rho)}$. Therefore $\tau_{E(\rho,\pi)} = \widetilde{\xi}_{E(\rho,\pi)}$.

Our goal is to prove that $\xi_{E(\pi)}$ is anisotropic. Let us suppose that $\xi_{E(\pi)}$ is isotropic. Then $\widetilde{\xi}_{E(\rho,\pi)}$ is isotropic too. Therefore $\tau_{E(\rho,\pi)}$ is isotropic. Hence the Pfister form $\tau_{E(\rho)}$ becomes isotropic over the function field of the Pfister form $\pi_{E(\rho)}$. Therefore either $\tau_{E(\rho)}$ or $\tau_{E(\rho)} = \pi_{E(\rho)}$ is hyperbolic.

Suppose first that $\tau_{E(\rho)}$ is hyperbolic. Since $\rho_{E(\sqrt{x})} = \langle\langle x, k\rangle\rangle_{E(\sqrt{x})}$ is isotropic we conclude that $\tau_{E(\sqrt{x})}$ is isotropic. This contradicts the assumption in this lemma.

Let now $\tau_{E(\rho)} = \pi_{E(\rho)}$. Then $(\tau \perp -\pi)_{an} \in W(E(\rho)/E)$. Hence $(\tau \perp -\pi)_{an} = \rho\lambda$ with λ a quadratic form over E ([S, Ch.4,5.6]). Since $\dim(\tau \perp -\pi)_{an} = 2^n - 4$ and $\dim(\rho) = 4$ we conclude that $\dim(\lambda) = (2^n - 4)/4 = 2^{n-2} - 1$. Since $n \geq 3$ we see that $\dim(\lambda)$ is odd and hence $[\lambda] \equiv [\langle 1\rangle] \pmod{I(E)}$. Since $\rho \in I^2(E)$ we have $[\rho\lambda] \equiv [\rho] \pmod{I^3(E)}$. Since $\tau, \pi \in P_n(E)$ and $n \geq 3$, we see that $[(\tau \perp -\pi)_{an}] \equiv 0 \pmod{I^3(E)}$. We have

$$[\rho] \equiv [\rho\lambda] = [(\tau \perp -\pi)_{an}] \equiv 0 \pmod{I^3(E)}.$$

Since $\dim(\rho) = 4 < 8$ we conclude that ρ is hyperbolic. Therefore $(\tau \perp -\pi)_{an} = \rho\lambda$ is hyperbolic. However $\dim(\tau \perp -\pi)_{an} = 2^n - 4 > 0$, a contradiction. □

Corollary 2.5. *Let π be an anisotropic n-fold Pfister form over the field F. If $n \geq 3$ then there is a unirational extension E/F such that $E(\pi)/E$ is not excellent.* □

This corollary completes the proof of Proposition 1.2 and Theorem 1.1.

Corollary 2.6. *Let $n \geq 3$. Then there are a field E, an n-fold Pfister form π over E, and a 2^n-dimensional form ψ over E such that ψ is an $E(\pi)$-minimal form.* □

Corollary 2.7. *Let $n \geq 3$. Then there are a field E and 2^n-dimensional forms ψ and π over E such that ψ is an $E(\pi)$-minimal form and ψ is not similar to π.* □

3. Nonstandard Splitting

An important problem in the theory of quadratic forms is to determine when an anisotropic quadratic form φ over F becomes isotropic over the function field $F(\psi)$ of another form ψ. There are some well-known situations when this occurs and we list some of them in the following two definitions.

Definition 3.1. Let φ and ψ be anisotropic quadratic forms. We say that the ordered pair φ, ψ is *elementary splitting* (or *elementary*) if one of the following conditions holds.

1) There is a $k \in F^*$ such that $k\psi \subset \varphi$;
2) There is a $k \in F^*$, such that $k\varphi \subset \psi$ and $\dim(\varphi) > \dim(\psi) - i_1(\psi)$;
3) There is a $\rho \in W(F(\psi)/F)$ such that $\dim(\rho) < 2\dim(\varphi)$ and $k\varphi \subset \rho$ for some $k \in F^*$.

Definition 3.2. Let φ and ψ be anisotropic quadratic forms. We say that the ordered pair φ, ψ is standard if there is a collection

$$\varphi_0 = \varphi, \varphi_1, \ldots, \varphi_{n-1}, \varphi_n = \psi$$

such that the pair φ_{i-1}, φ_i is elementary for each $i = 1, 2, \ldots, n$.

It is clear that if the pair (φ, ψ) is elementary splitting or standard, then $\varphi_{F(\psi)}$ is isotropic.

Examples 3.3. *Let φ and ψ be anisotropic quadratic forms such that $\varphi_{F(\psi)}$ is isotropic. Suppose that at least one of the following conditions holds*

a) *φ is a Pfister neighbor;*
b) *$\dim(\psi) \leq 3$, or $\psi \in GP_2(F)$;*
c) *$\dim(\varphi) \leq 5$;*

Then the pair φ, ψ is elementary.

Proof. a) Let φ be a Pfister neighbor of ρ. Then condition 3) of Definition 3.1 is fulfilled.

b) By the excellence property of the field extension $F(\psi)/F$ there exists an anisotropic form ξ over F such that $(\varphi_{F(\psi)})_{an} = \xi_{F(\psi)}$. Setting $\rho = \varphi \perp -\xi$ one can see that condition 3) of Definition 3.1 holds.

c) Let $\dim(\varphi) \leq 5$. We can suppose that φ is not a Pfister neighbor and $\psi \notin GP_2(F)$ (see a), b)). Then $\varphi_{F(\psi)}$ is isotropic if and only if φ contains a subform similar to ψ (see [H1, Th. 1, Main Theorem]). Therefore condition 1) of Definition 3.1 holds. $\square$

Example 3.4. *Let $F = \mathbb{R}(T)$, $\varphi = \langle T, T, T, 1, 1, 1, 1, 1\rangle$, $\psi = \langle T, T, 1, 1, 1, 1, 1, 1\rangle$. Then the pair φ, ψ is standard but not elementary.*

Proof. Let $\rho = \langle T, T, 1, 1, 1, 1, 1\rangle$. Since $\rho \subset \varphi$, the pair (φ, ρ) is elementary. Since $\rho \subset \psi$ and $\dim(\rho) = 7 > 8 - 2 = \dim(\psi) - i_1(\psi)$, we see that the pair (ρ, ψ) is elementary. Since the pairs (φ, ρ) and (ρ, ψ) are elementary, we see that the pair (φ, ψ) is standard. It follows from Lemma 3.7 below that the pair (φ, ψ) is not elementary. $\square$

In this section we construct a pair of anisotropic forms φ and ψ with $\varphi_{F(\psi)}$ isotropic which is not standard.

Lemma 3.5. *Let F be a field, $n \geq 3$, $x \in F^*$. Let $\pi, \tau \in P_n(F)$ be such that*

1) *$\pi \neq \tau$,*
2) *$\pi_{F(\sqrt{x})} = \tau_{F(\sqrt{x})}$,*
3) *$\pi_{F(\sqrt{x})}$ and $\tau_{F(\sqrt{x})}$ are anisotropic.*

Let $\varphi = \pi' \perp \langle x\rangle$ and $\psi = \tau' \perp \langle x\rangle$. Then

a) *ψ and φ are anisotropic,*
b) *$\varphi_{F(\psi)}$ and $\psi_{F(\varphi)}$ are isotropic,*
c) *$\varphi \not\sim \psi$.*

Proof. a) Obviously $\psi_{F(\sqrt{x})} = \pi_{F(\sqrt{x})}$ and $\psi_{F(\sqrt{x})} = \tau_{F(\sqrt{x})}$. It follows from condition 3) that φ and ψ are anisotropic.

b) Let us suppose that $\varphi_{F(\psi)}$ is anisotropic. Since $\varphi_{F(\sqrt{x})} = \pi_{F(\sqrt{x})}$ and $\psi_{F(\sqrt{x})} = \tau_{F(\sqrt{x})} = \pi_{F(\sqrt{x})}$ we see that $\varphi_{F(\psi,\sqrt{x})} = \pi_{F(\pi,\sqrt{x})}$. Since $\pi \in P_n(F)$ we conclude that $\varphi_{F(\psi,\sqrt{x})}$ is hyperbolic. Therefore $\varphi_{F(\psi)} = \langle\langle x\rangle\rangle\xi$ where ξ is a quadratic form over $F(\psi)$. Since $\dim(\xi) = 2^{n-1}$ is even, we have $\xi \in I(F(\psi))$. Therefore $\psi_{F(\psi)} = \langle\langle x\rangle\rangle\xi \in I^2(F(\psi))$. Hence $\psi \in I^2(F)$. Therefore $[\langle\langle x\rangle\rangle] = [\tau] - [\psi] \in I^2(F)$, a contradiction.

c) Suppose that $k\varphi = \psi$. Then $[k\pi] - [k\langle\langle x\rangle\rangle] = [k\varphi] = [\psi] = [\tau] - [\langle\langle x\rangle\rangle]$. Therefore $[\langle\langle x, k\rangle\rangle] = [\tau] - [k\pi] \in I^n(F) \subset I^3(F)$. Since $\dim(\langle\langle x, k\rangle\rangle) = 4 < 8$, we have $[\tau] - [k\pi] = [\langle\langle x, y\rangle\rangle] = 0$. Hence $\tau \sim \pi$. Since $\tau, \pi \in P_n(F)$ we see that $\tau = \pi$, a contradiction. $\square$

Lemma 3.6. *Let $\pi \in P_3(F)$ and $x \in F^*$ ($x \notin F^{*2}$) be such that $\pi_{F(\sqrt{x})}$ is anisotropic. Let $\varphi = \pi' \perp \langle x\rangle$. Suppose that ψ is an anisotropic quadratic form such that $\psi_{F(\varphi)}$ and $\varphi_{F(\psi)}$ are isotropic. Then $\dim(\psi) = 8$.*

By $C(\varphi)$ (resp. $C_0(\varphi)$) we will denote the Clifford algebra (resp. even Clifford algebra) of the quadratic form φ. If they are central simple we denote their classes in the Brauer group of the underlying field by $[C(\varphi)]$ (resp. $[C_0(\varphi)]$).

Proof. Since $\dim(\varphi) = 8$ and $\varphi_{F(\psi)}$ is isotropic, it follows from Hoffmann's theorem [H3, §1, Theorem 1] that $\dim(\psi) \leq 8$.

Suppose that $\dim(\psi) \leq 6$. Since $\dim(\varphi) = 8$ and $\psi_{F(\varphi)}$ is isotropic, it follows from Hoffmann's theorems [H1], [H2] that $\varphi \in GP_3(F)$. Therefore $x = \det(\varphi) = 1$, a contradiction.

Consider now the case $\dim(\psi) = 7$. Since $\pi_{F(\psi,\sqrt{x})} = \varphi_{F(\psi,\sqrt{x})}$ is isotropic we see that $\psi_{F(\sqrt{x})}$ is a Pfister neighbor of $\pi_{F(\sqrt{x})}$. Therefore $[C_0(\psi)_{F(\sqrt{x})}] = 0$. Hence there is $y \in F^*$ such that $[C_0(\psi)] = [\left(\frac{x,y}{F}\right)]$. Let $\rho = \langle\langle x, y\rangle\rangle$.

We claim that $\psi_{F(\rho)}$ is an anisotropic Pfister neighbor. To prove this we consider the quadratic form $\widetilde{\psi} = \psi \perp \langle\det(\psi)\rangle$. Since $\dim(\widetilde{\psi}) = 8$ and $[C(\widetilde{\psi}_{F(\rho)})] = [\left(\frac{x,y}{F(\rho)}\right)] = 0$ we have $\widetilde{\psi}_{F(\rho)} \in GP_3(F(\rho))$. If $\psi_{F(\rho)}$ is isotropic then $\widetilde{\psi}_{F(\rho)}$ is isotropic too and hence hyperbolic. Therefore, $(\widetilde{\psi})_{an} = \rho\mu$. Since $\dim(\widetilde{\psi}) = 6$ or 8 we must have $\dim \mu = 2$ which implies $\widetilde{\psi}_{an} = \widetilde{\psi} \in GP_3(F)$. Therefore $[C(\rho)] = [C_0(\psi)] = [C(\widetilde{\psi})] = 0$. Hence, ρ is hyperbolic and ψ stays anisotropic over $F(\rho)$, a contradiction.

Since $\psi_{F(\varphi)}$ is isotropic, $\psi_{F(\rho)}$ becomes isotropic over the functional field of the form $\varphi_{F(\rho)}$. Since $\psi_{F(\rho)}$ is an anisotropic Pfister neighbor and $\dim(\varphi_{F(\rho)}) = 8$ we see that $\varphi_{F(\rho)} \in GP_3(F(\rho)) \subset I^2(F(\rho))$. Since $W(F)/I^2(F) \to W(F(\rho))/I^2(F(\rho))$ is injective we have $\varphi \in I^2(F)$. Hence $x = \det(\varphi) = 1$, a contradiction. $\square$

Lemma 3.7. *Let φ and ψ be anisotropic 8-dimensional quadratic form such that $\psi \notin GP_3(F)$ and the pair φ, ψ is elementary. Then $\varphi \sim \psi$.*

Proof. Since the pair φ, ψ is elementary, one of conditions 1)–3) of Definition 3.1 holds. Since $\dim(\varphi) = \dim(\psi)$, both the conditions 1), 2) imply that $\varphi \sim \psi$. Now we suppose that condition 3) holds, i.e., there is $\rho \in W(F(\psi)/F)$ such that $\dim(\rho) < 2\dim(\varphi) = 16$ and $k\varphi \subset \rho$. Since $\dim(\psi) > 4$, the homomorphism $W(F)/I^3(F) \to W(F(\psi))/I^3(F(\psi))$ is injective. Hence $\rho \in I^3(F)$. Let $\sigma \in P_2(F)$ be such that ψ contains a Pfister neighbor of σ. Then $\rho \in W(F(\psi)/F) \subset W(F(\sigma)/F)$ and thus $\rho_{an} \cong \sigma\mu$ for some μ. If $\dim\mu$ is odd then $\sigma \equiv \sigma\mu = \rho \equiv 0 \pmod{I^3(F)}$, a contradiction. Thus $\dim\mu$ is even and $8|\dim(\rho_{an})$. Therefore $\dim(\rho_{an}) = 8$. Hence $\rho_{an} \in GP_3(F)$. Since $\rho_{F(\psi)}$ is hyperbolic, ψ is a Pfister neighbor in ρ_{an}. Since $\dim(\psi) = \dim(\rho_{an}) = 8$ we have $\psi \sim \rho_{an} \in GP_3(F)$, a contradiction. $\square$

Lemma 3.8. *Let $n = 3$, and let φ, ψ be as in Lemma 3.5. Then the pair φ, ψ is not standard.*

Proof. Assume that the pair φ, ψ is standard. Then there is a collection

$$\varphi_0 = \varphi, \varphi_1, \ldots, \varphi_{n-1}, \varphi_n = \psi$$

such that the pair φ_{i-1}, φ_i is elementary for each $i = 1, 2, \ldots, n$. Obviously, the quadratic forms $\varphi_{F(\varphi_i)}$ and $(\varphi_i)_{F(\psi)}$ are isotropic. Since $\psi_{F(\varphi)}$ is isotropic (see Lemma 3.5) and $(\varphi_i)_{F(\psi)}$ is isotropic, we see that $(\varphi_i)_{F(\varphi)}$ is isotropic too. Thus $\varphi_{F(\varphi_i)}$ and $(\varphi_i)_{F(\varphi)}$ are isotropic. It follows from Lemma 3.6 that $\dim(\varphi_i) = 8$.

Consider first the case $\psi_i \in GP_3(F)$. Since $(\varphi_i)_{F(\varphi)}$ and is isotropic, φ is a Pfister neighbor of ψ_i. Since $\dim(\varphi) = \dim(\psi_i) = 8$ we have $\varphi \sim \psi_i$. Hence $\varphi \in GP_3(F)$, a contradiction.

Thus we have proved that $\dim(\varphi_i) = 8$ and $\psi_i \notin GP_3(F)$ for each $i = 1, 2, \ldots, n$. It follows from Lemma 3.7 that $\varphi_{i-1} \sim \varphi_i$. We have

$$\varphi = \varphi_0 \sim \varphi_1 \sim \cdots \sim \varphi_n = \psi.$$

On the other hand, it follows from Lemma 3.5 that $\varphi \not\sim \psi$. The contradiction obtained proves the lemma. $\square$

Theorem 3.9. *For any field F there is a unirational field extension E/F and a pair of 8-dimensional anisotropic quadratic forms φ and ψ over E such that $\varphi_{E(\psi)}$ is isotropic, but the pair φ, ψ is not standard.*

Proof. Let $n = 3$. Let E, π and τ be such as in Lemma 2.3. Set $\varphi = \pi' \perp \langle x\rangle$, $\psi = \tau' \perp \langle x\rangle$. It is clear that all the conditions of Lemma 3.5 hold. Now the desired result follows immediately from Lemma 3.5 and Lemma 3.8. $\square$

References

[ELW1] Elman, R.; Lam, T.Y.; Wadsworth, A.R., *Amenable fields and Pfister extensions.*, Proc. of Quadratic Forms Conference (ed. G. Orzech). **46** (1977), Queen's Papers in Pure and Applied Mathematics, 445–491.

[ELW2] Elman, R.; Lam, T.Y.; Wadsworth, A.R., *Function fields of Pfister forms*, Invent.Math. **51** (1979), 61–75.

[H1] Hoffmann, D. W., *Isotropy of 5-dimensional quadratic forms over the function field of a quadric.*, Proc. Symp. Pure Math. **58.2** (1995), 217–225.

[H2] Hoffmann, D. W., *On 6-dimensional quadratic forms isotropic over the function field of a quadric.*, Comm. Alg. **22** (1994), 1999-2014.

[H3] Hoffmann, D. W., *Isotropy of quadratic forms over the function field of a quadric.*, Math. Z. **220** (1995), 461–466.
[H4] Hoffmann, D. W., *Twisted Pfister Forms*, Doc. Math. J. DMV **1** (1996), 67-102.
[Kn] Knebusch M., *Generic splitting of quadratic forms, II*, Proc. London Math. Soc. **34** (1977), 1–31.
[LVG] Lewis, D.W.; Van Geel, J., *Quadratic forms isotropic over the function field of a conic*, Indag. Math. **5** (1994), 325–339.
[R] Rost, M., *Quadratic forms isotropic over the function field of a conic*, Math. Ann. **288** (1990), 511–513.
[S] Scharlau, W., *Quadratic and Hermitian Forms*, Springer, Berlin, Heidelberg, New York, Tokyo, 1985.

O. T. Izhboldin
Department of Mathematics and Mechanics
St.-Petersburg State University
Petrodvorets, 198904, Russia

Two Interesting Oriented Matroids

Jürgen Richter-Gebert[1]

Received: February 2, 1996
Revised: February 19, 1996

Communicated by Günter M. Ziegler

Abstract. Oriented matroids are a combinatorial model for configurations in real vector spaces. A central role in the theory is played by the realizability problem: Given an oriented matroid, find an associated vector configuration. In this paper we present two closely related oriented matroids Ω_{14}^{+} and Ω_{14}^{-} of rank 3 with 14 elements that have interesting properties with respect to realizability. Ω_{14}^{+} and Ω_{14}^{-} differ in exactly one basis orientation.

The realizable oriented matroid Ω_{14}^{+} has at least two interesting properties: First it has a combinatorial symmetry that has no metric realization, and second it has a disconnected realization space. In other words, there are different realizations of Ω_{14}^{+} that cannot be continuously deformed into each other while staying in the same isotopy class. The oriented matroid Ω_{14}^{-} is non-realizable but it has no bi-quadratic final polynomial. In other words, the only known effective algorithmic method fails to prove the non-realizability of Ω_{14}^{-}.

1991 Mathematics Subject Classification: Primary 52B40; Secondary 14P10, 51A25, 52B30.

1 Introduction

Oriented matroids are combinatorial models for vector configurations in vector spaces over ordered fields. They form a basic combinatorial concept for treating many different objects on the borderline of combinatorics and geometry — such as convex polytopes, simplicial complexes, hyperplane-arrangements, quasi-crystals, etc. The *realizability question* is of fundamental importance in this theory: *When does a discrete structure have a geometric representation? What does the space of all representations look like?* Questions of this type occur in many different mathematical contexts (e.g. embedding of polyhedral manifolds, the theory of moduli spaces, Cairns' smoothing theory, etc.). The basic effects that arise here are often due to the properties of the

[1]Supported by a DFG Gerhardt-Hess-Forschungsförderungspreis awarded to G.M. Ziegler

underlying oriented matroids, and they can be profitably studied in this model. A systematic study of "small" oriented matroids that have interesting behavior with respect to realizability is a fruitful source for producing examples and counterexamples in many different mathematical disciplines. Here we present two new such oriented matroids.

Every vector configuration has an associated oriented matroid, but the converse is not true: there are oriented matroids that have no corresponding vector configuration. An oriented matroids is *realizable* if it corresponds to a vector configuration, and *non-realizable* otherwise. In this paper we present two closely related oriented matroids Ω_{14}^{+} and Ω_{14}^{-} of rank 3 with 14 elements that are interesting because of their properties with respect to realizability.

The oriented matroid Ω_{14}^{+} is realizable, but its realization space is not connected. The *realization space* of an oriented matroid χ is the set of all vector configurations X that have the associated oriented matroid χ, modulo linear equivalence. (For a more formal definition of realization spaces see Section 2). For a long time it was an outstanding open question whether oriented matroids with disconnected realization space exist. This problem was solved by N.E. Mnëv in a surprising way [6, 7]. He proved that for any basic semi-algebraic set V (defined over the rationals) there is an oriented matroid whose realization space is stably equivalent (in the sense of [9]) to V. Thus realization spaces can be homotopy equivalent to any finite simplicial complex (in particular they may have an arbitrary number of connected components). The examples produced by Mnëv's method in general involve a large number of points. At the same time P.Y. Suvorov [12] constructed an example of rank 3 with disconnected realization space that contains only 14 elements.

The oriented matroid Ω_{14}^{+} shares these properties with Suvorov's example, but it has the following additional nice properties:

- Ω_{14}^{+} is constructible. (After fixing the position of the points $x_1, \dots, x_4$ that form a projective basis and choosing a point $x_5 = (t+1)x_3 + (t-1)x_4$ each point x_i for $i = 6, \dots, 14$ is of the form $(x_a \vee x_b) \wedge (x_c \vee x_d)$ where "$\vee$" is the *join* operator and "$\wedge$" is the *meet* operator and a, b, c, d are indices that are smaller than i.)
- up to stable equivalence (see [9]) the realization space of Ω_{14}^{+} is an open interval from which one point has been deleted.
- Ω_{14}^{+} has rational realizations.
- Ω_{14}^{+} has a combinatorial symmetry of order two that has no metric realization. (The smallest example with this property, known so far, with 90 points, was constructed by P. Shor [11].)

It is still an open question whether there exists an oriented matroid with disconnected realization space and less than 14 points.

If we switch the orientation of one particular basis in Ω_{14}^{+} we obtain the *non-realizable* oriented matroid Ω_{14}^{-}. This oriented matroid has a remarkable property. It is the first known example of a non-realizable oriented matroid for which non-realizability cannot be proved by a bi-quadratic final polynomial.

Final polynomials [3, 5] are certificates for the non-realizability of matroids and oriented matroids. However, no algorithmic method for computing final polynomials is known to be both generally applicable and effective. Indeed, this is not surprising since the realizability problem is known to be NP-hard [11]. Bi-quadratic final polynomials (as introduced in [2] and [8]) are special kinds of final polynomials which can be computed very efficiently. The method of bi-quadratic final polynomials for the oriented matroid case was originally inspired by J. Bokowski [5], who suggested that one consider only inequalities of the form $[\ldots][\ldots] < [\ldots][\ldots]$ which are consequences of three-term Graßmann-Plücker polynomials and the signature of the oriented matroid. These inequalities have to be satisfied in the realizable case. If this system of these inequalities is inconsistent one has a bi-quadratic final polynomial. Deciding whether an oriented matroid has a bi-quadratic final polynomial can be translated into an LP-feasibility-problem and therefore solved in polynomial time. This is the first example of a non-realizable oriented matroid which cannot be certified to be non-realizable by a bi-quadratic final polynomial.

2 Realization spaces

Oriented matroids are combinatorial models for vector configurations in linear vector spaces over ordered fields. For an extensive introduction into oriented matroid theory we recommend [1] and [10]. Throughout the paper we will restrict ourselves to the case of vector configurations in $\mathbb{R}^3$, the case of oriented matroids of rank 3. Let $X = (x_1, \ldots, x_n) \in \mathbb{R}^{3n}$ be a configuration consisting of n vectors in $\mathbb{R}^3$. We set $E = \{1, \ldots, n\}$. To every triple of indices $(i, j, k) \in E^3$ we assign a sign

$$\chi_X(i, j, k) = \text{SIGN DET}(x_i, x_j, x_k).$$

The map $\chi_X : E^3 \to \{-1, 0, +1\}$ is called the oriented matroid of X. We omit the general definition of an oriented matroid (it can be found in [1] and [10]).

For us it is sufficient to know that an oriented matroid $\chi : E^3 \to \{-1, 0, +1\}$ is a sign map that models the combinatorial behavior of signs of determinants. In particular χ always satisfies the alternating determinant rules:

$$\chi(i, j, k) = \chi(k, i, j) = \chi(j, k, i) = -\chi(j, i, k) = -\chi(k, j, i) = -\chi(i, k, j).$$

Since χ is alternating it is sufficient to specify χ on the set

$$\Lambda(E, 3) = \{(i, j, k) \in E^3 \mid i < j < k\}.$$

An oriented matroid χ is *realizable* if there is a vector configuration X with $\chi_X = \chi$. If there is no such vector configuration, then χ is called *non-realizable.* Deciding the question whether an oriented matroid is realizable or not algorithmically is known to be an NP-hard problem [11].

For a realizable oriented matroid one is often interested not only in a particular realization, but also in the space of all realizations. There are various ways of describing this space, depending on how much of the actions on $\mathbb{R}^{3n}$ that preserve the oriented matroid of X are factored out. If at least a linear basis is fixed all these descriptions turn out to be isomorphic up to stable equivalence (compare [9]). We here use the version where a projective basis is fixed. Reorientation of a point i (i.e. reversing all signs $\chi(a,b,c)$ with $i \in \{a,b,c\}$) does not change the behavior of χ with respect to realizability: if $X = (x_1,\dots,x_n)$ is a realization of χ then we get a realization of the reversed situation if we replace x_i by $-x_i$. Hence, we may (up to relabeling, reorientation of points 1, 2, 3 or 4 and the assumption that χ has at least four points in general position) assume that we have $\chi(1,2,3) = \chi(1,2,4) = \chi(1,3,4) = \chi(2,3,4) = 1$.

Definition 2.1. Let $\chi\colon E^3 \to \{-1,0,+1\}$ be a rank 3 oriented matroid that satisfies $\chi(1,2,3) = \chi(1,2,4) = \chi(1,3,4) = \chi(2,3,4) = 1$. Let $x_1 = (1,0,0)$, $x_2 = (0,1,0)$, $x_3 = (1,0,1)$, and $x_4 = (0,1,1)$. The *realization space* of χ is the set of all $(x_5,\dots,x_n) \in \mathbb{R}^{3(n-4)}$ with $\chi_X = \chi$ for $X = (x_1,\dots,x_n)$.

3 Ω_{14}^+ has disconnected realization space

The configuration that we will study here is defined by the following construction sequence. The oriented matroid Ω_{14}^+ is the underlying oriented matroid for choices of the parameter t in $(-3+\sqrt{8},\ 0) \cup (0,\ 3-\sqrt{8})$.

$$
\begin{array}{lll}
x_1 & = (1,0,0), & \\
x_2 & = (0,1,0), & \\
x_3 & = (1,0,1), & \\
x_4 & = (0,1,1), & \\
x_5 & = (1-t)x_3 + (1+t)x_4, & \\
x_6 & = x_5x_2 \wedge x_1x_4 & = (1-t,2,2), \\
x_7 & = x_5x_1 \wedge x_2x_3 & = (-2,-1-t,-2), \\
x_8 & = x_6x_3 \wedge x_5x_1 & = (3-2t-t^2, 2+2t, 4), \\
x_9 & = x_7x_4 \wedge x_5x_2 & = (2-2t, 3+2t-t^2, 4), \\
x_{10} & = x_3x_4 \wedge x_8x_2 & = (-3+2t+t^2, -1-2t-t^2, -4), \\
x_{11} & = x_3x_4 \wedge x_9x_1 & = (-1+2t-t^2, -3-2t+t^2, -4), \\
x_{12} & = x_7x_{10} \wedge x_{11}x_2 & = (1-2t^2+t^4, -1+4t+10t^2+4t^3-t^4, 4+8t+4t^2), \\
x_{13} & = x_6x_{11} \wedge x_{10}x_1 & = (-1-4t+10t^2-4t^3-t^4, 1-2t^2+t^4, 4-8t+4t^2), \\
x_{14} & = x_1x_3 \wedge x_2x_4 & = (0,0,1)
\end{array}
$$

Here $x_\alpha x_\beta$ denotes the "join" of x_α and x_β, and $a \wedge b$ denotes the "meet". Both operations can be computed in terms of the standard cross-product in $\mathbb{R}^3$.

After fixing a projective basis consisting of the points $x_1,\dots,x_4$ the whole construction only depends on the choice of the parameter t. The following matrix gives coordinates for the situation $t = 0$ (the situation where x_5 is in the middle of x_3 and x_4).

$$
X_0 = \begin{pmatrix}
1 & 0 & 1 & 0 & 1 & 1 & 2 & 3 & 2 & 3 & 1 & 1 & -1 & 0 \\
0 & 1 & 0 & 1 & 1 & 2 & 1 & 2 & 3 & 1 & 3 & -1 & 1 & 0 \\
0 & 0 & 1 & 1 & 2 & 2 & 2 & 4 & 4 & 4 & 4 & 4 & 4 & 1
\end{pmatrix}
$$

We can visualize the situation if we normalize the last coordinate for $x_3,\dots,x_{14}$ to 1 by multiplying each vector with a suitable positive scalar. The situation in the plane $\{(x,y,1) \mid x,y \in \mathbb{R}\}$ gives an affine image of our vector configuration in $\mathbb{R}^3$. Figure 1 shows the affine situation for a value t slightly smaller than zero. The points x_1 and x_2 are the points at infinity that lie on the x-axis and y-axis. The little displacement of x_5 away from the symmetric position forces that the lines $(1,3)$, $(2,4)$ and $(12,13)$ not to be concurrent (as in the case $t = 0$).

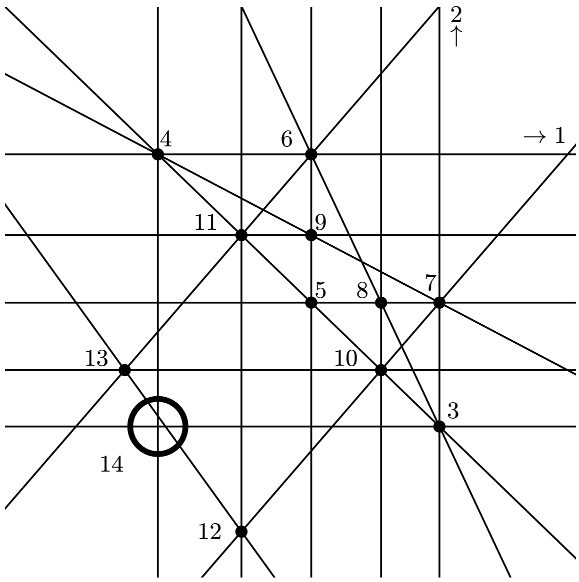

Figure 1

The whole construction sequence has a combinatorial symmetry that is induced by the permutation

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 2 & 1 & 4 & 3 & 5 & 7 & 6 & 9 & 8 & 11 & 10 & 13 & 12 & 14 \end{pmatrix}.$$

Evaluating the determinant $\textsc{det}(x_{12}, x_{13}, x_{14})$ we get

$$\textsc{det}(x_{12}, x_{13}, x_{14}) = 32\,t^2 - 64\,t^4 + 32\,t^6 = 32t^2(t^2-1)^2,$$

a polynomial that has a root which is actually a minimum at $t = 0$.

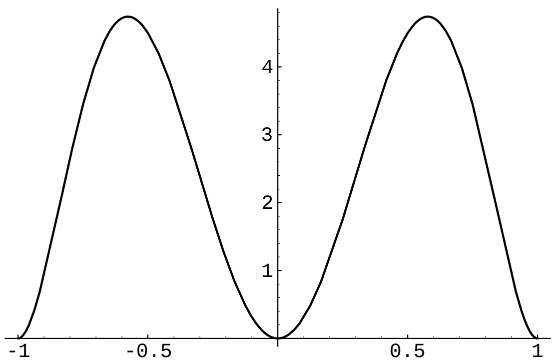

Figure 2

The fact that this polynomial is symmetric in t is already a consequence of the symmetry of the underlying construction of the configuration and of the symmetric choice of our basis $x_1, \ldots, x_4$. A graph of this polynomial is given in Figure 2.

We now define for all $(i, j, k) \in \Lambda(\{1, \ldots, 14\}, 3)$ and $\sigma \in \{-1, 0, +1\}$

$$\Omega_{14}^{\sigma}(i,j,k) := \begin{cases} \sigma & \text{if } (i,j,k) = (12,13,14), \\ \chi_{X_0}(i,j,k) & \text{otherwise.} \end{cases}$$

The oriented matroids Ω_{14}^{σ} have a combinatorial symmetry which is induced by π. For all $(i, j, k) \in \Lambda(\{1, \ldots, 14\}, 3)$ and $\sigma \in \{-1, 0, +1\}$ we have

$$\Omega_{14}^{\sigma}(\pi(i), \pi(j), \pi(k)) = -\Omega_{14}^{\sigma}(i, j, k).$$

A realization X of Ω_{14}^{σ} is *symmetric* if there is a linear involution $R: \mathbb{R}^3 \to \mathbb{R}^3$ with $R(x_i) = x_{\pi(i)}$ for $i \in \{1, \ldots, 14\}$.

Theorem 3.1. *The oriented matroids Ω_{14}^{σ} have the following properties:*

(i) *There is a polynomial function f from $((0,1) \setminus \{\frac{1}{2}\}) \times (0, \infty)^{10}$ to the realization space of Ω_{14}^{+} that is an isomorphism of semi-algebraic sets.*

(ii) *Ω_{14}^{+} has no symmetric realization.*

(iii) *Ω_{14}^{+} has rational realizations.*

(iv) *Ω_{14}^{-} is not realizable.*

Proof. The construction sequence at the beginning of this section shows that after the choice of the parameter t all points are determined up to multiplication by a positive number. The signs that are identical in Ω_{14}^{+}, Ω_{14}^{0}, and Ω_{14}^{-} are exactly taken for values of t in the open interval $(-3+\sqrt{8}, 3-\sqrt{8})$. (The basis that collapse at the end points of this open interval are (x_1, x_3, x_{12}) and (x_2, x_4, x_{13}).) We get realizations of Ω_{14}^{+} exactly for all choices of t in $I = (-3+\sqrt{8}, 0) \cup (0, 3-\sqrt{8})$. For $t = 0$ we get a realization of Ω_{14}^{0}. The factor $(0, \infty)^{10}$ in (i) is due to the fact that multiplication of any of the points $x_5, \ldots, x_{14}$ by a positive scalar does not change the underlying oriented matroid. This proves (i).

Assume that there was a symmetric realization X of Ω_{14}^{+}. After a suitable projective transformation we may assume that $x_1, \ldots, x_4$ are located at $(1,0,0)$, $(0,1,0)$, $(1,0,1)$, $(0,1,1)$, respectively, and that the reflection R is given by $R(x,y,z) = (y,x,z)$. Since x_5 is a fix-point of R it must be of the form $(x,x,z) \neq (0,0,0)$. Up to a positive multiple the only possible choice for x_5 is induced by $t = 0$ in our construction sequence. For $t = 0$ the determinant $\mathrm{DET}(x_{12}, x_{13}, x_{14})$ evaluates to zero. Hence, there is no symmetric realization. This proves (ii).

If we choose t as a rational number in $(-3+\sqrt{8}, 0) \cup (0, 3-\sqrt{8})$ we get a rational realization, as stated in (iii). Fact (iv) is a direct consequence of the fact that for $t \in (-3+\sqrt{8}, 3-\sqrt{8})$ the determinant $\mathrm{DET}(x_{12}, x_{13}, x_{14})$ is always positive or zero. $\square$

4 Final polynomials

Bi-quadratic final polynomials [2, 8] are special final polynomials that can be found by linear programming. They provide an effective tool to prove non-realizability for a large class of oriented matroids. Here we restrict ourselves to the case of realizability over $\mathbb{R}$ and to the case of oriented matroids in rank 3 on a ground set $E = \{1, \ldots, n\}$. Our starting point is the structure of three-term Graßmann-Plücker polynomials. For this the *brackets* $[i, j, k]$ with $i, j, k \in E$ are considered as formal variables. We identify brackets according to the alternating determinant rules:

$$[i, j, k] = [k, i, j] = [j, k, i] = -[j, i, k] = -[k, j, i] = -[i, k, j].$$

The polynomial ring in all brackets $\mathbb{R}[\{[\lambda] \mid \lambda \in E^3\}]$ modulo these identifications is abbreviated $B_{3,n}$. (This is a polynomial ring in $\binom{n}{3}$ generators.) For an oriented matroid χ and a bracket monomial $[\lambda^1] \cdot [\lambda^2] \cdot \ldots \cdot [\lambda^k]$ we write

$$\chi([\lambda^1] \cdot [\lambda^2] \cdot \ldots \cdot [\lambda^k]) := \chi(\lambda^1) \cdot \chi(\lambda^2) \cdot \ldots \cdot \chi(\lambda^k).$$

For a vector configuration $X = (x_1, \ldots, x_n) \in \mathbb{R}^{3n}$ and $(i, j, k) \in E^3$ we write

$$[i, j, k]_X = \mathrm{DET}(x_i, x_j, x_k).$$

Definition 4.1. Let χ be a rank 3 oriented matroid on a finite set E of cardinality $n > 3$, let $\tau \in E, \lambda = (a, b, c, d) \in E^4$ with $|\{\tau, a, b, c, d\}| = 5$ and let

$$\begin{array}{ll} A := (\tau, a, b), & B := (\tau, c, d), \\ C := (\tau, a, c), & D := (\tau, b, d), \\ E := (\tau, a, d), & F := (\tau, b, c). \end{array}$$

(1) The pair (τ, λ) is called χ-*normalized* if

$$\chi([A][B]) \geq 0, \quad \chi([C][D]) \geq 0, \quad \chi([E][F]) \geq 0.$$

(2) A χ-normalized pair (τ, λ) is called χ-*non-degenerate* if $\chi([C][D]) > 0$.

(3) For a χ-non-degenerate pair (τ, λ) we call

$$\begin{array}{lll} [A][B] < [C][D] & \text{a } \textit{bi-quadratic inequality} & \text{if } \chi([E][F]) > 0, \\ [A][B] = [C][D] & \text{a } \textit{bi-quadratic equation} & \text{if } \chi([E][F]) = 0, \\ [E][F] < [C][D] & \text{a } \textit{bi-quadratic inequality} & \text{if } \chi([A][B]) > 0, \\ [E][F] = [C][D] & \text{a } \textit{bi-quadratic equation} & \text{if } \chi([A][B]) = 0. \end{array}$$

In fact (as a consequence of the oriented matroid axioms) for any $\tau \in E$ and $\lambda \in E^4$ there is always a suitable permutation $\pi \in S_4$ of the elements in λ such that $(\tau, \pi(\lambda))$ is χ-normalized. Furthermore, if $[A][B] = [C][D]$ is a bi-quadratic equation, $[C][D] = [A][B]$ is a bi-quadratic equation as well.

The set of all bi-quadratic inequalities of χ will be denoted by $\mathcal{B}_\chi$ and the set of all its bi-quadratic equations will be denoted by $\mathcal{A}_\chi$. Each element in $\mathcal{B}_\chi \cup \mathcal{A}_\chi$ is called a bi-quadratic expression. The bi-quadratic expressions can be considered as natural consequences of Graßmann-Plücker relations in the realizable case, as we will see now.

Lemma 4.2. *For a vector configuration $X \in (\mathbb{R}^d)^n$ and its corresponding oriented matroid χ_X we have*

(i) $[A]_X[B]_X < [C]_X[D]_X$ *for all* $[A][B] < [C][D] \in \mathcal{B}_{\chi_X}$.

(ii) $[A]_X[B]_X = [C]_X[D]_X$ *for all* $[A][B] = [C][D] \in \mathcal{A}_{\chi_X}$

Proof.
(i): Assume that $[A][B] < [C][D]$ is a bi-quadratic inequality and let (τ, λ) be the corresponding χ-non-degenerate pair. Let $A, \ldots, F$ be defined as in Definition 4.1. We have $\chi([E][F]) = 1$. The polynomial $[A][B] - [C][D] + [E][F]$ is a Graßmann-Plücker-polynomial. Hence its evaluation is identical to zero for every configuration $X \in (\mathbb{R}^d)^n$:

$$[A]_X[B]_X - [C]_X[D]_X + [E]_X[F]_X = 0.$$

Since $\chi([E][F]) = 1$, in any realization X of χ we have $[A]_X[B]_X - [C]_X[D]_X < 0$. This proves the first part of the lemma.
(ii): Let $[A][B] = [C][D]$ be a bi-quadratic equation and let $(\tau, \lambda), E, F$ be defined as above. Then we have $\chi([E][F]) = 0$. Therefore in any realization X of χ we have $[A]_X[B]_X - [C]_X[D]_X = 0$. □

The following definition of *bi-quadratic final polynomials* is more general than the one given in [2], where only the uniform case (no zero determinants) was considered.

Definition 4.3. For an oriented matroid χ a non-empty collection of bi-quadratic inequalities

$$[A_i][B_i] < [C_i][D_i] \in \mathcal{B}_\chi; 1 \le i \le k$$

together with a (possibly empty) collection of bi-quadratic equations

$$[A_i][B_i] = [C_i][D_i] \in \mathcal{A}_\chi; k+1 \le i \le l$$

is called a *bi-quadratic final polynomial* if the following equality holds within the ring $B_{3,n}$ (where brackets are identified according to the alternating determinant rule):

$$\prod_{i=1}^{l}[A_i][B_i] = \prod_{i=1}^{l}[C_i][D_i].$$

Lemma 4.4. [2, Lemma 4.1] *If χ admits a bi-quadratic final polynomial, then χ is not realizable over* $\mathbb{R}$.

Proof. Assume on the contrary that χ admits a bi-quadratic final polynomial as defined above, and χ is realizable, i.e $\chi = \chi_X$ for a suitable vector configuration X. By Lemma 4.2 we have

$$\begin{aligned} [A_i]_X[B_i]_X &< [C_i]_X[D_i]_X \quad \text{for all } 1 \le i \le k, \text{ and} \\ [A_i]_X[B_i]_X &= [C_i]_X[D_i]_X \quad \text{for all } k+1 \le i \le l. \end{aligned}$$

At least one proper inequality appears. By definition the products on the left side are all positive and the products on the right side are positive as well. If we multiply all right and all left sides we obtain:

$$\prod_{i=1}^{l}[A_i]_X[B_i]_X < \prod_{i=1}^{l}[C_i]_X[D_i]_X.$$

On the other hand the fact that we have a bi-quadratic final polynomial implies

$$\prod_{i=1}^{l}[A_i]_X[B_i]_X = \prod_{i=1}^{l}[C_i]_X[D_i]_X.$$

This contradicts the assumption that χ was realizable. □

5 Ω_{14}^{-} has no bi-quadratic final polynomial

The main result of this section is:

Theorem 5.1. *Let χ^0, χ^+, χ^- be three oriented matroids that differ in exactly one basis $\mu \in \Lambda(E,3)$ with $\chi^\sigma(\mu) = \sigma$. If χ^0 and χ^- are realizable and χ^+ is not, then χ^+ cannot have a bi-quadratic final polynomial.*

Proof. Assume that a bi-quadratic final polynomial for χ^+ exists. Let

$$\{[A_i][B_i] < [C_i][D_i] \mid 1 \leq i \leq k\} \subseteq \mathcal{B}_{\chi^+}$$

together with

$$\{[A_i][B_i] = [C_i][D_i] \mid k+1 \leq i \leq l\} \subseteq \mathcal{A}_{\chi^+}$$

be a bi-quadratic final polynomial for χ^+ consisting of $k > 0$ bi-quadratic inequalities and $l - k \geq 0$ bi-quadratic equations. Since $[\tau, b, c][\tau, e, f] = [\tau, c, b][\tau, f, e]$ holds, we may assume that every bracket in the bi-quadratic final polynomial has positive signature. In each bi-quadratic expression the bracket $[\mu]$ can be contained at most once (since each three-term Graßmann-Plücker-polynomial contains each bracket at most once). Since we have a bi-quadratic final polynomial the overall number r of occurrences of $[\mu]$ on the right sides of the expressions equals the number of overall occurrences of $[\mu]$ on the left sides. Thus we may assume that the bi-quadratic expressions are sorted in a way that each expression of the form $[A_i][B_i] \leq [C_i][D_i]$ with $\mu \in \{A_i, B_i\}$ is directly followed by an expression $[A_{i+1}][B_{i+1}] \leq [C_{i+1}][D_{i+1}]$ with $\mu \in \{C_{i+1}, D_{i+1}\}$ (indices taken modulo r).

With suitable $\tau_i \in E$ and $\lambda_i := (\lambda_{i1}, \ldots, \lambda_{i4}) \in E^4$ we have

$$\begin{aligned} A_i &:= (\tau_i, \lambda_{i1}, \lambda_{i2}), \quad & B_i &:= (\tau_i, \lambda_{i3}, \lambda_{i4}), \\ C_i &:= (\tau_i, \lambda_{i1}, \lambda_{i3}), \quad & D_i &:= (\tau_i, \lambda_{i_2}, \lambda_{i_4}). \end{aligned}$$

With this choice the Graßmann-Plücker polynomials

$$\{\tau_i|\lambda_i\} := [A_i][B_i] - [C_i][D_i] + [E_i][F_i]$$

are χ-normalized and χ-non-degenerate. By Definition 4.1 we know that $\chi([E_i][F_i])$ is $+1$ for $1 \le i \le k$ and 0 for $k+1 \le i \le l$. Furthermore $\chi([A_i][B_i]) = 1$ and $\chi([C_i][D_i]) = 1$ for all $1 \le i \le l$. We define monomials

$$m_i := \prod_{j=1}^{i-1}([A_i][B_i]) \cdot \prod_{j=i+1}^{l}([C_i][D_i])$$

and consider the polynomial

$$p := \sum_{i=1}^{l}\Big(m_i \cdot \{\tau_i|\lambda_i\}\Big).$$

We have

$$m_i \cdot [A_i][B_i] = m_{i+1} \cdot [C_{i+1}][D_{i+1}].$$

Furthermore, since all bi-quadratic expressions together form a bi-quadratic final polynomial, we also have

$$m_l \cdot [A_l][B_l] = \prod_{i=1}^{l}([A_i][B_i]) = \prod_{i=1}^{l}([C_i][D_i]) = m_1 \cdot [C_1][D_1].$$

Thus, canceling pairwise vanishing summands in p yields:

$$p = \sum_{i=1}^{l}\Big(m_i \cdot [E_i][F_i]\Big).$$

(In fact p is an ordinary final polynomial for χ^+ in the sense of Bokowski & Sturmfels [1, 5].) Since all Graßmann-Plücker-polynomials that are involved were χ-normalized we get:

$$\chi(m_i \cdot [E_i][F_i]) = 1 \text{ for } i = 1, \ldots, k$$

and

$$\chi(m_i \cdot [E_i][F_i]) = 0 \text{ for } i = k+1, \ldots, l.$$

By our assumption on the order of the bi-quadratic expressions in each of the monomials $m_i = [\mu]^r \cdot m_i'$ the bracket $[\mu]$ occurs with degree r (the total number of occurrences of $[\mu]$ on the right side of bi-quadratic expressions). Thus if we consider the polynomial

$$p' := \sum_{i=1}^{l}\Big(m_i' \cdot \{\tau_i|\lambda_i\}\Big) = \sum_{i=1}^{l}\Big(m_i' \cdot [E_i][F_i]\Big).$$

each summand $m_i' \cdot [E_i][F_i]$ is either linear in $[\mu]$ (in case that $\mu \in \{E_i, F_i\}$) or does not contain $[\mu]$ at all. Furthermore (since $\chi^+(\mu) = 1$) we have $\chi(m_i' \cdot [E_i][F_i]) = 1$ for $i = 1, \ldots, k$ and $\chi(m_i' \cdot [E_i][F_i]) = 0$ for $i = k+1, \ldots, l$. Thus we have

$$p' = [\mu] \cdot \sum_{i=1}^{s} p_i + \sum_{i=1}^{l-s} q_i$$

with $\chi(p_i)$ and $\chi(q_i)$ all either zero or positive and at least one of these monomials positive. Observe that the p_i and q_i are independent on $[\mu]$ thus the corresponding signs $\chi(p_i)$ and $\chi(q_i)$ are identical for χ^+, χ^0 and χ^-.

We now replace the brackets of p' by the values of the actual determinants of a realization of χ^0 (we know that such a realization does exist). The polynomial p' is a linear combination of Graßmann-Plücker-polynomials, hence this expression must evaluate to zero. Since $\chi^0([\mu]) = 0$ and the monomials q_i evaluate to a non-negative number we can conclude that $\chi(q_i) = 0$ for all $i = 1, \dots, l - s$.

Using this information we now consider the case where we replace the brackets of p' by the values of the actual determinants of a realization of χ^- (we know that such a realization does also exist). The summands q_i for all $i = 1, \dots, l - s$ evaluate to zero. Each of the summands $[\mu] \cdot p_i$ for $i = 1, \dots, s$ evaluates either to zero or to a number with sign since $\chi^-([\mu]) = -1$. At least one non-zero summand occurs. Thus we have a non-empty collection of negative numbers summing up to zero. □

Corollary 5.2. *The oriented matroid Ω_{14}^- is not realizable and does not admit a bi-quadratic final polynomial.*

Proof. The non-realizability of Ω_{14}^- was proved in Theorem 3.1. Since Ω_{14}^+ and Ω_{14}^0 are realizable Theorem 5.1 applies and the corollary follows. □

References

[1] A. Björner, M. Las Vergnas, B. Sturmfels, N. White & G.M. Ziegler, *Oriented Matroids, Encyclopedia of Mathematics*, Vol. 46, Cambridge University Press 1993.

[2] J. Bokowski & J. Richter, *On the finding of final polynomials, Eur. J. Comb.*, 11 (1990), 21–34.

[3] J. Bokowski, J. Richter & B. Sturmfels, *Nonrealizability proofs in computational geometry, Discrete Comput. Geom.*, 5 (1990), 333–350.

[4] J. Bokowski & B. Sturmfels, *Programmsystem zur Realisierung orientierter Matroide,* Preprint, Universtät Köln, (1985), 33 p.

[5] J. Bokowski & B. Sturmfels, *Computational Synthetic Geometry,* Lecture Notes in Mathematics, 1355, Springer-Verlag, Berlin Heidelberg 1989.

[6] N.E Mnëv, *On manifolds of combinatorial types of projective configurations and convex polyhedra, Soviet Math. Doklady*, 32 (1985), 335–337.

[7] N.E Mnëv, *The universality theorems on the classification problem of configuration varieties and convex polytopes varieties,* in: Viro, O.Ya. (ed.): *Topology and Geometry — Rohlin Seminar*, Lecture Notes in Mathematics 1346, Springer-Verlag, Berlin Heidelberg 1988, 527–544.

[8] J. RICHTER-GEBERT, *On the Realizability Problem of Combinatorial Geometries – Decision Methods,* Dissertation, TH-Darmstadt, 1992, 144 p.

[9] J. RICHTER-GEBERT, *Realization spaces of* 4*-polytopes are universal,* Habilitationsschrift TU-Berlin, 1995; Preprint 448/1995, TU Berlin 1995, 112 p.

[10] J. RICHTER-GEBERT & G.M. ZIEGLER, *Oriented Matroids,* Preprint, TU Berlin, September 1995, 28 p.; CRC Handbook on "Discrete and Computational Geometry" (J.E. Goodman, J. O'Rourke, eds.), to appear.

[11] P. SHOR, *Stretchability of pseudolines is NP–hard,* in: *Applied Geometry and Discrete Mathematics – The Victor Klee Festschrift* (P. Gritzmann, B. Sturmfels, eds.), DIMACS Series in Discrete Mathematics and Theoretical Computer Science, Amer. Math. Soc., Providence, RI, 4 (1991), 531–554.

[12] P.Y. SUVOROV, *Isotopic but not rigidly isotopic plane systems of straight lines,* in: Viro, O.Ya. (ed.): *Topology and Geometry — Rohlin Seminar*, Lecture Notes in Mathematics 1346, Springer-Verlag, Heidelberg 1988, 545–556.

Jürgen Richter-Gebert
Technische Universität Berlin
FB Mathematik, Sekr. MA 6-1
Straße des 17. Juni 136
D-10623 Berlin
Germany
richter@math.tu-berlin.de

Manis Valuations and Prüfer Extensions I

Manfred Knebusch and Digen Zhang

Received: February 19, 1996

Communicated by Ulf Rehmann

Abstract. We call a commutative ring extension $A \subset R$ Prüfer, if A is an R-Prüfer ring in the sense of Griffin (Can. J. Math. 26 (1974)). These extensions relate to Manis valuations in much the same way as Prüfer domains to Krull valuations. We develop a basic theory of Prüfer extensions and give some examples. In the introduction we try to explain why Prüfer extensions deserve interest from a geometric viewpoint.

1991 Mathematics Subject Classification: Primary 13A18; Secondary 13B02, 13B30.

Contents

Introduction

If F is a formally real field then it is well known that the intersection of the real valuation rings of F is a Prüfer domain H, and that H has the quotient field F. {A valuation ring is called real if its residue class field is formally real.} H is the so called *real holomorphy ring* of F, cf. [B, §2], [S], [KS, Chap.III §12]. If F is the function field $k(V)$ of an algebraic variety V over a real closed field k (e.g. $k = \mathbb{R}$), suitable overrings of H in R can tell us a lot about the algebraic and the semi-algebraic geometry of $V(k)$.

These rings, of course, are again Prüfer domains. A very interesting and – to our opinion – still mysterious role is played by some of these rings which are related to the orderings of higher level of F, cf. e.g. [B_2], [B_3]. Here we meet a remarkable phenomenon. For orderings of level 1 (i.e. orderings in the classical sense) the usual procedure is to observe first that the convex subrings of ordered fields are valuation rings, and then to go on to Prüfer domains as intersections of such valuation rings, cf. e.g. [B], [S], [KS]. But for higher levels, up to now, the best method is, to construct directly a Prüfer domain A in F from a "torsion reordering" of F, and then to obtain the valuation rings necessary for analyzing the reordering as localizations $A_{\mathfrak{p}}$ of A, cf. [B_2, p.1956 f], [B_3]. Thus there is a two way traffic between valuations and Prüfer domains.

Less is done up to now for F the function field $k(V)$ of an algebraic variety V over a p-adically closed field k (e.g. $k = \mathbb{Q}_p$). But work of Kochen and Roquette (cf. §6 and §7 in the book [PR] by Prestel and Roquette) gives ample evidence, that also here Prüfer domains play a prominent role. In particular, every formally p-adic field F contains a "*p-adic holomorphy ring*", called the *Kochen ring*, in complete analogy to the formally real case [PR, §6]. Actually the Kochen ring has been found and studied much earlier than the real holomorphy ring ([Ko], [R_1]).

If R is a commutative ring (with 1) and k is a subring of R then we can still define a real holomorphy ring $H(R/k)$ consisting of those elements a of R which on the real spectrum of R (cf. [BCR], [B_1], [KS]) can be bounded by elements of k. {If R is a formally real field F and k the prime ring of F this coincides with the real holomorphy ring H from above}. These rings $H(R/k)$ have proved to be very useful in real semi-algebraic geometry. In particular, N. Schwartz and M. Prechtel have used them in order to complete a real closed space and, more generally, to turn a morphism between real closed spaces into a proper one in a universal way ([Sch, Chap V, §7], [Pt]).

The algebra of these holomorphy rings turns out to be particularly good natured if we assume that $1 + \Sigma R^2 \subset R^*$, i.e. that all elements $1 + a_1^2 + \cdots + a_n^2$ $(n \in \mathbb{N}, a_i \in R)$ are units in R. This is a natural condition in real algebra. The rings used by Schwartz and Prechtel, consisting of abstract semi-algebraic functions, fulfill the condition automatically. More generally, if A is any commutative ring (always with 1) then the localization $S^{-1}A$ with respect to the multiplicative set $S = 1 + \Sigma A^2$ is a ring R fulfilling the condition, and R has the same real spectrum as A. Thus for many problems in real geometry we may replace A by R.

Recently V. Powers has proved that, if $1 + \Sigma R^2 \subset R^*$, the real holomorphy ring

$H(R/k)$ with respect to any subring k is an R-Prüfer ring, as defined by Griffin in 1973 [G_2].*) More generally V. Powers proved that, if $1 + \Sigma R^{2d} \subset R^*$ for some even number $2d$, every subring A of R containing the elements $\frac{1}{1+q}$ with $q \in \Sigma R^{2d}$ is R-Prüfer ([P, Th.1.7], cf. also [BP]).

An R-Prüfer ring is related to Manis valuations on R in much the same way as a Prüfer domain is related to valuations of its quotient field. Why shouldn't we try to repeat the success story of Prüfer domains and real valuations on the level of relative Prüfer rings and Manis valuations? Already Marshall in his important paper [Mar] has followed such a program. He has worked there with "Manis places" in a ring R with $1 + \Sigma R^2 \subset R^*$, and has related them to the points of the real spectrum Sper$R$.

We mention that Marshall's notion of Manis places is slightly misleading. By his definition these places do not correspond to Manis valuations but to a broader class of valuations which we call "special valuations", cf. §1 of the present paper. But then V. Powers (and independently one of us, D.Z.) observed that, in the case $1+\Sigma R^2 \subset R^*$, the places of Marshall in fact do correspond to the Manis valuations of R [P]. {In §1 of the present paper we prove that every special valuation of R is Manis under a much weaker condition on R, cf. Theorem 1.1.}

The program to study Manis valuations and relative Prüfer rings in rings of real functions has gained new impetus and urgency from the fact, that the theory of orderings of higher level has recently been pushed from fields to rings leading to *real spectra of higher level*. These spectra in turn have already proved to be useful for ordinary real semi-algebraic geometry. We mention an opus magnum by Ralph Berr [Be], where spectra of higher level are used in a fascinating way to classify the singularities of real semi-algebraic functions.

p-adic semi-algebraic geometry seems to be accessible as well. L. Bröcker and H.-J. Schinke have brought the theory of p-adic spectra to a rather satisfactory level by studying the "L-spectrum" L-spec A of a commutative ring A with respect to a given non-Archimedean local field L (e.g. $L = \mathbb{Q}_p$). There seems to be no major obstacle in sight which prevents us from defining and studying rings of semialgebraic functions on a constructible (or even pro-constructible) subset X of L-spec A. Here "semi-algebraic" means definability in a model theoretic sense plus a suitable continuity condition. Relative Prüfer subrings of such rings should be quite interesting.

The present paper is the first version of Chapter I of a book in preparation, devoted to a study of relative Prüfer rings and Manis valuations, with an eye to applications in real and p-adic geometry. In this chapter we present the basic theory and some examples.

Now, there exists already a rich theory of "*Prüfer rings with zero divisors*" also started by Griffin [G_1], cf. the books [LM], [Huc], and the literature cited there. But this theory seems not to be tailored to geometric needs. A Prüfer ring with zero divisors A is the same as an R-Prüfer ring with $R = \operatorname{Quot} A$, the total quotient ring of A. While this is a reasonable notion from the viewpoint of ring theory it may be artificial from a geometric viewpoint. A typical situation in real geometry is the following. R

*) The definition by Griffin needs a slight modification, cf. Def.1 in §5 below.

is the ring of (continuous) semialgebraic functions on a semialgebraic set M over a real closed field k or, more generally, the set of abstract semialgebraic functions on a pro-constructible subset X of a real spectrum (cf. [Sch], [Sch_1]). Although the ring R has very many zero divisors we have experience that in some sense R behaves nearly as well as a field, cf. e.g. our notion of "convenient ring extensions" in §6 of the present paper. Now, if A is a subring of R, then it is natural and interesting from a geometric viewpoint to study the R-Prüfer rings $B \supset A$, while the total quotient rings Quot$A$ and Quot$B$ seem to bear little geometric relevance.

Except in a paper by P.L. Rhodes from 1991 [Rh] very little seems to be done on relative Prüfer rings in general, and in the original paper of Griffin the proofs of important facts [G_2, Prop.6, Th.7] are omitted. Moreover the paper by Rhodes has a gap in the proof of his main theorem. {[Rh, Th.2.1], condition (5b) there is apparently not a characterization of Prüfer extensions. Any algebraic field extension is a counterexample.} Thus we have been careful about a foundation of this theory.

In §1 and §2 we gather what we need about Manis valuations. Then in §3 and §4 we develop an auxiliary theory of "weakly surjective" ring homomorphisms. These form a class of epimorphisms in the category of commutative rings close to the flat epimorphisms studied by D. Lazard and others in the sixties, cf. [L], [Sa_1], [A]. In §5 the up to then independent theories of Manis valuations and weakly surjective homomorphisms are brought together to study Prüfer extensions. {We call a ring extension $A \subset R$ Prüfer, if A is R-Prüfer in the sense of Griffin.} It is remarkable that, although Prüfer extensions are defined in terms of Manis valuations (cf. §5, Def.1 below), they can be characterized entirely in terms of weak surjectivity. Namely, a ring extension $A \subset R$ is Prüfer iff every subextension $A \subset B$ is weakly surjective (cf. Th.5.2 below). A third way to characterize Prüfer extensions is by multiplicative ideal theory, as we will explicate in Chapter II of our planned book.

Our first major result on Prüfer extensions is Theorem 5.2 giving various characterizations of these extensions which sometimes make it easy to recognize a given ring extension as Prüfer, cf. the examples in §6. We then establish various permanence properties of the class of Prüfer extensions. For example we prove for Prüfer extensions $A \subset B$ and $B \subset C$ that $A \subset C$ is again Prüfer (Th.5.6).

At the end of §5 we prove that any commutative ring A has a universal Prüfer extension $A \subset P(A)$ which we call the *Prüfer hull* of A. Every other Prüfer extension $A \hookrightarrow R$ can be embedded into $A \hookrightarrow P(A)$ in a unique way. The Prüfer rings with zero divisors are just the rings A with $P(A)$ containing the total quotient ring Quot$A$. Prüfer hulls mean new territory leading to many new open questions. We will pursue some of them in later chapters of our planned book.

In §6 we prove theorems which give us various examples of Manis valuations and Prüfer extensions. We illustrate how naturally they come up in algebraic geometry over a field k which is not algebraically closed (§6, Example 5, Th.6.5, Th.6.9), and in real algebraic and semialgebraic geometry (§6, Examples 3 and 10). Perhaps our best result here is Theorem 6.8 giving a far-reaching generalization of an old lemma by A. Dress (cf. [D, Satz 2′]). This lemma states for F a field, in which -1 is not a square, that the subring of F generated by the elements $1/(1+a^2)$, $a \in F$, is Prüfer in F.

Dress's innocent looking lemma seems to have inspired generations of real algebraists (cf. e.g. [La, p.86], [KS, p.163]) and also ring theorists, cf. [Gi_1].

We finally prove in §7 for various Prüfer extensions $A \subset R$ that, if $\mathfrak{a}$ is a finitely generated A-submodule of R with $R\mathfrak{a} = R$, then some power $\mathfrak{a}^d$ (with d specified) is principal. Our main result here (Theorem 7.8) is a generalization of a theorem by P. Roquette [R, Th.1] which states this for R a field (cf. also [Gi_1]). Roquette used his theorem to prove by general principles that the Kochen ring of a formally p-adic field is Bézout [loc.cit]. Similar applications should be possible in p-adic semialgebraic geometry. Roquette's paper has been an inspiration for our whole work since it indicates well the ubiquity of Prüfer domains in algebraic geometry over a non algebraically closed field.

Important topics missing in the present paper are multiplicative ideal theory, the characterization of a given Prüfer extension $A \subset R$ by a suitable lattice of ideals of A, approximation theory for Manis valuations and, finally, the construction of a "Manis valuation spectrum", i.e. a suitable space whose points are the Manis valuations of a given ring R. (One needs a condition on the ring R to establish this spectrum, otherwise one has to be content with the valuation spectrum Spev$R$, cf. [HK].) We will deal with these topics in later chapters of our planned book. A good deal of multiplicative ideal theory and the characterization business has already been done by Rhodes [Rh].

We have been forced to change some of the terminology used by ring theorists, say in the books of Larsen-McCarthy [LM] and of Huckaba [Huc]. While these authors mean by valuation on a ring a Manis valuation we use the word "valuation" in the much broader sense of Bourbaki [Bo, Chap.VI, §3]. It is true that Manis valuations are the really good ones for computations. But the central notion is the Bourbaki valuation, since only with these valuations one can build an honest spectral space, the valuation spectrum [HK]. Valuation spectra have already proved to be immensely useful both in algebraic geometry (cf. [HK]) and rigid analytic geometry (e.g. [Hu_1], [Hu_2]). The closely related real valuation spectra (cf. [Hu_3, §1]) seem to be the natural basic spaces for endeavors in real algebra concerning valuations and Prüfer extensions.

Some notations. In this paper all rings are commutative with 1. For A a ring we denote the group of units of A by A^*. We denote the total quotient ring of A by Quot$A$. For $\mathfrak{p}$ a prime ideal of A we denote the field $\mathrm{Quot}(A/\mathfrak{p})$ by $k(\mathfrak{p})$.

$\mathbb{N} = \{1, 2, 3, \dots\}$, $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$. If A and B are sets then $A \subset B$ means that A is a subset of B and $A \subsetneqq B$ means that A is a proper subset of B. If two subsets M and N of some set X are given then $M \setminus N$ denotes the complement of $M \cap N$ in M.

§1 Valuations on rings

Let R be a ring and Γ an (additive) totally ordered Abelian group. We extend Γ to an ordered monoid $\Gamma \cup \infty := \Gamma \cup \{\infty\}$ by the rules $\infty + x = x + \infty = \infty$ for all $x \in \Gamma \cup \infty$ and $x < \infty$ for all $x \in \Gamma$.

Definition 1 (Bourbaki [Bo, VI. 3.1]).
A *valuation on* R with values in Γ is a map $v: R \to \Gamma \cup \infty$ such that:

(1) $v(xy) = v(x) + v(y)$ for all $x, y \in R$.
(2) $v(x+y) \geq \min\{v(x), v(y)\}$ for all $x, y \in R$.
(3) $v(1) = 0$ and $v(0) = \infty$.

If $v(R) = \{0, \infty\}$ then v is said to be *trivial*, otherwise v is called *non-trivial*.

We recall some very basic facts[1] about valuations on rings and fix notations. Let $v: R \to \Gamma \cup \infty$ be a valuation on R.

The subgroup of Γ generated by $v(R) \setminus \{\infty\}$ is called the *value group* of v and is denoted by Γ_v. The set $v^{-1}(\infty)$ is a prime ideal of R. It is called the *support* of v and is denoted by $\operatorname{supp} v$. v induces a valuation $\hat{v}: k(\operatorname{supp} v) \to \Gamma \cup \infty$ on the quotient field $k(\operatorname{supp} v)$ of $R/\operatorname{supp} v$. We denote by $\mathfrak{o}_v$ the valuation ring of $k(\operatorname{supp} v)$ corresponding to $\hat{v}$, by $\mathfrak{m}_v$ its maximal ideal, and by $\kappa(v)$ its residue class field, $\kappa(v) := \mathfrak{o}_v/\mathfrak{m}_v$.

Notice that $\hat{v}(\mathfrak{o}_v) = (\Gamma_v)_+ \cup \{\infty\}$, where $(\Gamma_v)_+$ denotes the set of nonnegative elements in Γ_v. (We use such a notation for any ordered Abelian group.)

We further denote by A_v the set $\{x \in R \mid v(x) \geq 0\}$ and by $\mathfrak{p}_v$ the set $\{x \in R \mid v(x) > 0\}$. Clearly A_v is a subring of R and $\mathfrak{p}_v$ is a prime ideal of A_v. We call A_v the *valuation ring* of v and $\mathfrak{p}_v$ the *center* of v.

Definition 2. Two valuations v, w on R are said to be *equivalent*, in short, $v \sim w$, if the following equivalent conditions are satisfied:

(1) There is an isomorphism $f: \Gamma_v \cup \{\infty\} \to \Gamma_w \cup \{\infty\}$ of ordered monoids with $w(x) = f(v(x))$ for all $x \in R$.
(2) $v(a) \geq v(b) \Longleftrightarrow w(a) \geq w(b)$ for all $a, b \in R$.
(3) $\operatorname{supp} v = \operatorname{supp} w$ and $\mathfrak{o}_v = \mathfrak{o}_w$.

By abuse of language we will often regard equivalent valuations as "equal".

Definition 3. a) The *characteristic subgroup* $c_v(\Gamma)$ of Γ with respect to v is the smallest convex subgroup of Γ (convex with respect to the total ordering of Γ) which contains all elements $v(x)$ with $x \in R$, $v(x) \leq 0$. Clearly $c_v(\Gamma)$ is the set of all $\gamma \in \Gamma$ such that $v(x) \leq \gamma \leq -v(x)$ for some $x \in R$ with $v(x) \leq 0$.
b) v is called *special*,[2] if $c_v(\Gamma_v) = \Gamma_v$. (We replaced Γ by Γ_v.)

If H is any convex subgroup of Γ containing $c_v(\Gamma)$ then we obtain from v a new valuation $v|H: R \to \Gamma_\infty$ putting $(v|H)(x) = v(x)$ if $v(x) \in H$ and $v(x) = \infty$ else. Taking $H = c_v\Gamma$ we obtain from v a special valuation $w = v|c_v\Gamma$. Notice that $A_w = A_v$, $\mathfrak{p}_w = \mathfrak{p}_v$.

Definition 4 (cf. [M]). v is called a *Manis valuation* on R, if $v(R) = \Gamma_v \cup \infty$.[3]

---

[1] For this we refer to [Bo, VI.3.1] and [HK, §1]

[2] The word "special" alludes to the fact that such a valuation has no proper primary specialization in the valuation spectrum of R, cf. [HK, §1].

[3] Since we often identify equivalent valuations we have slightly altered the definition in [M]. Manis demands that $v(R)=\Gamma\cup\infty$.

Manis valuation will be in the focus of the present paper. Notice that every Manis valuation is special, but that the converse is widely false.

Example. Let R be the polynomial ring $k[x]$ in one variable x over some field k. Consider the valuation $v\colon R \to \mathbb{Z} \cup \infty$ with $v(f) = -\deg f$ for any $f \in R \setminus \{0\}$. This valuation is special but definitely not Manis. $\square$

One of our primary observations is that nevertheless there are many interesting rings, on which every special valuation is Manis. For example this holds if for every $x \in R$ the element $1 + x^2$ is a unit in R. More generally we have the following theorem.

Theorem 1.1. *Let k be a subring of R. Assume that for every $x \in R \setminus k$ there exists some monic polynomial $F(T) \in k[T]$ (one variable T) with $F(x) \in R^*$. Then every special valuation v on R with $A_v \supset k$ is Manis.*

Proof. We may assume that v is non trivial. Let $x \in R$ be given with $v(x) \neq 0, \infty$. We have to find some $y \in R$ with $v(y) = -v(x)$. Since v is special there exists some $a \in R$ with $v(ax) < 0$. Let $F(T) = T^d + c_1 T^{d-1} + \cdots + c_d$ be a polynomial with $c_1, \ldots, c_d \in k$ and $F(ax) \in R^*$. Since $v(ax) < 0$, but $v(c_i) \geq 0$ for $i = 1, \ldots, d$, we have $v(F(ax)) = dv(ax)$. The element $y := \frac{a^d x^{d-1}}{F(ax)}$ does the job.[4] q.e.d.

We return to valuations in general. Up to the end of this section we will keep the following

Notations. *$v\colon R \to \Gamma \cup \infty$ is a valuation on some ring R, $A := A_v$, $\mathfrak{p} := \mathfrak{p}_v$, $\mathfrak{q} :=$* $\operatorname{supp} v$, *$\bar{R} := R/\mathfrak{q}$, $\bar{A} := A/\mathfrak{q}$, $\bar{\mathfrak{p}} := \mathfrak{p}/\mathfrak{q}$. $\pi\colon R \to \bar{R}$ is the evident epimorphism from R to $\bar{R}$. We have a unique valuation $\bar{v}\colon \bar{R} \to \Gamma \cup \infty$ on $\bar{R}$ such that $\bar{v} \circ \pi = v$.*

We have $A_{\bar{v}} = \bar{A}$, $\mathfrak{p}_{\bar{v}} = \bar{\mathfrak{p}}$, $\operatorname{supp} \bar{v} = \{0\}$, $\Gamma_{\bar{v}} = \Gamma_v$, $\mathfrak{o}_v = \mathfrak{o}_{\bar{v}}$. It is evident that v is special iff $\bar{v}$ is special, and that v is Manis iff $\bar{v}$ is Manis. Looking at the valuation $\hat{v}$ on the quotient field $k(\mathfrak{q})$ of $\bar{R}$ (which extends $\bar{v}$) one now obtains by an easy exercise

Proposition 1.2.

a) *v is Manis iff $k(\mathfrak{q}) = \bar{R} \cdot \mathfrak{o}_v^*$.*
b) *v is special iff $k(\mathfrak{q}) = \bar{R} \cdot \mathfrak{o}_v$.*

Here $\bar{R} \cdot \mathfrak{o}_v^*$ (resp. $\bar{R} \cdot \mathfrak{o}_v$) denotes the set of products xy with $x \in \bar{R}$, $y \in \mathfrak{o}_v^*$ (resp. $\mathfrak{o}_v$). The set $\bar{R} \cdot \mathfrak{o}_v$ is also the subring of $k(\mathfrak{q})$ generated by $\bar{R}$ and $\mathfrak{o}_v$. $\square$

Definition 5. v is called *local* if the pair $(A, \mathfrak{p})$ is local, i.e. $\mathfrak{p}$ is the unique maximal ideal of A.

Proposition 1.3 (cf. [G_2, Prop. 5]).*The following are equivalent.*

i) *v is Manis and local.*
ii) *The pair $(R, \mathfrak{q})$ is local.*
iii) *v is local and $\mathfrak{q}$ is a maximal ideal of R.*

[4] We are indebted to Roland Huber for this simple argument.

PROOF. i) $\Rightarrow$ ii): Let $x \in R \setminus \mathfrak{q}$ be given. Since v is Manis there exists some $y \in R$ with $v(xy) = 0$. Since v is local this implies that xy is a unit of A, hence also a unit of R. Thus x is a unit of R.
ii) $\Rightarrow$ i): $\bar{v}$ is a valuation of the field $\bar{R}$. Thus $\bar{v}$ is Manis, which implies that v is Manis. Let $x \in A \setminus \mathfrak{p}$ be given. Then x is a unit in R. We have $v(x^{-1}) = -v(x) = 0$. Thus $x^{-1} \in A$, $x \in A^*$.
i), ii) $\Rightarrow$ iii): trivial.
iii) $\Rightarrow$ i): $\bar{v}$ is a valuation of the field $\bar{R}$. From this we conclude again that v is Manis.
□

If S is any multiplicative subset of R with $S \cap \mathfrak{q} = \emptyset$ then we denote by v_S the unique "extension" of v to a valuation on $S^{-1}R$, defined by

$$v_S\left(\frac{a}{s}\right) = v(a) - v(s) \qquad (a \in R, s \in S).$$

For $w = v_S$ we have $\Gamma_w = \Gamma_v$ and $c_w(\Gamma) \supset c_v(\Gamma)$. Thus if v is Manis then v_S is Manis and if v is special then v_S is special. v_S has the support $S^{-1}\mathfrak{q}$.

We now consider the special case $S = A \setminus \mathfrak{p}$. Then

$$v_S\left(\frac{a}{s}\right) = v(a) \qquad (a \in R, s \in S).$$

Thus for $w = v_S$ we now have $A_w = S^{-1}A = A_\mathfrak{p}$ and $\mathfrak{p}_w = S^{-1}\mathfrak{p} = \mathfrak{p}_\mathfrak{p}$, and we see that v_S is a local valuation. Moreover $A \setminus \mathfrak{p}$ is the smallest saturated multiplicative subset S of R such that v_S is local. We write $S^{-1}R = R_\mathfrak{p}$.

DEFINITION 6. The valuation v_S with $S = A \setminus \mathfrak{p}$ is called *the localization of* v, and is denoted by $\tilde{v}$.

We have $\tilde{v}(R_\mathfrak{p}) = v(R)$, $\Gamma_{\tilde{v}} = \Gamma_v$, $c_v\Gamma = c_{\tilde{v}}\Gamma$. Thus v is Manis iff $\tilde{v}$ is Manis and v is special iff $\tilde{v}$ is special. Applying Proposition 3 [5] to $\tilde{v}$ we obtain

PROPOSITION 1.4. *The following are equivalent.*

i) *v is Manis.*
ii) *$\mathfrak{q}$ is the unique ideal of R which is maximal among all ideals of R which do not meet $A \setminus \mathfrak{p}$.*
iii) *$\mathfrak{q}$ is maximal among all ideals of R which do not meet $A \setminus \mathfrak{p}$.* □

If S is a (non empty) multiplicative subset of R then we denote by $\mathrm{Sat}_R(S)$ the set of all elements of R which divide some element of S ("saturum of S in R"). Recall from basic commutative algebra that, if T is a second multiplicative subset of R, then $S^{-1}R = T^{-1}R$ iff $\mathrm{Sat}_R(S) = \mathrm{Sat}_R(T)$.

The following characterization of Manis valuations can be deduced from Proposition 4, but we will give an independent proof.

[5] Reference to Prop.1.3 in this section. In later sections we will refer to this proposition as "Prop.1.3", instead of "Prop.3".

Proposition 1.5. *The following are equivalent.*

i) v *is Manis.*
ii) $Sat_R(A \setminus \mathfrak{p}) = R \setminus \mathfrak{q}$.
iii) $R_\mathfrak{p} = R_\mathfrak{q}$.

Proof. The multiplicative set $R \setminus \mathfrak{q}$ is saturated. Thus the equivalence ii) $\Longleftrightarrow$ iii) is evident from what has been said above.
i) $\Longleftrightarrow$ ii): v is Manis $\Longleftrightarrow$ For every $x \in R \setminus \mathfrak{q}$ there exists some $y \in R$ with $v(x) + v(y) = 0$, i.e. with $xy \in A \setminus \mathfrak{p} \Longleftrightarrow R \setminus \mathfrak{q} = \mathrm{Sat}_R(A \setminus \mathfrak{p})$. □

Proposition 1.6. *If* v *is Manis then* $\mathfrak{o}_v = \bar{A}_{\bar{\mathfrak{p}}}$.

Proof. We may pass from v to $\bar{v}$. Thus we assume without loss of generality that $\mathfrak{q} = 0$. We have $\mathfrak{o}_v = \mathfrak{o}_{\tilde{v}}$ and v is Manis iff $\tilde{v}$ is Manis. Thus we may assume without loss of generality that v is also local. Now R is a field (cf. Prop. 3), and $\mathfrak{o}_v = A = A_\mathfrak{p}$. □

Definition 7. We say that v *has maximal support* if $\mathfrak{q}$ is a maximal ideal of R.

Proposition 1.7. v *has maximal support iff* $\bar{v}$ *is local and Manis. Then* v *is also a Manis valuation on* R.

Proof. If v has maximal support, then $\bar{v}$ is a valuation on the field $\bar{R}$. Thus $\bar{v}$ is certainly Manis and local. Since $\bar{v}$ is Manis, also v is Manis.

If $\bar{v}$ is local and Manis then, applying Proposition 3 to $\bar{v}$, we learn that the pair $(\bar{R}, \{0\})$ is local. This means that $\mathfrak{q}$ is a maximal ideal of R. □

Definition 8. An additive subgroup M of R is called v-*convex*, if for any elements $x \in M$, $y \in R$ with $v(x) \le v(y) (\le v(0) = \infty)$ it follows that $y \in M$.

If M is a v-convex additive subgroup of R, then certainly $ax \in M$ for any $a \in A$, $x \in M$, i.e. M is an A-submodule of R. We now have a closer look at the v-convex ideals of A.

Clearly $\mathfrak{q}$ is a v-convex ideal of A and is contained in any other v-convex ideal of A. Also $\mathfrak{p}$ is v-convex and $I \subset \mathfrak{p}$ for every v-convex ideal $I \neq A$.

Proposition 1.8. *If* v *has maximal support then every* A*-submodule of* R *containing* $\mathfrak{q}$ *is* v*-convex.*

Proof. Let I be an A-submodule of R containing $\mathfrak{q}$, and $\bar{I} := I/\mathfrak{q}$. It is easy to see that I is v-convex iff $\bar{I}$ is $\bar{v}$-convex. Since v has maximal support, $\bar{v}$ is a valuation on the field $\bar{R} := R/\mathfrak{q}$. From classical valuation theory we conclude that $\bar{I}$ is $\bar{v}$-convex. □

Corollary 1.9. *If* v *is a local Manis valuation then every* A*-submodule of* R *containing* $\mathfrak{q}$ *is* v*-convex.*

Proof. By Proposition 3 we know that v has maximal support. □

Proposition 1.10. [M, Prop. 3]. *Assume that the valuation v is Manis. Then a prime ideal $\mathfrak{r}$ of A is v-convex iff $\mathfrak{q} \subset \mathfrak{r} \subset \mathfrak{p}$.*

Proof. Replacing v by $\bar{v}$ we assume without loss of generality that $\mathfrak{q} = 0$. Since $v(A \setminus \mathfrak{p}) = \{0\}$ it is evident that the v-convex prime ideals $\mathfrak{r}$ of A correspond uniquely with the $\tilde{v}$-convex prime ideals $\mathfrak{r}'$ of $A_{\mathfrak{p}}$ via $\mathfrak{r}' = \mathfrak{r}_{\mathfrak{p}}$. Thus we may pass from v to $\tilde{v}$ and assume without loss of generality that v is local. All prime ideals (in fact, all ideals) of A are v-convex (Cor. 9). q.e.d.

Proposition 1.11. *Assume that v is a non trivial Manis valuation. The following are equivalent.*

i) *Every ideal I of A with $\mathfrak{q} \subset I \subset \mathfrak{p}$ is v-convex.*
ii) *Any two ideals I, J of A with $\mathfrak{q} \subset I \subset \mathfrak{p}$ and $\mathfrak{q} \subset J \subset \mathfrak{p}$ are comparable by inclusion.*
iii) *$\bar{A}$ is a (Krull)valuation domain.*
iv) *$\mathfrak{p}$ is the unique maximal ideal of A which contains $\mathfrak{q}$.*
v) *v has maximal support.*
vi) *Every ideal I of A containing $\mathfrak{q}$ is v-convex.*

Proof. We assume without loss of generality that $\mathfrak{q} = \{0\}$. Now R is an integral domain.

i) $\Rightarrow$ ii) is evident, since for any two v-convex ideals I and J of A we have $I \subset J$ or $J \subset I$. (This holds more generally for v-convex additive subgroups I, J of R.)

ii) $\Rightarrow$ iii): We verify: If $x \in A$, $y \in A$ then $Ax \subset Ay$ or $Ay \subset Ax$. This will imply that A is a valuation domain. We assume without loss of generality that $v(x) \leq v(y)$. If $x \in \mathfrak{p}$ then also $y \in \mathfrak{p}$. The ideals Ax and Ay are comparable by our assumption ii). There remains the case that $x \notin \mathfrak{p}$. We choose an element $c \neq 0$ in $\mathfrak{p}$. Then $xc \in \mathfrak{p}$ and $v(xc) \leq v(yc)$. As we have proved this implies $Ayc \subset Axc$ or $Axc \subset Ayc$. Since R is a domain we conclude that $Ay \subset Ax$ or $Ax \subset Ay$.

iii) $\Longrightarrow$ iv): trivial. iv) $\Longrightarrow$ v) is evident by Proposition 7, and v) $\Longrightarrow$ vi) is evident by Proposition 8. Clearly vi) $\Rightarrow$ i). □

Definition 9. A valuation $w: R \to \Gamma' \cup \infty$ is called *coarser* than v (or a *coarsening* of v) if there exists an order preserving homomorphism[6] $f: \Gamma_v \to \Gamma_w$ such that, for all $x \in R$, $w(x) = f(v(x))$ (put $f(\infty) = \infty$).

If H is a convex subgroup of Γ then the quotient Γ/H is a totally ordered Abelian group in such a way that the natural projection from Γ to Γ/H is an order preserving homomorphism. We have $(\Gamma/H)_+ = (\Gamma_+ + H)/H$. From v we obtain a coarsening $w: R \to (\Gamma/H) \cup \infty$ putting $w(x) := x + H$ for all $x \in R$. (Read $\infty + H = \infty$.) This valuation w is denoted by v/H.

Remarks 1.12. a) v/H has the center $\mathfrak{p}_H := \{x \in R \mid v(x) > H\}$, and this is a v-convex prime ideal of A. $\{v(x) > H$ means $v(x) > \gamma$ for every $\gamma \in H\}$. If $\Gamma_+ \subset v(R)$

[6] This means f is a homomorphism of Abelian groups with $f(\alpha) \geq f(\beta)$ if $\alpha \geq \beta$. The homomorphism f is necessarily surjective.

(e.g. v is Manis and $\Gamma = \Gamma_v$) then the v-convex prime ideals $\mathfrak{r}$ of A correspond uniquely with the convex subgroups H of Γ via $\mathfrak{r} = \mathfrak{p}_H$.

b) Assume (without loss of generality) that $\Gamma = \Gamma_v$. The coarsenings w of v correspond, up to equivalence, uniquely with the convex subgroups H of Γ via $w = v/H$. We have $A \subset A_w$, $\mathfrak{p} \supset \mathfrak{p}_w$, $\operatorname{supp} w = \mathfrak{q}$, $\hat{w} = \hat{v}/H$, $\bar{w} = \bar{v}/H$, $\tilde{w} = (\tilde{v}/H)^{\sim}$. If S is a multiplicative subset of R with $S \cap \mathfrak{q} = \emptyset$ then $v_S/H = (v/H)_S$. If v is special then v/H is special. If v is Manis then v/H is Manis.

All this is either trivial or can be verified in a straightforward way.

How do we obtain the ring A_w from $A_v = A$ if $w = v/H$? In order to give a satisfactory answer, at least in special cases, we need a definition which will be widely used also later on.

Definition 10. Let B be a subring of R, let S be a multiplicative subset of B and let $j_S \colon R \to S^{-1}R$ denote the localization map $x \mapsto \frac{x}{1}$ of R with respect to S. For any B-submodule M of R we define

$$M_{[S]} := j_S^{-1}(S^{-1}M).$$

Clearly $M_{[S]}$ is the set of all $x \in R$ such that $sx \in M$ for some $s \in S$. We call $M_{[S]}$ the *saturation of* M *(in* R*) by* S.[7] In the case $S = B \setminus \mathfrak{r}$ with $\mathfrak{r}$ a prime ideal of B we usually write $j_{\mathfrak{r}}$ and $M_{[\mathfrak{r}]}$ instead of j_S, $M_{[S]}$.

Notice that $B_{[S]}$ is a subring of R and $M_{[S]}$ is a $B_{[S]}$-submodule of R. If M is an ideal of B then $M_{[S]}$ is an ideal of $B_{[S]}$. If M is a prime ideal of B with $M \cap S = \emptyset$ then $M_{[S]}$ is a prime ideal of $B_{[S]}$.

Proposition 1.13. *Let S be a multiplicative subset of $A \setminus \mathfrak{q}$, and let H denote the convex subgroup of Γ generated by $v(S)$, i.e. the smallest convex subgroup of Γ containing $v(S)$. Let $w := v/H$ and $\mathfrak{r} := \mathfrak{p}_H$. Then*

$$\begin{aligned} A_w &= A_{[S]} = A_{[\mathfrak{r}]}, \\ \mathfrak{p}_w &= \mathfrak{r} = \{x \in R \mid v(x) > v(S)\}. \end{aligned}$$

Proof. We already stated above that $\mathfrak{p}_w = \mathfrak{p}_H = \mathfrak{r}$. This ideal coincides with the set of all $x \in R$ with $v(x) > v(S)$. It is evident that $A_{[S]} \subset A_w$. Let now $x \in A_w$ be given. There exists some element $\gamma \in H_+$ with $v(x) \geq -\gamma$, and some element $s \in S$ with $\gamma \leq v(s)$. We obtain $v(xs) \geq 0$, i.e. $xs \in A$. This proves that $A_w = A_{[S]}$. We have $S \subset A \setminus \mathfrak{r}$, thus $A_{[S]} \subset A_{[\mathfrak{r}]}$. Let $x \in A_{[\mathfrak{r}]}$ be given. We choose $y \in A \setminus \mathfrak{r}$ with $xy \in A$. There exists some $\gamma \in H_+$ with $v(y) \leq \gamma$ and some $s \in S$ with $\gamma \leq v(s)$. We have

$$0 \leq v(x) + v(y) \leq v(x) + v(s) = v(sx).$$

Thus $sx \in A$, $x \in A_{[S]}$. This proves $A_{[S]} = A_{[\mathfrak{r}]}$. q.e.d.

[7] $M_{[S]}$ is called the "S-component of M" in [LM].

REMARK. The converse of Proposition 13 for the case of non-trivial Manis valuations is also true (Th.2.6.ii).

COROLLARY 1.14. *Assume that $\Gamma_+ \subset v(R)$ (e.g. v Manis and $\Gamma_v = \Gamma$). Let H be a convex subgroup of Γ, $w := v/H$ and $\mathfrak{r} := \mathfrak{p}_H$. We have $A_w = A_{[\mathfrak{r}]}$ and $\mathfrak{p}_w = \mathfrak{r}$.*

PROOF. Apply Prop. 13 to the set $S := \{x \in R \mid v(x) \in H_+\}$. □

PROPOSITION 1.15. *Let I be an A-submodule of R with $\mathfrak{q} \subset I$. Assume that v is Manis. Then I is v-convex iff $I = I_{[\mathfrak{p}]}$.*

PROOF. Assume first that I is v-convex. We have $I \subset I_{[\mathfrak{p}]}$. Let $x \in I_{[\mathfrak{p}]}$ be given. We choose $d \in A \setminus \mathfrak{p}$ with $dx \in I$. We have $v(x) = v(dx)$. Since I is v-convex this implies $x \in I$. Thus $I = I_{[\mathfrak{p}]}$.

Assume now that $I = I_{[\mathfrak{p}]}$. This means $I = j_{\mathfrak{p}}^{-1}(I_{\mathfrak{p}})$ with $j_{\mathfrak{p}}$ the localization map from R to $R_{\mathfrak{p}}$. As always let $\tilde{v}: R_{\mathfrak{p}} \to \Gamma \cup \infty$ denote the localization of v. We have $A_{\tilde{v}} = A_{\mathfrak{p}}$, $\operatorname{supp} \tilde{v} = \mathfrak{q}_{\mathfrak{p}}$. Since $\tilde{v}$ is local, every $A_{\mathfrak{p}}$-submodule of $R_{\mathfrak{p}}$ containing $\mathfrak{q}_{\mathfrak{p}}$ is $\tilde{v}$-convex (Cor. 1.9). In particular $I_{\mathfrak{p}}$ is $\tilde{v}$-convex. Since $I = j_{\mathfrak{p}}^{-1}(I_{\mathfrak{p}})$ and $v = \tilde{v} \circ j_{\mathfrak{p}}$ we conclude that I is v-convex. □

We briefly discuss a process of restriction which gives us special valuations on subrings of R.

Let B be a subring of R. The restriction $u = v|B: B \to \Gamma \cup \infty$ of the map $v: R \to \Gamma \cup \infty$ is a valuation on B. Let $\Delta := c_u(\Gamma)$ and $w := u|\Delta$. Then $w: B \to \Delta \cup \infty$ is a special valuation on B.

DEFINITION 11. We call w the *special restriction* of v to B, and denote this valuation by $v|_B$.

For $w = v|_B$ we have $A_w = A \cap B$, $\mathfrak{p}_w = \mathfrak{p} \cap B$, $\operatorname{supp} w \supset \mathfrak{q} \cap B$. Notice also that $v|_B = (v|c_v\Gamma)|_B$. Thus in essence our restriction process deals with special valuations.

In the case that v is Manis the question arises, under which conditions on B the special restriction $v|_B$ is again Manis. We need an easy lemma.

LEMMA 1.16. *If $v: R \to \Gamma \cup \infty$ is special and $(\Gamma_v)_+ \subset v(R)$, then v is Manis.*

PROOF. This is a consequence of Proposition 2. By that proposition $k(\mathfrak{q}) = \bar{R}\mathfrak{o}_v$. From $(\Gamma_v)_+ \subset v(R) = \bar{v}(\bar{R})$ we conclude that $\mathfrak{o}_v \subset \bar{R}\mathfrak{o}_v^*$, hence $k(\mathfrak{q}) = \bar{R}\mathfrak{o}_v^*$, and this means that v is Manis. □

PROPOSITION 1.17. *Assume that v is Manis and that B is a subring of R containing $\mathfrak{p} = \mathfrak{p}_v$. Then the special restriction $v|_B: B \to \Delta \cup \infty$ of v is again Manis. If v is surjective (i.e. $\Gamma = \Gamma_v$) then $v|_B$ is surjective.*

PROOF. We assume without loss of generality that v is surjective. Let $u := v|B$ and $w := v|_B$. Let $\gamma \in \Delta$ be given with $\gamma > 0$. There exists some $a \in \mathfrak{p}_v$ with $v(a) = \gamma$. Since $\mathfrak{p}_v \subset B$ we have $a \in B$, hence $v(a) = u(a) = w(a)$. {Recall that for any $x \in B$

with $u(x) \in \Delta$ we have $w(x) = u(x)$.} This proves that $\Delta_+ \subset w(B)$. By the lemma w is Manis. $\square$

Scholium 1.18. Let $v: R \to \Gamma \cup \infty$ be a Manis valuation and H a convex subgroup of Γ. Let $w: = v/H$ and $B: = A_w$. We have

$$A_w = \{x \in R \mid v(x) \geq h \quad \text{for some} \quad h \in H\} =: A_H$$
$$\mathfrak{p}_w = \{x \in R \mid v(x) > h \quad \text{for all} \quad h \in H\} =: \mathfrak{p}_H.$$

Let $v_H: B \to \Delta \cup \infty$ denote the special restriction $v|_B$ of v. Here $\Delta = c_{v|B}(\Gamma) \subset H$. v_H has support $\mathfrak{p}_H$, hence gives us a Manis valuation $\overline{v_H}: A_H/\mathfrak{p}_H \to \Delta \cup \infty$ of support zero. If v is surjective then $\Delta = H$.

The proof of all this is a straightforward exercise. Later we will prove a converse to these statements (Prop. 2.8).

Using Lemma 16 from above we can prove a converse to Proposition 6.

Proposition 1.19. *Assume that the valuation v on R is special and that $\mathfrak{o}_v = \bar{A}_{\bar{\mathfrak{p}}}$ (cf. notations above). Then v is Manis.*

Proof. Replacing A by $\bar{A} = A/\mathfrak{q}$ and v by $\bar{v}$ we assume without loss of generality that $\mathfrak{q} = 0$. Now R is an integral domain, and $A \subset R \subset K$ with K the quotient field of R. We also assume without loss of generality that $\Gamma = \Gamma_v$. The valuation $v: R \to \Gamma \cup \infty$ extends to the valuation $\hat{v}: K \twoheadrightarrow \Gamma \cup \infty$, and $\hat{v}$ has the valuation ring $\mathfrak{o}_v$. We have $v(A \setminus \mathfrak{p}) = \{0\}$, hence $v(A) = \hat{v}(A_\mathfrak{p}) = \hat{v}(\mathfrak{o}_v) = \Gamma_+$. By Lemma 16 we conclude that v is Manis. $\square$

§2 Valuation subrings and Manis pairs

As before let R be a ring (commutative, with 1).

Definition 1. a) A *valuation subring of* R is a subring A of R such that there exists some valuation $v: R \to \Gamma \cup \infty$ with $A = A_v$. A *valuation pair in* R (also called "R-valuation pair") is a pair $(A, \mathfrak{p})$ consisting of a subring A of R and a prime ideal $\mathfrak{p}$ of A such that $A = A_v$, $\mathfrak{p} = \mathfrak{p}_v$ for some valuation v of R.
b) We speak of a *Manis subring* A of R and a *Manis pair* $(A, \mathfrak{p})$ in R respectively if here v can be chosen as a Manis valuation of R.

Two bunches of questions come to mind immediately. 1) How can a valuation subring or a Manis subring of R be characterized ring theoretically? Ditto for pairs.
2) How far is a valuation v determined by the associated ring A_v or pair $(A_v, \mathfrak{p}_v)$?

As stated in §1 the pair $(A_v, \mathfrak{p}_v)$ does not change if we pass from v to the associated special valuation $v|c_v\Gamma$. Thus, starting from now, we will concentrate on special valuations.

If $A = R$ then a special valuation v with $A_v = A$ must be trivial, and any prime ideal $\mathfrak{p}$ of R occurs as the center (= support) of such a valuation v. The valuation v is completely determined by $(R, \mathfrak{p})$ and is Manis. These pairs $(R, \mathfrak{p})$ are called the *trivial Manis pairs* in R.

If $A \neq R$ and A is a valuation subring of R then clearly $R \setminus A$ is a multiplicatively closed subset of R. P. Samuel started an investigation of such subrings of R. We quote one of his very remarkable results.

Definition 2. Let A be a subring of R with $A \neq R$ and $S := R \setminus A$ multiplicatively closed. We define the following subsets $\mathfrak{p}_A$ and $\mathfrak{q}_A$ of A. $\mathfrak{p}_A$ is the set of all $x \in A$ such that there exists some $s \in S$ with $sx \in A$, and $\mathfrak{q}_A$ is the set of all $x \in A$ with $sx \in A$ for all $s \in R \setminus A$.

Clearly $\mathfrak{q}_A \subset \mathfrak{p}_A$. Also $\mathfrak{q}_A = \{x \in R \mid rx \in A \text{ for all } r \in R\}$. Thus $\mathfrak{q}_A$ is the biggest ideal of R contained in A, called the *conductor* of A in R.

Theorem 2.1. [Sa, Th.1 and Th.2]. *Let A be a proper subring of R with $R \setminus A$ multiplicatively closed.*

i) *$\mathfrak{p}_A$ is a prime ideal of A and $\mathfrak{q}_A$ is a prime ideal both of A and R.*
ii) *A is integrally closed in R.*
iii) *If R is a field then A is a valuation domain, and R is the quotient field of A.* □

If v is a special nontrivial valuation then the support of v is determined by the ring A_v alone. More precisely we have the following proposition, whose proof is an easy exercise.

Proposition 2.2. *Let v be a non trivial valuation on R and $A := A_v$. Then $\mathfrak{q}_A \supset \operatorname{supp} v$. The valuation v is special iff $\mathfrak{q}_A = \operatorname{supp} v$.* □

We cannot expect that a special valuation v is determined up to equivalence by the pair $(A, \mathfrak{p}) := (A_v, \mathfrak{p}_v)$, as is already clear from the example in §1. But this holds if v is Manis. Indeed, if v is also non trivial, then we see from Prop. 2 and Prop.1.6 that $\mathfrak{o}_v = \bar{A}_{\bar{\mathfrak{p}}}$ with $\bar{A} = A/\mathfrak{q}_A$, $\bar{\mathfrak{p}} = \mathfrak{p}/\mathfrak{q}_A$. Even more is true. The following proposition implies that v is determined up to equivalence by A alone. The proof is again an easy exercise.

Proposition 2.3. *Let v be a non trivial valuation on R and $A := A_v$. Then $\mathfrak{p}_A \subset \mathfrak{p}_v$. If v is Manis then $\mathfrak{p}_A = \mathfrak{p}_v$.* □

We have the following important characterization of Manis pairs.

Theorem 2.4 ([M, Prop. 1], or [Huc, Th. 5.1]). *Let A be a subring of R and $\mathfrak{p}$ a prime ideal of A. The following are equivalent.*

i) *$(A, \mathfrak{p})$ is a Manis pair in R.*
ii) *If B is a subring of R and $\mathfrak{q}$ a prime ideal of B with $A \subset B$ and $\mathfrak{q} \cap A = \mathfrak{p}$ then $A = B$.* [1)]

[1)] In [M] and [Huc] it is not assumed that $\mathfrak{q}$ is a prime ideal. It can be proved easily that their condition can be changed to our condition (ii).

iii) *For every $x \in R \setminus A$ there exists some $y \in A$ with $xy \in A \setminus \mathfrak{p}$.* $\square$

There also exists a satisfying characterization of the valuation subrings of R in ring theoretic terms, due to Samuel and Griffin [e.g.Huc, Th.5.5], but we do not need this here.

We give a characterization of local Manis pairs in a classical style.

Theorem 2.5. *Let $A \subset R$ be a ring extension, $A \neq R$.*
i) The following are equivalent
(1) *Every $x \in R \setminus A$ is a unit in R and $x^{-1} \in A$.*
(2) *A has a unique maximal ideal $\mathfrak{p}$ (hence is local) and $(A, \mathfrak{p})$ is Manis in R.*
ii) If (1), (2) hold, then R is a local ring with maximal ideal $\mathfrak{q} := \mathfrak{q}_A$ and $A_{\mathfrak{q}} = R_{\mathfrak{p}} = R$. Moreover, $\mathfrak{p} = \mathfrak{q} \cup \{x^{-1} | x \in R \setminus A\}$.

Proof. Assume that (1) holds. Then $R \setminus A$ is closed under multiplication. Indeed, let $x, y \in R \setminus A$ be given. Then $(xy)y^{-1} \in R \setminus A$, but $y^{-1} \in A$, hence $xy \in R \setminus A$. We introduce the prime ideals $\mathfrak{p} := \mathfrak{p}_A$ and $\mathfrak{q} := \mathfrak{q}_A$ (cf. Def. 2). If $\mathfrak{M}$ is any maximal ideal of R then $\mathfrak{M} \cap (R \setminus A) = \emptyset$, since $R \setminus A \subset R^*$, and $\mathfrak{M} \subset A$. Thus $\mathfrak{M}$ is contained in the conductor $\mathfrak{q}$ of A in R, and we conclude that $\mathfrak{M} = \mathfrak{q}$. Thus $\mathfrak{q}$ is the only maximal ideal of R. Let K denote the field $R/\mathfrak{q}$ and $\overline{A}$ the subring $A/\mathfrak{q}$ of K. For every $z \in K \setminus \overline{A}$ the inverse z^{-1} is contained in $\overline{A}$. Thus $\overline{A}$ is a valuation domain with quotient field K. We conclude that A is Manis in R, and then, that $(A, \mathfrak{p})$ is a Manis pair in R (cf. Prop. 3). Since $(R, \mathfrak{q})$ is local we learn from Proposition 1.3 that $(A, \mathfrak{p})$ is local.

Now assume that (2) holds. We know from Proposition 1.3 that R is local with maximal ideal $\mathfrak{q} := \mathfrak{q}_A$. Thus $R \setminus A \subset R \setminus \mathfrak{q} = R^*$. Since $(A, \mathfrak{p})$ is Manis in R we have $x^{-1} \in \mathfrak{p} \subset A$ for every $x \in R \setminus A$, and it is also clear that $\mathfrak{p} = \mathfrak{q} \cup \{x^{-1} | x \in R \setminus A\}$.

We have $A \setminus \mathfrak{q} \subset R^*$, hence $A_{\mathfrak{q}} \subset R$. If $x \in R \setminus A$ then $x = \frac{1}{y}$ with $y \in A \setminus \mathfrak{q}$. Thus $x \in A_{\mathfrak{q}}$. This proves that $A_{\mathfrak{q}} = R$. Since $A \setminus \mathfrak{p} \subset R^*$ also $R_{\mathfrak{p}} = R$.

Assume now that (2) holds. We know from Proposition 1.3 that R is local with maximal ideal $\mathfrak{q} := \mathfrak{q}_A$. Thus $R \setminus A \subset R \setminus \mathfrak{q} = R^*$. Since $(A, \mathfrak{p})$ is Manis in R we have $x^{-1} \in \mathfrak{p}$ for every $x \in R \setminus A$, a fortiori $x^{-1} \in A$. $\square$

Let $v : R \longrightarrow \Gamma \cup \infty$ and w be valuations on R. We have called w coarser than v if w is equivalent to v/H for some convex subgroup H of v (§1, Def. 9 and Remark 1.12). How can the coarsening relation be expressed in terms of the pairs $(A_v, \mathfrak{p}_v)$, $(A_w, \mathfrak{p}_w)$ if both v and w are Manis?

Theorem 2.6 (cf. [M, Prop.4] for a weaker statement). *Assume that $v : R \longrightarrow \Gamma \cup \infty$ and w are two non-trivial Manis valuations of R.*

i) *The following are equivalent:*
(1) *w is coarser than v.*
(2) *$supp\,(v) = supp\,(w)$ and $\mathfrak{o}_v \subset \mathfrak{o}_w$.*
(3) *$A_v \subset A_w$ and $\mathfrak{p}_w \subset \mathfrak{p}_v$.*
(4) *$\mathfrak{p}_w$ is an ideal of A_v contained in $\mathfrak{p}_v$.*

ii) *Let* $A := A_v$, $\mathfrak{p} := \mathfrak{p}_v$, *and let* $\mathfrak{r}$ *be a prime ideal of* A *with* $\operatorname{supp} v \subset \mathfrak{r} \subset \mathfrak{p}$. *Let* H *denote the convex subgroup of* Γ *generated by* $v(A \setminus \mathfrak{r})$ *and* $w := v/H$. *Then* $\mathfrak{r} = \mathfrak{p}_H = \mathfrak{p}_w$ *and* $A_{[\mathfrak{r}]} = A_w = A_H$.[2)]

Proof: (1) $\iff$ (2): We may assume in advance that $\operatorname{supp} v = \operatorname{supp} w$. It is now evident that w is coarser than v iff $\hat{w}$ is coarser than $\hat{v}$. By classical valuation theory this holds iff the valuation ring $\mathfrak{o}_v$ of $\hat{v}$ is contained in $\mathfrak{o}_w$.

(2) $\Longrightarrow$ (3): Replacing R by $R/\operatorname{supp} v$ we assume without loss of generality that $\operatorname{supp} v = \operatorname{supp} w = \{0\}$. In the quotient field K of R we have $\mathfrak{o}_v \cap R = A_v$, $\mathfrak{o}_w \cap R = A_w$, $\mathfrak{m}_v \cap R = \mathfrak{p}_v$ and $\mathfrak{m}_w \cap R = \mathfrak{p}_w$. By assumption $\mathfrak{o}_v \subset \mathfrak{o}_w$. This implies $\mathfrak{m}_v \supset \mathfrak{m}_w$. We conclude that $A_v \subset A_w$ and $\mathfrak{p}_v \supset \mathfrak{p}_w$.

(3) $\Longrightarrow$ (2): We verify first that $\operatorname{supp}(v) = \operatorname{supp}(w)$. We know that $\operatorname{supp}(v) = \{x \in R \mid xR \subset A_v\}$ and $\operatorname{supp}(w) = \{x \in R \mid xR \subset A_w\}$ (cf. Proposition 2). Using the assumption $A_v \subset A_w$ we conclude $\operatorname{supp} v \subset \operatorname{supp} w$. Since v, w are Manis valuations, it is also evident that $\operatorname{supp}(v) = \{x \in R \mid xR \subset \mathfrak{p}_v\}$ and $\operatorname{supp}(w) := \{x \in R \mid xR \subset \mathfrak{p}_w\}$. Using the assumption $\mathfrak{p}_v \supset \mathfrak{p}_w$ we conclude that $\operatorname{supp} v \supset \operatorname{supp} w$. Thus indeed $\operatorname{supp}(v) = \operatorname{supp}(w)$.

In order to prove that $\mathfrak{o}_v \subset \mathfrak{o}_w$ we may replace R by $R/\operatorname{supp} v$. Thus we may assume that $\operatorname{supp} v = \operatorname{supp} w = \{0\}$. Now we know from Proposition 1.6 that $\mathfrak{o}_v = (A_v)_{\mathfrak{p}_v}$ and $\mathfrak{o}_w = (A_w)_{\mathfrak{p}_w}$. The inclusions $A_v \subset A_w$ and $\mathfrak{p}_v \supset \mathfrak{p}_w$ imply that $\mathfrak{o}_v \subset \mathfrak{o}_w$.

(3) $\Longrightarrow$ (4): trivial.

(4) $\Longrightarrow$ (3): Since w is Manis we have $A_w = \{x \in R \mid x\mathfrak{p}_w \subset \mathfrak{p}_w\}$. Now $\mathfrak{p}_w$ is an ideal of A_v. Thus $A_v \subset A_w$.

ii): We know from Prop.1.10 that the ideal $\mathfrak{r}$ is v-convex, and from Remark 1.12.a that $\mathfrak{r} = \mathfrak{p}_H$. Let $w := v/H$ and $B := A_w$. We have $B = A_H$ (cf. 1.18) and $\mathfrak{p}_w = \mathfrak{p}_H = \mathfrak{r}$.

It remains to prove that $B = A_{[\mathfrak{r}]}$. Let $x \in A_{[\mathfrak{r}]}$ be given. We choose some $d \in A \setminus \mathfrak{r}$ with $dx \in A$. Since $A \subset A_w$, $\mathfrak{r} = \mathfrak{p}_w$, we have $w(dx) \geq 0$, $w(d) = 0$, hence $w(x) \geq 0$, i.e. $x \in B$. This proves that $A_{[\mathfrak{r}]} \subset B$. Let now $x \in B$ be given. Suppose that $x \notin A_{[\mathfrak{r}]}$. Since $x \notin A$ there exists some $x' \in \mathfrak{p}$ with $xx' \in A \setminus \mathfrak{p} \subset A \setminus \mathfrak{r} \subset A$. Since $x \notin A_{[\mathfrak{r}]}$ we have $x' \in \mathfrak{r}$. Thus $x\mathfrak{r} \not\subset \mathfrak{r}$. This is a contradiction, since $\mathfrak{r}$ is an ideal of B and $x \in B$. Thus $x \in A_{[\mathfrak{r}]}$. We have proved $B = A_{[\mathfrak{r}]}$. q.e.d.

Corollary 2.7. *Let* $v : R \to \Gamma \cup \infty$ *be a Manis valuation and* $A := A_v$, $\mathfrak{p} = \mathfrak{p}_v$. *The coarsenings* w *of* v *correspond uniquely, up to equivalence, with the prime ideals* $\mathfrak{r}$ *of* A *between* $\operatorname{supp} v$ *and* $\mathfrak{p}$ *via* $\mathfrak{r} = \mathfrak{p}_w$. *Also* $A_{[\mathfrak{r}]} = A_w$.

Proof. If v is trivial then $\operatorname{supp} v = \mathfrak{p}$, and all assertions are evident. Assume now that v is not trivial. For the trivial coarsening t of v we have $\mathfrak{p}_t = \operatorname{supp} t = \operatorname{supp} v$ and $A_{[\mathfrak{p}_t]} = R$. If w is a non trivial coarsening of v then $\mathfrak{p}_w$ is an ideal of A with $\operatorname{supp} v \subsetneqq \mathfrak{p}_w \subset \mathfrak{p}$ (cf. Th.6.i). This ideal is prime in A since it is prime in the ring $A_w \supset A$. Conversely, if $\mathfrak{r}$ is a prime ideal of A with $\operatorname{supp} v \subsetneqq \mathfrak{r} \subset \mathfrak{p}$ then, by Theorem 6.ii, there exists a coarsening w of v with $\mathfrak{p}_w = \mathfrak{r}$, $A_w = A_{[\mathfrak{r}]}$, and w is not trivial.

2) Recall the notations from 1.12 and 1.18.

Finally, if w and w' are two nontrivial coarsenings of v with $\mathfrak{p}_w = \mathfrak{p}_{w'} = \mathfrak{r}$, then $A_w = \{x \in R \mid x\mathfrak{r} \subset \mathfrak{r}\} = A_{w'}$, and we learn from (3) in Theorem 6.i (or by a direct argument), that $w \sim w'$. $\square$

We establish a converse to the construction 1.18.

Proposition 2.8. *Let w be a non-trivial Manis valuation on R and u a Manis valuation on $A_w/\mathfrak{p}_w$. Let A and $\mathfrak{p}$ denote the pre-images of A_u and $\mathfrak{p}_u$ in A_w under the natural homomorphism $\varphi: A_w \to A_w/\mathfrak{p}_w$.*

i) *$(A, \mathfrak{p})$ is a Manis pair in R iff $\operatorname{supp} u = \{0\}$.*

ii) *If this holds, let $v: R \twoheadrightarrow \Gamma \cup \infty$ be a surjective valuation with $A_v = A$, $\mathfrak{p}_v = \mathfrak{p}$. Then Γ has a convex subgroup H, uniquely determined by w and u, such that w is equivalent to v/H and u is equivalent to $\overline{v_H}$ (cf. 1.18).*

Proof. We have $\mathfrak{p}_w \subset \mathfrak{p} \subset A \subset A_w \subset R$.

a) We assume that $\operatorname{supp} u = \{0\}$ and prove that the pair $(A, \mathfrak{p})$ is Manis in R. Let $x \in R \setminus A$ be given. By Theorem 4 we are done if we find some $y \in \mathfrak{p}$ with $xy \in A \setminus \mathfrak{p}$.

Case 1: $x \in A_w$. Since $\varphi(x) \notin A_u$ there exists some $y \in \mathfrak{p}$ with $\varphi(x)\varphi(y) \in A_u \setminus \mathfrak{p}_u$, hence $xy \in A \setminus \mathfrak{p}$.

Case 2: $x \in R \setminus A_w$. Since w is Manis there exists some $y \in \mathfrak{p}_w$ with $xy \in A_w \setminus \mathfrak{p}_w$. We have $\varphi(xy) \neq 0$. Since u has support zero there exists some $z \in A_w$ with $\varphi(xy)\varphi(z) \in A_u \setminus \mathfrak{p}_u$, hence $xyz \in A \setminus \mathfrak{p}$. Clearly $yz \in \mathfrak{p}_w \subset \mathfrak{p}$.

b) Assume now that $(A, \mathfrak{p})$ is Manis in R, and that $v: R \twoheadrightarrow \Gamma \cup \infty$ is a surjective valuation with $A_v = A$, $\mathfrak{p}_v = \mathfrak{p}$. We verify that u has support zero and prove the second part of the proposition. Since w is not trivial, we know from Theorem 6 that w is a coarsening of v. There is a unique convex subgroup H of Γ with $w \sim v/H$, and $A_w = A_H$, $\mathfrak{p}_w = \mathfrak{p}_H$ (notations from 1.18). We obtain from v and H a Manis valuation $v_H: A_w \twoheadrightarrow H \cup \infty$ with support $\mathfrak{p}_w$, as explained in 1.18. The pair associated to v_H is $(A, \mathfrak{p})$. Thus $v_H \sim u \circ \varphi$ and $\overline{v_H} \sim u$. In particular $\operatorname{supp} u = \operatorname{supp} \overline{v_H} = \{0\}$. $\square$

We now consider the following situation: A is a subring of R and $\mathfrak{p}$ is a prime ideal of A. We are looking for criteria that the pair $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ (cf. §1, Def. 10) is Manis.

We need an easy lemma.

Lemma 2.9. a) $R_\mathfrak{p} = R_{(\mathfrak{p}_{[\mathfrak{p}]})}$.

b) *If M is an A-submodule of R then $M_\mathfrak{p} = (M_{[\mathfrak{p}]})_{\mathfrak{p}_{[\mathfrak{p}]}}$.*

c) *If M is an A-submodule of R and $\mathfrak{r}$ is a prime ideal of A contained in $\mathfrak{p}$, then*

$$M_{[\mathfrak{r}]} = (M_{[\mathfrak{p}]})_{[\mathfrak{r}_{[\mathfrak{p}]}]}.$$

Proof. We have $R_\mathfrak{p} = S^{-1}R$ and $R_{(\mathfrak{p}_{[\mathfrak{p}]})} = T^{-1}R$ with $S = A \setminus \mathfrak{p}$, $T = A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. Notice that $S \subset T$. Let $x \in T$ be given. Choose some $d \in S$ with $dx \in A$. Then $dx \in A \setminus \mathfrak{p} = S$. This proves that $\operatorname{Sat}_R(S) = \operatorname{Sat}_R(T)$, and we conclude that $S^{-1}R = T^{-1}R$.

If M is an A-submodule of R, then $M_{[\mathfrak{p}]}$ is an $A_{[\mathfrak{p}]}$-submodule of R, and $M_{\mathfrak{p}} = S^{-1}M$, $(M_{[\mathfrak{p}]})_{\mathfrak{p}_{[\mathfrak{p}]}} = T^{-1}M_{[\mathfrak{p}]}$. Clearly $S^{-1}M \subset T^{-1}M_{[\mathfrak{p}]}$. (N.B. Both are subsets of $S^{-1}R = T^{-1}R$.) Also $T^{-1}M_{[\mathfrak{p}]} = S^{-1}M_{[\mathfrak{p}]}$. Let $z \in S^{-1}M_{[\mathfrak{p}]}$ be given. Write $z = \frac{x}{s}$ with $x \in M_{[\mathfrak{p}]}$, $s \in S$. We choose some $d \in S$ with $dx = m \in M$. We have $z = \frac{m}{sd} \in M_{\mathfrak{p}}$. This proves part b) of the lemma. The last statement c) follows from the obvious equality $M_{\mathfrak{r}} = (M_{\mathfrak{p}})_{\mathfrak{r}_{\mathfrak{p}}}$ by taking pre-images under the various localization maps. □

Proposition 2.10. *$(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R iff $(A_{\mathfrak{p}}, \mathfrak{p}_{\mathfrak{p}})$ is a Manis pair in $R_{\mathfrak{p}}$. In this case, if $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ comes from the Manis valuation v on R, then $(A_{\mathfrak{p}}, \mathfrak{p}_{\mathfrak{p}})$ comes from the localization $\tilde{v}$ of v defined in §1 (Def. 6). {Recall from the lemma that $A_{\mathfrak{p}} = (A_{[\mathfrak{p}]})_{\mathfrak{p}_{[\mathfrak{p}]}}, \mathfrak{p}_{\mathfrak{p}} = (\mathfrak{p}_{[\mathfrak{p}]})_{\mathfrak{p}_{[\mathfrak{p}]}}$.} With $\mathfrak{q} := A \cap \operatorname{supp} v$ we have $\operatorname{supp} v = \mathfrak{q}_{[\mathfrak{p}]}, \operatorname{supp} \tilde{v} = \mathfrak{q}_{\mathfrak{p}}$.*

Proof. a) Assume first that there exists a Manis valuation $v: R \to \Gamma \cup \infty$ with $A_v = A_{[\mathfrak{p}]}, \mathfrak{p}_v = \mathfrak{p}_{[\mathfrak{p}]}$. Let $\tilde{v}: R_{\mathfrak{p}_v} \to \Gamma \cup \infty$ denote the localization of v. Then $\tilde{v}$ is again Manis and $A_{\tilde{v}} = (A_v)_{\mathfrak{p}_v}, \mathfrak{p}_{\tilde{v}} = (\mathfrak{p}_v)_{\mathfrak{p}_v}$, $\operatorname{supp} \tilde{v} = (\operatorname{supp} v)_{\mathfrak{p}_v}$ (cf. §1). By part a) of the lemma above we have $R_{\mathfrak{p}_v} = R_{\mathfrak{p}}$, $A_{\tilde{v}} = A_{\mathfrak{p}}, \mathfrak{p}_{\tilde{v}} = \mathfrak{p}_{\mathfrak{p}}$. Let $\mathfrak{q} := A \cap \operatorname{supp} v$. Certainly $\mathfrak{q}_{[\mathfrak{p}]} \subset \operatorname{supp} v$. Let $x \in \operatorname{supp} v$ be given. We have $x \in A_v = A_{[\mathfrak{p}]}$. We choose some $d \in A \setminus \mathfrak{p}$ with $dx \in A$. Then $v(dx) = \infty$, thus $dx \in A \cap \operatorname{supp} v = \mathfrak{q}$, $x \in \mathfrak{q}_{[\mathfrak{p}]}$. This proves $\operatorname{supp} v = \mathfrak{q}_{[\mathfrak{p}]}$. Using part b) of the lemma we obtain $\operatorname{supp} \tilde{v} = \mathfrak{q}_{\mathfrak{p}}$.
b) Assume finally that $w: R_{\mathfrak{p}} \to \Gamma \cup \infty$ is a Manis valuation with $A_w = A_{\mathfrak{p}}$, $\mathfrak{p}_w = \mathfrak{p}_{\mathfrak{p}}$. Let $j_T: R \to R_{\mathfrak{p}}$ denote the localization map of R with respect to $T := A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. Let v denote the valuation $w \circ j_T$ on R. We have $v(T) = \{0\}$. Thus $v(R) = w(R_{\mathfrak{p}}) = \Gamma_w$, and we conclude that v is Manis. Also $A_v = j_T^{-1}(A_w) = A_{[\mathfrak{p}]}$, $\mathfrak{p}_v = j_T^{-1}(\mathfrak{p}_w) = \mathfrak{p}_{[\mathfrak{p}]}$, and w coincides with the localization $\tilde{v}$ of v. q.e.d.

Proposition 2.11. *Let $\mathfrak{r}$ be a prime ideal of A contained in $\mathfrak{p}$. Assume that $v: R \to \Gamma \cup \infty$ is a valuation with $A_v = A_{[\mathfrak{p}]}$, $\mathfrak{p}_v = \mathfrak{p}_{[\mathfrak{p}]}$, $A \cap \operatorname{supp} v \subset \mathfrak{r}$. Let H denote the convex subgroup of Γ generated by $v(A \setminus \mathfrak{r})$ and let $w := v/H$. Then $A_w = A_{[\mathfrak{r}]}$, $\mathfrak{p}_w = \mathfrak{r}_{[\mathfrak{p}]} = \mathfrak{r}_{[\mathfrak{r}]}$. Thus, if $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R the same holds for $(A_{[\mathfrak{r}]}, \mathfrak{r}_{[\mathfrak{r}]})$.*

Proof. By the last statement in Prop. 10 we have $\operatorname{supp} v \subset \mathfrak{r}_{[\mathfrak{p}]}$. It follows from Proposition 1.13 and part c) of lemma 9 above that $A_w = A_{[\mathfrak{r}]}$, $\mathfrak{p}_w = \mathfrak{r}_{[\mathfrak{p}]}$. It is evident that $\mathfrak{r}_{[\mathfrak{p}]} \subset \mathfrak{r}_{[\mathfrak{r}]} \subset \mathfrak{p}_w$. Thus $\mathfrak{r}_{[\mathfrak{p}]} = \mathfrak{r}_{[\mathfrak{r}]}$. □

We now state a criterion which will play a key role for the theory of relative Prüfer rings in §5.

Theorem 2.12. *Assume that A is integrally closed in R. The following are equivalent.*

i) *$(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R.*
ii) *For each $x \in R$ there exists some polynomial $F[T] \in A[T] \setminus \mathfrak{p}[T]$ with $F(x) = 0$.*

Proof. i) $\Rightarrow$ ii): We first consider the case that $x \in A_{[\mathfrak{p}]}$. We choose some $s \in A \setminus \mathfrak{p}$ with $sx = a \in A$. The polynomial $F(T) := sT - a$ fulfills the requirements. Let now $x \in R \setminus A_{[\mathfrak{p}]}$. Since $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair there exists some $y \in \mathfrak{p}_{[\mathfrak{p}]}$ with $xy \in A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. We choose elements s and t in $A \setminus \mathfrak{p}$ with $ty \in \mathfrak{p}$, $sxy \in A$. We have $sxy \in A \setminus \mathfrak{p}$. Put $a_0 := sty \in \mathfrak{p}$, $a_1 := -stxy \in A \setminus \mathfrak{p}$. The polynomial $F(T) := a_0T + a_1$ fulfills the requirements.

ii) $\Rightarrow$ i): We verify the property (iii) in Theorem 4. Let $x \in R \setminus A_{[\mathfrak{p}]}$ be given. We look for an element $y \in \mathfrak{p}_{[\mathfrak{p}]}$ with $xy \in A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. Let

$$F(T) := a_0T^n + a_1T^{n-1} + \cdots + a_n$$

be a polynomial of minimal degree $n \geq 1$ in $A[T] \setminus \mathfrak{p}[T]$ with $F(x) = 0$. From $F(x) = 0$ we deduce that $b := a_0x$ is integral over A. Thus $b \in A$. Since $x \notin A_{[\mathfrak{p}]}$ we conclude that $a_0 \in \mathfrak{p}$. Suppose that $n > 1$. We put

$$G(T) := a_0T - b \quad \text{in the case} \quad b \notin \mathfrak{p},$$

and

$$G(T) := (b + a_1)T^{n-1} + a_2T^{n-2} + \cdots + a_n$$

in the case $b \in \mathfrak{p}$. In both cases

$$G(T) \in A[T] \setminus \mathfrak{p}[T] \quad \text{and} \quad G(x) = 0.$$

This contradicts the minimality of n. Thus $n = 1$, $F(T) = a_0T + a_1$. Since $a_0 \in \mathfrak{p}$, certainly $a_1 \in A \setminus \mathfrak{p}$. For $y := a_0$ we have $y \in \mathfrak{p}_{[\mathfrak{p}]}$, $xy \in A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. q.e.d.

Essentially as a consequence of Theorems 4 and 12 we derive still another criterion for a pair $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ to be Manis in R. In the case of Krull valuation rings (i.e. R a field) such a criterion had been observed by Gilmer [Gi, Th. 19.15]. We need (a special case of) an easy lemma.

Lemma 2.13. *Let $(B, \mathfrak{q})$ be a Manis pair in R. Let I be a B-submodule of R with $I \cap B \subset \mathfrak{q}$. Then $I \subset \mathfrak{q}$.*

Proof. Suppose there exists an $x \in I$ with $x \notin \mathfrak{q}$, hence $x \notin B$. Since $(B, \mathfrak{q})$ is Manis there exists some $y \in B$ with $xy \in B \setminus \mathfrak{q}$. Then $xy \notin I$. On the other hand $x \in I$ and $y \in B$, a contradiction. □

Theorem 2.14 (cf. [Gi, Th. 19.15] for R a field). *Assume that A is integrally closed in R, and let $\mathfrak{p}$ be a prime ideal of A. The following are equivalent.*

i) *$(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R.*
ii) *If B is a subring of R containing $A_{[\mathfrak{p}]}$ and $\mathfrak{q}, \mathfrak{q}'$ are prime ideals of B with $\mathfrak{q} \subset \mathfrak{q}'$ and $\mathfrak{q} \cap A_{[\mathfrak{p}]} = \mathfrak{q}' \cap A_{[\mathfrak{p}]} \subset \mathfrak{p}_{[\mathfrak{p}]}$, then $\mathfrak{q} = \mathfrak{q}'$.*
ii′) *If B is a subring of R containing $A_{[\mathfrak{p}]}$ and $\mathfrak{q} \subset \mathfrak{q}'$ are prime ideals of B lying over $\mathfrak{p}_{[\mathfrak{p}]}$, then $\mathfrak{q} = \mathfrak{q}'$.*
iii) *If B is a subring of R containing A and $\mathfrak{q}, \mathfrak{q}'$ are prime ideals of B with $\mathfrak{q} \subset \mathfrak{q}'$ and $\mathfrak{q} \cap A = \mathfrak{q}' \cap A \subset \mathfrak{p}$ then $\mathfrak{q} = \mathfrak{q}'$.*
iii′) *If B is a subring of R containing A and $\mathfrak{q} \subset \mathfrak{q}'$ are prime ideals of B lying over $\mathfrak{p}$ then $\mathfrak{q} = \mathfrak{q}'$.*
iv) *There exists only one Manis pair $(B, \mathfrak{q})$ in R over $(A, \mathfrak{p})$, i.e. with $A \subset B$ and $\mathfrak{q} \cap A = \mathfrak{p}$.*

v) *For every subring B of R containing A there exists at most one prime ideal $\mathfrak{q}$ of B over $\mathfrak{p}$.*

vi) *For every Manis pair $(B, \mathfrak{q})$ in R over $(A, \mathfrak{p})$ the field extension $k(\mathfrak{p}) \subset k(\mathfrak{q})$ is algebraic.*

PROOF. The implication i) $\Rightarrow$ ii) is evident by the preceding lemma. The implications ii) $\Rightarrow$ ii$'$) and iii) $\Rightarrow$ iii$'$) are trivial.
ii$'$) $\Rightarrow$ iii$'$): If $\mathfrak{q}$ and $\mathfrak{q}'$ are prime ideals of B over $\mathfrak{p}$ with $\mathfrak{q} \subset \mathfrak{q}'$, then $\mathfrak{q}_{[\mathfrak{p}]}$ and $\mathfrak{q}'_{[\mathfrak{p}]}$ are prime ideals of $B_{[\mathfrak{p}]}$ over $\mathfrak{p}_{[\mathfrak{p}]}$ with $\mathfrak{q}_{[\mathfrak{p}]} \subset \mathfrak{q}'_{[\mathfrak{p}]}$. Thus $\mathfrak{q}_{[\mathfrak{p}]} = \mathfrak{q}'_{[\mathfrak{p}]}$. Intersecting with B we obtain $\mathfrak{q} = \mathfrak{q}'$. ii) $\Rightarrow$ iii): The proof is similar.
iii$'$) $\Rightarrow$ i): Suppose that $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is not Manis in R. By Theorem 12 there exists some $x \in R$ such that $F(x) \neq 0$ for every polynomial $F(T) \in A[T] \setminus \mathfrak{p}[T]$. We introduce the subring $B := A[x]$ of R and the surjective ring homomorphism $\varphi: A[T] \longrightarrow B$ over A with $\varphi(T) = x$. The kernel of φ is contained in $\mathfrak{p}[T]$. This implies that the ideals $\mathfrak{q}$ and $\mathfrak{q}'$ of B defined by

$$\mathfrak{q} := \varphi(\mathfrak{p}[T]) = \mathfrak{p}[x] = \mathfrak{p}B, \quad \mathfrak{q}' := \varphi(\mathfrak{p} + TA[T]) = \mathfrak{p} + xB = \mathfrak{q} + xB,$$

both are prime and lie over $\mathfrak{p}$. Since $\mathfrak{q} \neq \mathfrak{q}'$ this contradicts the assumption iii$'$). Thus $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is Manis in R.
i) $\Rightarrow$ iv): Let $(B, \mathfrak{q})$ be a Manis pair in R over $(A, \mathfrak{p})$. It is easily verified that $(B, \mathfrak{q})$ is a pair over $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$. Since $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is Manis in R we conclude by Theorem 4 that $(B, \mathfrak{q}) = (A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$.
iv) $\Rightarrow$ v): Assume that B is a subring of R containing A and $\mathfrak{q}_1$, $\mathfrak{q}_2$ are prime ideals of B over $\mathfrak{p}$. We extend the pairs $(B, \mathfrak{q}_1)$ and $(B, \mathfrak{q}_2)$ to maximal pairs $(C, \mathfrak{q}'_1)$ and $(D, \mathfrak{q}'_2)$ in R. These pairs are Manis in R by Theorem 4. They both lie over $(A, \mathfrak{p})$, hence $(C, \mathfrak{q}'_1) = (D, \mathfrak{q}'_2)$. Intersecting with B we obtain $\mathfrak{q}_1 = \mathfrak{q}_2$.
v) $\Rightarrow$ iii$'$): trivial.
i) $\Rightarrow$ vi): Since (i) and (iv) hold we know that $(B, \mathfrak{q}) := (A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is the only Manis pair in R over $(A, \mathfrak{p})$. We have $k(\mathfrak{p}) = k(\mathfrak{q})$.
vi) $\Rightarrow$ iii$'$): Suppose that $(B, \mathfrak{q}_1)$ and $(B, \mathfrak{q}_2)$ are pairs in R over $(A, \mathfrak{p})$ with $\mathfrak{q}_1 \subsetneqq \mathfrak{q}_2$. We choose a maximal pair $(C, \mathfrak{r})$ in R over $(B, \mathfrak{q}_1)$. Then $(C, \mathfrak{r})$ is Manis, hence $k(\mathfrak{r})$ is algebraic over $k(\mathfrak{p})$. It follows that $k(\mathfrak{q}_1)$ is algebraic over $k(\mathfrak{p})$. We choose an element $x \in \mathfrak{q}_2 \setminus \mathfrak{q}_1$. Since $k(\mathfrak{q}_1)$ is algebraic over $k(\mathfrak{p})$ we have a relation

$$(*) \qquad \sum_{i=0}^{n} a_i x^i = b$$

with $a_0, a_1, \ldots, a_n \in A$, $a_n \notin \mathfrak{p}$, $b \in \mathfrak{q}_1$. Let B' denote the subring $A[b, a_n x]$ of B, and $\mathfrak{q}'_1 := \mathfrak{q}_1 \cap B'$, $\mathfrak{q}'_2 := \mathfrak{q}_2 \cap B'$. We have $\mathfrak{q}'_1 \subsetneqq \mathfrak{q}'_2$, since $a_n x \in \mathfrak{q}'_2 \setminus \mathfrak{q}'_1$. But $\mathfrak{q}'_1 \cap A = \mathfrak{q}'_2 \cap A = \mathfrak{p}$. We learn from the relation $(*)$ that $B'/\mathfrak{q}'_1$ is integral over $A/\mathfrak{p}$. But the ring $B'/\mathfrak{q}'_1$ contains the prime ideal $\mathfrak{q}'_2/\mathfrak{q}'_1 \neq \{0\}$ with $(\mathfrak{q}'_2/\mathfrak{q}'_1) \cap A/\mathfrak{p} = \{0\}$. Such a situation is impossible in an integral ring extension (cf. [Bo, V §2, n^o 1]). Thus (iii$'$) is valid. $\square$

§3 Weakly surjective homomorphisms

In section §5 we will start our theory of "Prüfer extensions". In the terminology developed there the Prüfer rings (with zero divisors) of the classical literature (e.g. [LM], [Huc]) are those commutative rings A which are Prüfer in their total quotient rings Quot$A$. In the present section and the following one we develop an auxiliary theory of "weakly surjective" ring extensions. The inclusions $A \subset \mathrm{Quot}A$ are (very special) examples of such extensions.

Definition 1. i) Let $\varphi: A \to B$ be a ring homomorphism. We call φ *locally surjective* (abbreviated: ls) if for every prime ideal $\mathfrak{q}$ of B the induced homomorphism $\varphi_\mathfrak{q}: A_{\varphi^{-1}(\mathfrak{q})} \to B_\mathfrak{q}$ is surjective. We call φ *weakly surjective* (abbreviated: ws) if for every prime ideal $\mathfrak{p}$ of A with $\mathfrak{p}B \neq B$ the induced homomorphism $\varphi_\mathfrak{p}: A_\mathfrak{p} \to B_\mathfrak{p}$ is surjective.
ii) If A is a subring of a ring B, then we say that A is *locally surjective in* B (resp. *weakly surjective in* B) if the inclusion mapping $A \hookrightarrow B$ is ls (resp. ws).

At first glance "locally surjective" seems to be a more natural notion than "weakly surjective", but it is the latter notion which will be needed below.

Of course, a surjective homomorphism is both weakly surjective and locally surjective. We now prove that weak surjectivity is a stronger property than local surjectivity.

Proposition 3.1. *If $\varphi: A \to B$ is weakly surjective then φ is locally surjective.*

This follows from

Lemma 3.2. *Let $\varphi: A \to B$ be a ring homomorphism. Let $\mathfrak{q}$ be a prime ideal of B and $\mathfrak{p} := \varphi^{-1}(\mathfrak{q})$. Assume that $\varphi_\mathfrak{p}: A_\mathfrak{p} \to B_\mathfrak{p}$ is surjective. Then the natural map $B_\mathfrak{p} \to B_\mathfrak{q}$ is an isomorphism, in short, $B_\mathfrak{p} = B_\mathfrak{q}$. Furthermore $\mathfrak{p}B_\mathfrak{p} = \mathfrak{p}B_\mathfrak{q} = \mathfrak{q}B_\mathfrak{q}$.*

Proof of the lemma. One easily retreats to the case that A is a subring of B and φ is the inclusion $A \hookrightarrow B$. Now $\mathfrak{p} = \mathfrak{q} \cap A$ and $A_\mathfrak{p} = B_\mathfrak{p}$. We have $\mathfrak{p}A_\mathfrak{p} = \mathfrak{p}B_\mathfrak{p} \subset \mathfrak{q}B_\mathfrak{p}$. Since $\mathfrak{p}A_\mathfrak{p}$ is the maximal ideal of $A_\mathfrak{p}$ and $(\mathfrak{q}B_\mathfrak{p}) \cap B = \mathfrak{q}$, hence $\mathfrak{q}B_\mathfrak{p} \neq B_\mathfrak{p}$, we have $\mathfrak{p}B_\mathfrak{p} = \mathfrak{q}B_\mathfrak{p}$. The natural homomorphism $B \to B_\mathfrak{p}$ maps $B \setminus \mathfrak{q}$ into the group of units of $B_\mathfrak{p}$, hence factors through a homomorphism from $B_\mathfrak{q}$ to $B_\mathfrak{p}$. This homomorphism is inverse to the natural map from $B_\mathfrak{p}$ to $B_\mathfrak{q}$. □

Example 3.3. If S is a multiplicative subset of a ring A then the localization map $A \to S^{-1}A$ is weakly surjective. □

Example 3.4. Let K be a field. The diagonal homomorphism $K \to K \times K$, $x \mapsto (x, x)$, is locally surjective but not weakly surjective, as is easily verified.

Proposition 3.5. *If $\varphi: A \to B$ is locally surjective and B is an integral domain then φ is weakly surjective.*

Proof. Let $\mathfrak{p}$ be a prime ideal of A with $\mathfrak{p}B \neq B$. We choose a prime ideal $\mathfrak{q}$ of B containing $\mathfrak{p}B$. Let $\mathfrak{r} := \varphi^{-1}(\mathfrak{q})$. We have a natural commuting triangle

$$\begin{array}{ccc} A_{\mathfrak{r}} & \xrightarrow{\varphi_{\mathfrak{r}}} & B_{\mathfrak{r}} \\ & {\scriptstyle \varphi_{\mathfrak{q}}} \searrow \quad \swarrow {\scriptstyle \psi} & \\ & B_{\mathfrak{q}} & . \end{array}$$

$\varphi_{\mathfrak{q}}$ is surjective since φ is ls. On the other hand ψ is injective since B is a domain. Thus ψ is bijective and $\varphi_{\mathfrak{r}}$ is surjective. (We have $B_{\mathfrak{r}} = B_{\mathfrak{q}}$, $\varphi_{\mathfrak{r}} = \varphi_{\mathfrak{q}}$.) Since $\mathfrak{p} \subset \mathfrak{r}$ also $\varphi_{\mathfrak{p}}$ is surjective. □

Proposition 3.6. *Every locally surjective homomorphism is an epimorphism in the category $\mathcal{R}$ of rings (commutative, with 1).*

Proof. Assume that $\varphi: A \to B$ is locally surjective, and that $\psi_1: B \to C$, $\psi_2: B \to C$ are two ring homomorphisms with $\psi_1 \circ \varphi = \psi_2 \circ \varphi$. For every prime ideal $\mathfrak{q}$ of B the map $\varphi_{\mathfrak{q}}: A_{\varphi^{-1}(\mathfrak{q})} \to B_{\mathfrak{q}}$ is surjective, thus $\psi_{1\mathfrak{q}} = \psi_{2\mathfrak{q}}$. We conclude that $\psi_1 = \psi_2$ (cf. [Bo, Chap II, §3]). □

A fortiori every ws map is an epimorphism in $\mathcal{R}$. We now verify that this class of epimorphisms has pleasant formal properties.

Proposition 3.7. *Let $\varphi: A \to B$ and $\psi: B \to C$ be ring homomorphisms.*
a) If both φ and ψ are weakly surjective then $\psi \circ \varphi$ is weakly surjective.
b) If $\psi \circ \varphi$ is weakly surjective then ψ is weakly surjective.

Proof. a): Let $\mathfrak{p}$ be a prime ideal of A with $\mathfrak{p}C \neq C$. We choose a prime ideal $\mathfrak{r}$ of C containing $\mathfrak{p}C$. Let $\mathfrak{q} := \psi^{-1}(\mathfrak{r})$ and $\tilde{\mathfrak{p}} := \varphi^{-1}(\mathfrak{q})$. The map $\varphi_{\tilde{\mathfrak{p}}}: A_{\tilde{\mathfrak{p}}} \to B_{\tilde{\mathfrak{p}}}$ is surjective. By lemma 3.2 we know that $B_{\mathfrak{q}} = B_{\tilde{\mathfrak{p}}}$. Thus also $C_{\tilde{\mathfrak{p}}} = C \otimes_A A_{\tilde{\mathfrak{p}}} = C \otimes_B (B \otimes_A A_{\tilde{\mathfrak{p}}}) = C \otimes_B B_{\tilde{\mathfrak{p}}} = C \otimes_B B_{\mathfrak{q}} = C_{\mathfrak{q}}$, and $\psi_{\tilde{\mathfrak{p}}} = \psi_{\mathfrak{q}}$, which is surjective. We conclude that $(\psi \circ \varphi)_{\tilde{\mathfrak{p}}} = \psi_{\tilde{\mathfrak{p}}} \circ \varphi_{\tilde{\mathfrak{p}}}$ is surjective.
b): Let $\mathfrak{q}$ be a prime ideal of B with $\mathfrak{q}C \neq C$. Let $\mathfrak{p} := \varphi^{-1}(\mathfrak{q})$. The map $\psi_{\mathfrak{p}} \circ \varphi_{\mathfrak{p}} = (\psi \circ \varphi)_{\mathfrak{p}}$ is surjective. Thus $\psi_{\mathfrak{p}}$ is surjective. Since $\varphi(A \setminus \mathfrak{p}) \subset B \setminus \mathfrak{q}$ also $\psi_{\mathfrak{q}}$ is surjective.
□

Proposition 3.8. *If $\varphi: A \to B$ and $\psi: B \to C$ are ring homomorphisms and φ is ws then $\psi\varphi(A)$ is ws in $\psi(B)$.*

Proof. We have a commuting square

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ {\scriptstyle p} \downarrow & & \downarrow {\scriptstyle q} \\ \psi\varphi(A) & \underset{i}{\hookrightarrow} & \psi(B) \end{array}$$

with i an inclusion mapping and surjections p and q. Since φ and q are ws, the composite $q \circ \varphi = i \circ p$ is ws. Thus also i is ws. □

Corollary 3.9. *Let $\varphi: A \to B$ a ring homomorphism. φ is ws iff $\varphi(A)$ is ws in B.*

Proof. Applying Proposition 8 with $\psi = id_B$ we see that weak surjectivity of φ implies weak surjectivity of the inclusion mapping $i: \varphi(A) \hookrightarrow B$. Conversely, if i is ws, then φ is ws, since $\varphi = i \circ p$ with p a surjection. □

It is also easy to verify the corollary directly by using Definition 1.

PROPOSITION 3.10. *Let*

$$\begin{array}{ccc} A & \xrightarrow{\varphi} & B \\ \alpha\downarrow & & \downarrow\beta \\ C & \xrightarrow[\psi]{} & D \end{array}$$

be a commuting square of ring homomorphisms. Assume that φ is ws and $D = \beta(B) \cdot \psi(C)$. Then ψ is ws.

PROOF. Let $\mathfrak{q} \in \mathrm{Spec} C$ be given with $\psi(\mathfrak{q})D \neq D$, and let $\mathfrak{p} := \alpha^{-1}(\mathfrak{q})$. The commuting square above "extends" to a commuting square

$$\begin{array}{ccc} A_{\mathfrak{p}} & \xrightarrow{\tilde{\varphi}} & B_{\mathfrak{p}} \\ \tilde{\alpha}\downarrow & & \downarrow\tilde{\beta} \\ C_{\mathfrak{q}} & \xrightarrow[\tilde{\psi}]{} & D_{\mathfrak{q}} \end{array}$$

with $\tilde{\varphi} = \varphi_{\mathfrak{p}}$, $\tilde{\psi} = \psi_{\mathfrak{q}}$. We have $\mathfrak{p}B \neq B$. The map $\tilde{\varphi}$ is surjective. We are done, if we verify that $\tilde{\psi}$ is surjective.

Let $\xi \in D_{\mathfrak{q}}$ be given. Write $\xi = \frac{x}{s}$ with $x \in D$, $s \in C \setminus \mathfrak{q}$. Since $D = \beta(B)\psi(C)$ we have an equation

$$x = \sum_{i \in I} \beta(b_i)\psi(c_i)$$

with finite index set I, $b_i \in B$, $c_i \in C$. This equation gives us

$$\xi = \sum_{i \in I} \tilde{\beta}(\frac{b_i}{1})\tilde{\psi}(\frac{c_i}{s}).$$

Since $\tilde{\varphi}$ is surjective we have elements $a_i \in A \quad (i \in I)$ and an element $t \in A \setminus \mathfrak{p}$ with $\frac{b_i}{1} = \tilde{\varphi}(\frac{a_i}{t})$ for every $i \in I$. Then

$$\xi = \tilde{\psi}(\frac{y}{s\alpha(t)})$$

with $y := \sum\limits_{\alpha \in I} \alpha(a_i)c_i$. This proves that $\tilde{\psi}$ is surjective. $\square$

In order to understand weakly surjective homomorphisms it suffices by Cor. 9 to analyze weakly surjective ring extensions.

In the following R is a ring and A is a subring of R.

DEFINITION 2. An *R-overring of A* is a subring B of R with $A \subset B$.

PROPOSITION 3.11.

a) *Let B_1 and B_2 be R-overrings of A. If A is ws both in B_1 and B_2 then A is ws in B_1B_2.*

b) *There exists a unique R-overring $M(A,R)$ of A such that A is ws in $M(A,R)$ and $M(A,R)$ contains every R-overring of A in which A is ws.*

PROOF. a) Since $A \hookrightarrow B_1$ is ws, the inclusion $B_2 \hookrightarrow B_1B_2$ is ws, as follows from Proposition 10. Since also $A \hookrightarrow B_2$ is ws, the composite $A \hookrightarrow B_2 \hookrightarrow B_1B_2$ is ws (Prop. 7).
b) Let $\mathfrak{A}$ denote the set of all R-overrings of A in which A is ws. Then $\mathfrak{A}$ is an upward directed system of subrings of R. Let $M(A,R)$ denote the union of all these subrings, which is again a subring of R. A is ws in $M(A,R)$ by the following general remark, which is immediate from Definition 1.

REMARK 3.12. Let $(B_i | i \in I)$ be an upward directed system of R-overrings of A. If A is ws in each B_i then A is ws in $\bigcup\limits_{i \in I} B_i$.

DEFINITION 3. We call $M(A,R)$ the *weakly surjective hull* of A in R.

We now derive criteria for a homomorphism to be weakly surjective. Without essential loss of generality we concentrate on ring extensions. Let R be a ring and A a subring of R. Recall from §2 that for $\mathfrak{p}$ a prime ideal of A we denote by $A_{[\mathfrak{p}]}$ the pre-image of $A_{\mathfrak{p}}$ under the localization map $R \to R_{\mathfrak{p}}$.

NOTATION. If $x \in R$ then $(A\colon x)$ denotes the ideal of A consisting of all $a \in A$ with $ax \in A$.

THEOREM 3.13 (cf. [G_1, Prop. 10] in the case $R = \mathrm{Quot}A$). *Let B be an R-overring of A. The following are equivalent.*

(1) *A is weakly surjective in B.*
(2) *$B_{[\mathfrak{q}]} = A_{[\mathfrak{q} \cap A]}$ for every prime ideal $\mathfrak{q}$ of B.*
(2′) *$B_{[\mathfrak{q}]} = A_{[\mathfrak{q} \cap A]}$ for every maximal ideal $\mathfrak{q}$ of B.*
(3) *$B \subset A_{[\mathfrak{p}]}$ for every prime ideal $\mathfrak{p}$ of A with $\mathfrak{p}B \neq B$.*
(4) *$(A\colon x)B = B$ for every $x \in B$.*

PROOF. (1) $\Longleftrightarrow$ (3): We verify the following: For any $\mathfrak{p} \in \mathrm{Spec}A$

$$B \subset A_{[\mathfrak{p}]} \Longleftrightarrow B_{\mathfrak{p}} = A_{\mathfrak{p}}.$$

Then we will be done according to Def. 1.

$\Rightarrow$: If $B \subset A_{[\mathfrak{p}]}$, then $B_{\mathfrak{p}} \subset (A_{[\mathfrak{p}]})_{\mathfrak{p}} = A_{\mathfrak{p}}$.
$\Leftarrow$: If $B_{\mathfrak{p}} = A_{\mathfrak{p}}$ then the pre-image $A_{[\mathfrak{p}]}$ of $A_{\mathfrak{p}}$ under the localization map $R \to R_{\mathfrak{p}}$ contains B.

(3) $\Rightarrow$ (2): Let $\mathfrak{q} \in \mathrm{Spec}B$ and $\mathfrak{p} := \mathfrak{q} \cap A$. Of course, $A_{[\mathfrak{p}]} \subset B_{[\mathfrak{q}]}$. In order to prove the converse inclusion we first remark that $\mathfrak{p}B \subset \mathfrak{q}$, hence $\mathfrak{p}B \neq B$. By hypothesis $B \subset A_{[\mathfrak{p}]}$. Let $x \in B_{[\mathfrak{q}]}$ be given. Choose $b \in B \setminus \mathfrak{q}$ with $bx =: b_1 \in B$. We then have elements a, a_1 in $A \setminus \mathfrak{p}$ with $ab \in A$, $a_1b_1 \in A$. Since $a \in B \setminus \mathfrak{q}$, also $ab \in B \setminus \mathfrak{q}$, hence $ab \in A \cap (B \setminus \mathfrak{q}) = A \setminus \mathfrak{p}$. Also $a_1ab \in A \setminus \mathfrak{p}$. From $(a_1ab)x = a(a_1bx) = a(a_1b_1) \in A$ we see that $x \in A_{[\mathfrak{p}]}$.
(2) $\Rightarrow$ (2′): trivial.

$(2') \Rightarrow (4)$: Let $x \in B$ be given. Suppose that $(A\colon x)B \neq B$. We choose a maximal ideal $\mathfrak{q}$ of B containing $(A\colon x)B$. Let $\mathfrak{p} := \mathfrak{q} \cap A$. Then $(A\colon x) \subset \mathfrak{p}$. But it follows from $(2')$ that $x \in A_{[\mathfrak{p}]}$, i.e. $(A\colon x) \not\subset \mathfrak{p}$. This contradiction proves that $(A\colon x)B = B$.
$(4) \Rightarrow (3)$: Let $\mathfrak{p}$ be a prime ideal of A with $\mathfrak{p}B \neq B$. Suppose there exists some $x \in B$ with $x \notin A_{[\mathfrak{p}]}$. Then $(A\colon x) \subset \mathfrak{p}$. Thus $(A\colon x)B \subset \mathfrak{p}B \subsetneqq B$. This contradicts the assumption (4). We conclude that $B \subset A_{[\mathfrak{p}]}$. □

Remarks. In the case of domains Richman [Ri, §2] has studied the properties (3), (4) under the name "good extensions". If $A \subset B$ and B is a domain then good means the same as weakly surjective and as locally surjective. Theorem 13 has a close relation to work of Lazard [L, Chap. IV] and Akiba [A], cf. Theorem 4.4 in the next section.

Definition 4. [Lb, §2.3].

a) An ideal $\mathfrak{a}$ of a ring C is called *dense in* C if its annulator ideal $\mathrm{Ann}_C(\mathfrak{a})$ is zero.
b) A *ring of quotients of* A is a ring $B \supset A$ such that $(A\colon x)B$ is dense in B for every $x \in B$.

We recall the following important fact from Lambek's book [Lb, §2.3]. For any ring A there exists a ring of quotients $Q(A)$ of A, explicitly constructed in [Lb], such that for any other ring of quotients B of A there exists a unique homomorphism from B to $Q(A)$ over A. Every such homomorphism is injective. $Q(A)$ is called the *complete ring of quotients* of A. Of course $Q(A)$ contains the total quotient ring $\mathrm{Quot}(A)$ {also called the "classical" quotient ring of A}. For A Noetherian it is known that $\mathrm{Quot}A = Q(A)$, cf. [A, Prop. 1], but in general these two extensions of A may be different.

From condition (4) in Theorem 13 it is clear that, if $A \subset B$ is a weakly surjective ring extension, then B is a ring of quotients of A. Thus every weakly surjective ring extension of A embeds into $Q(A)$ in a unique way.

Definition 5. The *weakly surjective hull* $M(A)$ of A is defined as the ws hull $M(A, Q(A))$ of A in $Q(A)$.

From our discussion of the hulls $M(A,R)$ above the following is evident.

Proposition 3.14. *For every weakly surjective ring extension $A \subset B$ there exists a unique homomorphism $B \to M(A)$ over A, and this is a monomorphism.* □

Thus, without serious abuse, we may regard any ws extension $A \subset B$ as a subextension of $A \subset M(A)$. In particular, $A \subset \mathrm{Quot}A \subset M(A)$.

Remark 3.15. If C is any subring of $M(A)$ containing A then $M(C) = M(A)$. In particular, $MM(A) = M(A)$.

Proof. Since C is ws in $M(A)$ we have embeddings $C \subset M(A) \subset M(C)$. Now A is ws in $M(A)$ and $M(A)$ is ws in $M(C)$, hence A is ws in $M(C)$. Due to the maximality of $M(A)$ we have $M(C) = M(A)$. □

Caution. In general, if C is a subring of $M(A)$ containing A, then A is not necessarily ws in C (cf. §5).

Corollary 3.16. *Let $A \subset B_1$ and $A \subset B_2$ be weakly surjective extensions. Then there exists at most one homomorphism $\lambda: B_1 \to B_2$ over A, and λ is injective.*

Proof. We have unique homomorphisms $\mu_i: B_i \to M(A)$ over A $(i = 1, 2)$, and they both are injective. If $\lambda: B_1 \to B_2$ is a homomorphism over A, this implies that $\mu_2 \circ \lambda = \mu_1$. Thus λ is injective and is uniquely determined by μ_1 and μ_2. □

Of course, the uniqueness of λ is a priori clear, since $A \hookrightarrow B_1$ is an epimorphism (Prop. 6).

§4 More on weakly surjective extensions

Having set the stage we discuss some properties of weakly surjective ring extensions. We are mainly interested in functorial properties and the behavior of ideals.

In the following *we assume that $A \subset B$ is a weakly surjective ring extension.*

Proposition 4.1. *Every weakly surjective ring extension $A \subset B$ is flat (i.e., B is a flat A-module).*

Proof. Let $\alpha: M' \to M$ be an injective homomorphism of A-modules.
We verify that $\alpha \otimes_A B: M' \otimes_A B \to M \otimes_A B$ is again injective. Let $\mathfrak{q}$ be a prime ideal of B and $\mathfrak{p} := \mathfrak{q} \cap A$. Then $A_\mathfrak{p} = B_\mathfrak{q}$, thus

$$(\alpha \otimes_A B)_\mathfrak{q} = (\alpha \otimes_A B) \otimes_B B_\mathfrak{q} = \alpha \otimes_A B_\mathfrak{q} = \alpha \otimes_A A_\mathfrak{p}.$$

Since $A \to A_\mathfrak{p}$ is flat the homomorphism $(\alpha \otimes_A B)_\mathfrak{q}$ is injective. Since this holds for every $\mathfrak{q} \in \mathrm{Spec} B$ we conclude that $\alpha \otimes_A B$ is injective. □

Proposition 4.2. *Let $A \subset B_1$ and $A \subset B_2$ be weakly surjective ring extensions.*

a) *Then the natural map $A \to B_1 \otimes_A B_2$ is injective and weakly surjective, hence may be regarded as a ws extension.*

b) *If both $A \subset B_1$ and $A \subset B_2$ are subextensions of a ring extension $A \subset R$, then the natural map $B_1 \otimes_A B_2 \to B_1 B_2$ is an isomorphism, in short, $B_1 \otimes_A B_2 = B_1 B_2$.*

Proof. a) Since B_1 is flat over A the natural map $B_1 \to B_1 \otimes_A B_2$ is injective. Also $B_2 \to B_1 \otimes_A B_2$ and $A \to B_1 \otimes B_2$ are injective. We regard A, B_1, B_2 as subrings of $B_1 \otimes_A B_2$ and conclude from Propositions 3.7.a and 3.8. that A is ws in $B_1 \otimes B_2$.
b) In the situation $B_1 \subset R$, $B_2 \subset R$ the ring A is also ws in $B_1 B_2$. The natural map $\lambda: B_1 \otimes_A B_2 \to B_1 B_2$ is a surjective homomorphism over A. By Cor.3.16 λ is also injective, hence is an isomorphism. □

Example 4.3. If $\varphi: A \to B$ is a weakly surjective homomorphism then the natural map $B \otimes_A B \longrightarrow B, x \otimes y \longmapsto xy$, is an isomorphism.

This follows from the proposition since $B \otimes_A B = B \otimes_{\varphi(A)} B$. The statement is just a reformulation of the fact, already known to us (Prop. 3.6), that φ is an epimorphism, cf. e.g. [St, p. 380].

We now invoke the important work of Lazard in his thesis [L] and of Akiba [A]. We have seen that every injective weakly surjective homomorphism is a flat epimorphism (in the category $\mathcal{R}$ of rings). By [L, IV. Prop. 2.4] or [A, Th.1] the converse also holds.

Theorem 4.4 (Lazard, Akiba). *An injective homomorphism φ is weakly surjective iff φ is a flat epimorphism.* $\square$

Proposition 4.5. *Let $A \subset B$ be a weakly surjective extension and C a subring of B containing A. Then $A \subset C$ is weakly surjective iff C is flat over A.*

Proof. We know already that weak surjectivity of $A \hookrightarrow C$ implies flatness. Conversely, if $A \hookrightarrow C$ is flat then $A \hookrightarrow C$ is epimorphic by the theory of Lazard [L, IV Cor. 3.2], hence is ws. $\square$

Up to very minor points also the results to follow, up to Proposition 10, are contained in Lazard's thesis [L], and many more. For the convenience of the reader we give short proofs in the present frame work. Our focus is different from Lazard's, since we only strive for the understanding of a special class of flat epimorphic extensions, the Prüfer extensions to be defined in §5.

As before we are given a ws extension $A \subset B$.

Proposition 4.6. *Let $\mathfrak{b}$ be an ideal of B and $\mathfrak{a} := \mathfrak{b} \cap A$. Then $\mathfrak{b} = \mathfrak{a}B$.*

Proof. Let $\mathfrak{c} := \mathfrak{a}B$. Then $\mathfrak{c} \subset \mathfrak{b}$ and $\mathfrak{c} \cap A = \mathfrak{a}$. We have a commuting triangle of natural homomorphisms

$$\begin{tikzcd} & B/\mathfrak{c} \arrow[dd, "\lambda"] \\ A/\mathfrak{a} \arrow[ur, "\alpha"] \arrow[dr, "\beta"'] & \\ & B/\mathfrak{b} \end{tikzcd}$$

with α and β injective (and λ surjective). Both α and β are ws. Thus λ is injective (hence an isomorphism) by Cor. 3.16. This means that $\mathfrak{c} = \mathfrak{b}$. $\square$

The nil radical of a ring C will be denoted by $\operatorname{Nil} C$.

Example 4.7. $\operatorname{Nil} B = (\operatorname{Nil} A)B$.

Indeed, we have $(\operatorname{Nil} B) \cap A = \operatorname{Nil} A$. $\square$

Theorem 4.8. *Let $\mathfrak{p}$ be a prime ideal of A with $\mathfrak{p}B \neq B$. Then $\mathfrak{q} := \mathfrak{p}B$ is a prime ideal of B. This is the unique prime ideal of B lying over $\mathfrak{p}$. If B is given as a subextension of an extension $A \subset R$ then $\mathfrak{q} = \mathfrak{p}_{[\mathfrak{p}]} \cap B$.*

PROOF. We have $A_\mathfrak{p} = B_\mathfrak{p}$. Thus $\mathfrak{p}B_\mathfrak{p}$ is the unique maximal ideal of $B_\mathfrak{p}$. Let $\mathfrak{q}$ denote the pre-image of $\mathfrak{p}B_\mathfrak{p}$ under the localization map $B \to B_\mathfrak{p}$. From the natural commuting triangle

$$\begin{array}{ccc} A & \hookrightarrow & B \\ & \searrow \quad \swarrow & \\ & A_\mathfrak{p} = B_\mathfrak{p} & \end{array}$$

we read off that $\mathfrak{q} \cap A = \mathfrak{p}$. By Prop. 6 we have $\mathfrak{p}B = \mathfrak{q}$. Thus $\mathfrak{p}B$ is a prime ideal. Now assume that $A \subset B \subset R$. Then $B \subset A_{[\mathfrak{p}]}$ by Theorem 3.13. $\mathfrak{q}' := \mathfrak{p}_{[\mathfrak{p}]} \cap B$ is a prime ideal of B with $\mathfrak{q}' \cap A = \mathfrak{p}_{[\mathfrak{p}]} \cap A = \mathfrak{p}$. Thus $\mathfrak{q}' = \mathfrak{q}$. □

REMARK 4.9. If $\mathfrak{p}B = B$ then certainly $\mathfrak{p}B \neq \mathfrak{p}_{[\mathfrak{p}]} \cap B$.

Let $X(B/A)$ denote the image of the restriction map $\mathfrak{q} \mapsto \mathfrak{q} \cap A$ from $\mathrm{Spec}B$ to $\mathrm{Spec}A$. We endow $X(B/A)$ with the subspace topology in $\mathrm{Spec}A$. It follows from Theorem 8 that $X(B/A)$ is the set of all $\mathfrak{p} \in \mathrm{Spec}A$ with $\mathfrak{p}B \neq B$.

PROPOSITION 4.10. *The restriction map SpecB $\to$ SpecA is a homeomorphism from SpecB to $X(B/A)$. The set $X(B/A)$ is pro-constructible and dense in SpecA. It is closed under generalizations in SpecA.*

PROOF. We use the framework of spectral spaces, cf. [Ho] or e.g. [KS, Chap. III]. The restriction map $\mathrm{Spec}B \to \mathrm{Spec}A$ is spectral. Thus $X(B/A)$ is pro-constructible in $\mathrm{Spec}A$, hence is itself a spectral space. Again by Theorem 8 the restriction map $r: \mathrm{Spec}B \to X(B/A)$ is bijective. If $x, y \in \mathrm{Spec}B$ and $r(y)$ is a specialization of $r(x)$ then y is a specialization of x. Since r is spectral this implies that r is a homeomorphism.

Since A is a subring of B, the image $X(B/A)$ of the restriction map contains all minimal prime ideals of A and is dense in $\mathrm{Spec}A$. If $\mathfrak{p} \in \mathrm{Spec}A$ and $\mathfrak{p}B \neq B$, then $\mathfrak{r}B \neq B$ for the prime ideals $\mathfrak{r}$ of A contained in $\mathfrak{p}$. Thus $X(B/A)$ is closed under generalizations. {This already follows from the fact that $A \hookrightarrow B$ is flat, hence the "going down theorem" holds for prime ideals.} □

We briefly discuss relations between weakly surjective extensions and integral extensions.

PROPOSITION 4.11(cf. [G$_1$, Prop. 11]). *If a ring homomorphism $\varphi: A \to B$ is both weakly surjective and integral then φ is surjective.*

PROOF. Replacing A by $\varphi(A)$ we assume without loss of generality that $A \subset B$ and φ is the inclusion mapping. We have to prove that $A = B$.

Suppose there exists an element $x \in B \setminus A$. Then $(A{:}\,x)$ is a proper ideal of A. Since B is integral over A, this implies that $(A{:}\,x)B \neq B$. This contradicts property (4) in Theorem 3.13. Thus $A = B$. □

PROPOSITION 4.12. ([Ri, §4] for R a field, [G$_1$, Prop. 11] for $R = \mathrm{Quot}A$). *Assume that $A \subset B \subset R$ are ring extensions, and that A is weakly surjective in B. For the integral closures $\tilde{A}$ and $\tilde{B}$ of A and B in R the following holds.*

i) $\tilde{B} = \tilde{A} \cdot B$.
ii) *$\tilde{A}$ is weakly surjective in $\tilde{B}$.*

Proof. The argument in [Ri] (p.797, proof of Prop.1) extends to our more general situation.

$\square$

§5 Basic theory of relative Prüfer rings

Let R be a ring and A a subring of R.

Definition 1 [G_2, §4] *) A is called an *R-Prüfer ring*, or a *Prüfer subring of R*, if $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R for every maximal ideal $\mathfrak{p}$ of A. We then also say that A is *Prüfer in R*, or that R is a *Prüfer extension* of A.

N.B. According to Prop. 2.10 this holds iff $(A_\mathfrak{p}, \mathfrak{p}_\mathfrak{p})$ is a Manis pair in $R_\mathfrak{p}$ for every maximal ideal $\mathfrak{p}$ of A.

In particular, if R is a field, we arrive at the classical notion of a Prüfer domain.

Proposition 5.1. *Assume that A is Prüfer in R.*

i) *For every prime ideal $\mathfrak{p}$ of A the pair $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is Manis in R.*
ii) *The following are equivalent.*

(1) *A is a Manis subring of R.*
(2) *A is a valuation subring of R.*
(3) *$R \setminus A$ is multiplicatively closed, i.e. $(R \setminus A)(R \setminus A) \subset R \setminus A$.*

Moreover, if $A \neq R$ and (1) – (3) hold then $(A, \mathfrak{p}_A)$ is a Manis pair of R. $\{\mathfrak{p}_A$ had been defined in §2, Def.2.$\}$

Proof. i) Let $\mathfrak{p}$ be a prime ideal of A. We choose a maximal ideal $\mathfrak{m} \supset \mathfrak{p}$. There exists a Manis valuation v on R with $A_v = A_{[\mathfrak{m}]}$, $\mathfrak{p}_v = \mathfrak{m}_{[\mathfrak{m}]}$. If $A \cap \operatorname{supp} v \not\subset \mathfrak{p}$, then we choose some $s \in (\operatorname{supp} v) \cap (A \setminus \mathfrak{p})$. We have $sR \subset A_{[\mathfrak{m}]} \subset A_{[\mathfrak{p}]}$, and we conclude that $A_{[\mathfrak{p}]} = R$. Thus $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is certainly Manis in R in this case. Assume now that $A \cap \operatorname{supp} v \subset \mathfrak{p}$. Then it follows from Prop.2.11 that $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is Manis in R.

In assertion (ii) the implications (1) $\Rightarrow$ (2) $\Rightarrow$ (3) are trivial. We prove (3) $\Rightarrow$ (1). We may assume $A \neq R$. Let $\mathfrak{p} := \mathfrak{p}_A$. Let $x \in A_{[\mathfrak{p}]}$ be given. There exists some $d \in A \setminus \mathfrak{p}$ with $dx \in A$. If $x \notin A$ this would imply $d \in \mathfrak{p}$ by definition of $\mathfrak{p} = \mathfrak{p}_A$. Thus $x \in A$. This proves $A_{[\mathfrak{p}]} \subset A$, i.e. $A_{[\mathfrak{p}]} = A$. Then $\mathfrak{p}_{[\mathfrak{p}]} \subset A$, hence $\mathfrak{p} = \mathfrak{p}_{[\mathfrak{p}]} \cap A = \mathfrak{p}_{[\mathfrak{p}]}$. Since A is Prüfer in R we conclude that the pair $(A, \mathfrak{p})$ is Manis in R. $\square$

*) It turned out that Griffin's definition is not quite "correct". He only demands that the $A_{[\mathfrak{p}]}$ are Manis subrings of R. For a reasonable theory it is necessary to include a condition on the $\mathfrak{p}_{[\mathfrak{p}]}$, cf. also [Gr, p.285].

The following theorem gives a bunch of criteria for a given ring extension $A \subset R$ to be Prüfer. It is here that the theory of Manis valuations and the theory of weakly surjective ring extensions, displayed in §1, §2 and in §3, §4 respectively, come together.

THEOREM 5.2. *The following are equivalent.*

(1) *A is an R-Prüfer ring.*
(2) *A is weakly surjective in every R-overring.*
(2′) *A is weakly surjective in $A[x]$ for every $x \in R$.*
(3) *If B is any R-overring of A then $(A{:}\,x)B = B$ for every $x \in B$.*
(4) *Every R-overring of A is integrally closed in R.*
(5) *A is integrally closed in R, and $A[x] = A[x^n]$ for every $x \in R$ and $n \in \mathbb{N}$.*
(5′) *A is integrally closed in R, and $A[x] = A[x^2]$ for every $x \in R$.*
(6) *A is integrally closed in R. For every $x \in R$ there exists a polynomial $F[T] = \sum_{i=0}^{d} a_i T^i$ with all $a_i \in A$, $a_j = 1$ for at least one index j, such that $F(x) = 0$.*
(7) *A is integrally closed in R. For every $x \in R$ and every maximal ideal $\mathfrak{p}$ of A there exists a polynomial $F_{x,\mathfrak{p}}(T) \in A[T] \setminus \mathfrak{p}[T]$ such that $F_{x,\mathfrak{p}}(x) = 0$.*
(8) *$(A{:}\,x) + x(A{:}\,x) = A$ for every $x \in R$.*
(9) *A is integrally closed in R. For every overring B of R the restriction map $SpecB \to SpecA$ is injective.*
(9′) *A is integrally closed in R. If B is an R-overring of A and $\mathfrak{q} \subset \mathfrak{q}'$ are prime ideals of B with $\mathfrak{q} \cap A = \mathfrak{q}' \cap A$ then $\mathfrak{q} = \mathfrak{q}'$.*
(10) *A is integrally closed in R. For every prime ideal $\mathfrak{p}$ of A there exists a unique Manis pair $(B, \mathfrak{q})$ in R over $(A, \mathfrak{p})$, i.e. with $A \subset B$, $\mathfrak{q} \cap A = \mathfrak{p}$.*
(11) *For every R-overring B of A the inclusion map $A \hookrightarrow B$ is an epimorphism (in the category of rings).*
(11′) *For every $x \in R$ the inclusion map $A \hookrightarrow A[x]$ is an epimorphism.* □

REMARKS. The equivalence of (1), (2), (3), (4) had already been stated by Griffin [G_2, Prop.6, Th.7], but he made additional assumptions and did not present the proofs. On the other hand, Griffin weakened the hypothesis that our rings have unit elements. The equivalence of (1), (4), (8) has been proved by Eggert for $R = Q(A)$, the complete ring of quotients of A [Eg, Th.2]. The equivalence of (1) and any of the conditions (4) - (7) is a generalization of classical results for R a field (cf. e.g. [E, Th.11.10]). The equivalence of (1) and (11) for R a field has been proved by Storrer [St_1]. The equivalence of (1), (2), (4), (8) has been stated in full generality by Rhodes [Rh, Th.2.1]. Unfortunately his proof contains a gap (cf. Introduction to the present paper). E.D. Davis studied extensions $A \subset R$ with property (4) under the name "normal pairs". In the case of domains some of our results in this section can be read off from his paper [Da].

PROOF. (1) ⇒ (2): Let B be an R-overring of A and $\mathfrak{q}$ a prime ideal of B. Let $\mathfrak{p} := \mathfrak{q} \cap A$. We verify that $A_{[\mathfrak{p}]} = B_{[\mathfrak{q}]}$ and then will be done by Theorem 3.13. Of course, $A_{[\mathfrak{p}]} \subset B_{[\mathfrak{q}]}$. Let $x \in R \setminus A_{[\mathfrak{p}]}$ be given. We prove that $x \notin B_{[\mathfrak{q}]}$, and then will be done.

Since $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R there exists an element y of $\mathfrak{p}_{[\mathfrak{p}]}$ with $xy \in A_{[\mathfrak{p}]} \setminus \mathfrak{p}_{[\mathfrak{p}]}$. We choose elements a and c in $A \setminus \mathfrak{p}$ with $a(xy) \in A$ and $cy \in \mathfrak{p}$. We have $a(xy) \in A \setminus \mathfrak{p}$. Suppose that $x \in B_{[\mathfrak{q}]}$. Then there exists some $b \in B \setminus \mathfrak{q}$ with $bx \in B$. We have $a(bx)(cy) \in \mathfrak{q}$. On the other hand, $a(bx)(cy) = bc(axy) \in B \setminus \mathfrak{q}$. This contradiction proves that $x \notin B_{[\mathfrak{q}]}$.

(2) $\Rightarrow$ (2′): trivial.

(2) $\Leftrightarrow$ (3): Clear from Th. 3.13.

(2′) $\Rightarrow$ (3): Let $x \in B$. Then $(A{:}\,x)A[x] = A[x]$. A fortiori $(A{:}\,x)B = B$.

(2) $\Rightarrow$ (4): Let B be an R-overring of A, and let $C = \tilde{B}$ denote the integral closure of B in R. By (2) A is ws in C. Thus B is ws in C (Prop. 3.7.b). Prop. 4.11 tells us that $C = B$, i.e. B is integrally closed in R.

(4) $\Rightarrow$ (5): x is integral over $A[x^n]$. By assumption (4) the subring $A[x^n]$ is integrally closed in R. Thus $x \in A[x^n]$.

(5) $\Rightarrow$ (5′): trivial.

(5′) $\Rightarrow$ (6): For every $x \in R$ we have a relation $x = \sum\limits_{i=0}^{m} a_i x^{2i}$ with $m \in \mathbb{N}_0$, $a_i \in A$.

(6) $\Rightarrow$ (7): trivial.

(7) $\Rightarrow$ (1): Theorem 2.12 tells us that $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is a Manis pair in R for every $\mathfrak{p} \in \operatorname{Spec} A$.

(1) $\Rightarrow$ (8): Suppose there exists some $x \in R$ with $I{:}= (A{:}\,x) + x(A{:}\,x) \neq A$. We choose a maximal ideal $\mathfrak{m}$ of A containing I. Then $x \in R \setminus A_{[\mathfrak{m}]}$ since $(A{:}\,x) \subset \mathfrak{m}$. By (1) and Theorem 2.4 (iii) there exists some $x' \in \mathfrak{m}_{[\mathfrak{m}]}$ with $xx' \in A_{[\mathfrak{m}]} \setminus \mathfrak{m}_{[\mathfrak{m}]}$. We then choose some $d \in A \setminus \mathfrak{m}$ with $dx' \in \mathfrak{m}$ and $dxx' \in A \setminus \mathfrak{m}$. It follows that $dx' \in (A{:}\,x)$ and $dxx' \in x(A{:}\,x) \subset \mathfrak{m}$, a contradiction. Thus (8) holds.

(8) $\Rightarrow$ (1): We prove for a given prime ideal $\mathfrak{p}$ of A that the pair $(A_{[\mathfrak{p}]}, \mathfrak{p}_{[\mathfrak{p}]})$ is Manis in R by verifying condition (iii) in Theorem 2.4. Let $x \in R \setminus A_{[\mathfrak{p}]}$. Then $(A{:}\,x) \subset \mathfrak{p}$. By (8) we know that $x(A{:}\,x) \not\subset \mathfrak{p}$. Thus there exists some $x' \in (A{:}\,x) \subset \mathfrak{p}$ with $xx' \in A \setminus \mathfrak{p}$.

The equivalence of (1), (9), (9′), (10) is evident from Theorem 2.14. The implication (2′) $\Rightarrow$ (11′) follows from the fact that every weakly surjective map is an epimorphism (cf. Prop.3.6).

(11′) $\Rightarrow$ (11): Suppose there exists an R-overring B of A such that the inclusion map $A \hookrightarrow B$ is not an epimorphism. Then there exist two ring homomorphisms φ_1, φ_2 from B to some ring C with $\varphi_1|A = \varphi_2|A$ but $\varphi_1 \neq \varphi_2$. We choose some $x \in B$ with $\varphi_1(x) \neq \varphi_2(x)$. The restrictions $\varphi_1|A[x]$ and $\varphi_2|A[x]$ are different, but $\varphi_1|A = \varphi_2|A$. This contradicts the assumption (11′).

(11) $\Rightarrow$ (4): Let B be an R-overring of A, and let $x \in R$ be integral over B. We want to prove that $x \in B$. The inclusion $A \hookrightarrow B[x]$ is an epimorphism. Thus (for purely categorial reasons) also the inclusion $B \hookrightarrow B[x]$ is an epimorphism. By an easy proposition of Lazard [L, Chap. IV, Prop.1.7], $B[x] = B$. $\square$

From condition (4) in this theorem one obtains immediately

Corollary 5.3. *Let B be an R-overring of A. If A is Prüfer in R then B is Prüfer in R and A is Prüfer in B.* $\square$

From condition (8) in the theorem we obtain

COROLLARY 5.4. *If A is Prüfer in R then for any $x \in R$ the ideal $(A\colon x)$ is generated by two elements.*

Indeed, we have elements a and b in $(A\colon x)$ with $1 = a + xb$. If $u \in (A\colon x)$ then $u = ua + (ux)b$. Thus $(A\colon x) = Aa + Ab$. □

Theorem 2 contains the fact that every R-Prüfer ring is integrally closed in R. The reader might ask for a more direct proof of this statement. Indeed this follows from the definition of R-Prüfer rings and an elementary fact which holds without any assumption about our subring A of R.

REMARK 5.5. If M is an A-submodule of R, then

$$M = \bigcap_{\mathfrak{p} \in \Omega} A_{[\mathfrak{p}]} \cdot M = \bigcap_{\mathfrak{p} \in \Omega} M_{[\mathfrak{p}]},$$

with Ω denoting the set of maximal ideals of A. In particular $A = \bigcap\limits_{\mathfrak{p} \in \Omega} A_{[\mathfrak{p}]}$.

PROOF. Of course, $M \subset A_{[\mathfrak{p}]}M \subset M_{[\mathfrak{p}]}$ for every $\mathfrak{p} \in \Omega$. Let $x \in \bigcap\limits_{\mathfrak{p} \in \Omega} M_{[\mathfrak{p}]}$ be given. Consider the ideal $\mathfrak{a}\colon= \{a \in A \mid ax \in M\}$. For every $\mathfrak{p} \in \Omega$ the intersection $\mathfrak{a} \cap (A \setminus \mathfrak{p})$ is not empty, i.e. $\mathfrak{a} \not\subset \mathfrak{p}$. Thus $\mathfrak{a} = A$, i.e. $x \in M$. □

We now look for permanence properties of relative Prüfer rings.

THEOREM 5.6 [Rh, Prop.3.1.3]. *Assume that A is a Prüfer subring of B and B is a Prüfer subring of C. Then A is Prüfer in C.*

PROOF (cf. [Rh, loc.cit]). We verify for a given prime ideal $\mathfrak{p}$ of A that the pair $(A_\mathfrak{p}, \mathfrak{p}_\mathfrak{p})$ is Manis in $C_\mathfrak{p}$. Replacing A, B, C by $A_\mathfrak{p}$, $B_\mathfrak{p}$, $C_\mathfrak{p}$ we assume without loss of generality that A is local and $\mathfrak{p}$ is the maximal ideal of A. We will apply Theorem 2.5. By this theorem (or Prop.1.3) B is local, and the maximal ideal $\mathfrak{q}$ of B is contained in $\mathfrak{p}$. Let $x \in C \setminus A$ be given. If $x \in B$ then, by Theorem 2.5, $x \in B^*$ and $x^{-1} \in \mathfrak{p}$. If $x \notin B$ then, by the same theorem, $x \in C^*$ and $x^{-1} \in \mathfrak{q} \subset \mathfrak{p}$. Thus in both cases x is a unit in C and $x^{-1} \in A$. We conclude, again by Theorem 2.5, that $(A, \mathfrak{p})$ is Manis in C. □

PROPOSITION 5.7. *Assume that A is a Prüfer subring of B. Then, for any ring homomorphism $\psi\colon B \to D$ the ring $\psi(A)$ is Prüfer in $\psi(B)$.*

PROOF. Let C' be a subring of $\psi(B)$ containing $\psi(A)$. We verify that $\psi(A)$ is weakly surjective in C', and then will be done by condition (2) in Theorem 2. Indeed, $C\colon= \psi^{-1}(C')$ is a subring of B containing A. Thus A is weakly surjective in C. By Proposition 3.8 $\psi(A)$ is weakly surjective in $\psi(C) = C'$. □

PROPOSITION 5.8 [Rh, Prop.3.1.1]. *Let $A \subset R$ be a ring extension and I an ideal of R contained in A. Then A is Prüfer in R iff A/I is Prüfer in R/I.*

PROOF. If A is Prüfer in R then the preceding proposition tells us that A/I is Prüfer in R/I. Assume now that the latter holds. We verify condition (4) in Theorem 2 and then will be done.

Let B be an R-overring of A. Then B/I is an R/I-overring of A/I. Thus B/I is integrally closed in R/I. Let $x \in R$ be integral over B. Then $x + I \in B/I$. Since $I \subset B$ we conclude that $x \in B$. Thus B is integrally closed in R. $\square$

THEOREM 5.9. *Let $\varphi: R \to R'$ be an integral ring homomorphism. Let A be a Prüfer subring of R, and let A' denote the integral closure of $\varphi(A)$ in R'. Then A' is a Prüfer subring of R'.*

PROOF. We verify condition (7) in Theorem 2. Let an element x of R' and a prime ideal $\mathfrak{q}$ of R' be given. Let $\mathfrak{p} := \varphi^{-1}(\mathfrak{q})$. We look for a polynomial $G(T) \in A[T] \setminus \mathfrak{p}[T]$ with $G^{\varphi}(x) = 0$, where $G^{\varphi}(T)$ denotes the polynomial obtained from $G(T)$ by applying φ to the coefficients.

We start with a polynomial

$$F(T) = T^n + a_1 T^{n-1} + \cdots + a_n \in R[T]$$

such that $F^{\varphi}(x) = 0$. Such a polynomial exists since φ is integral. Let $v: R \twoheadrightarrow \Gamma \cup \infty$ denote the Manis valuation on R with $A_v = A_{[\mathfrak{p}]}$, $\mathfrak{p}_v = \mathfrak{p}_{[\mathfrak{p}]}$. We choose an index $r \in \{1, \dots, n\}$ with $v(a_r) = \mathrm{Min}\{v(a_i) \mid 1 \le i \le n\}$. We distinguish two cases.

CASE 1: $v(a_r) = \infty$. Now certainly $a_i \in A_{[\mathfrak{p}]}$ for $i = 1, 2, \dots, n$. We choose some $d \in A \setminus \mathfrak{p}$ with $da_i \in A$ for all i. The polynomial $G(T) := dF(T)$ does the job.

CASE 2: $v(a_r) < \infty$. We choose some $b \in R$ with $v(ba_r) = 0$. This is possible since v is Manis. We have

$$ba_i \in A_{[\mathfrak{p}]} \quad \text{for every} \quad i \in \{1, \dots, n\}$$

and $ba_r \notin \mathfrak{p}_{[\mathfrak{p}]}$. We choose some $c \in A \setminus \mathfrak{p}$ with $cba_i \in A$ for $i = 1, \dots, n$. The polynomial $G(T) := cbF(T)$ does the job. $\square$

REMARK. Since $\varphi(A)$ is weakly surjective in $\varphi(R)$ we conclude from Prop.4.12 that $R' = A' \cdot \varphi(R)$.

THEOREM 5.10. *Let A be a subring of R and B, C be two R-overrings of A. Assume that A is Prüfer in B and weakly surjective in C. Then C is Prüfer in BC.*

PROOF. We pick a prime ideal $\mathfrak{q}$ of C and verify that $(C_{\mathfrak{q}}, \mathfrak{q}_{\mathfrak{q}})$ is a Manis pair in $(BC)_{\mathfrak{q}}$.

Let $\mathfrak{p} := \mathfrak{q} \cap A$. Then $A_{\mathfrak{p}} = C_{\mathfrak{p}} = C_{\mathfrak{q}}$ and $\mathfrak{q}_{\mathfrak{q}} = \mathfrak{q}_{\mathfrak{p}} = \mathfrak{p}_{\mathfrak{p}}$ (cf. Lemma 3.2). Thus $C \setminus \mathfrak{q}$ is the saturum of the multiplicative set $A \setminus \mathfrak{p}$ in C. Notice also that $BC = B \otimes_A C$ (Prop. 4.2). Thus $(BC)_{\mathfrak{p}} = B_{\mathfrak{p}} \otimes_{A_{\mathfrak{p}}} C_{\mathfrak{p}} = B_{\mathfrak{p}}$, more precisely, the subrings $(BC)_{\mathfrak{p}}$ and $B_{\mathfrak{p}}$ of $R_{\mathfrak{p}}$ are equal. We conclude that $B_{\mathfrak{q}} = B_{\mathfrak{p}}$, $(BC)_{\mathfrak{q}} = (BC)_{\mathfrak{p}} = B_{\mathfrak{p}}$, $(C_{\mathfrak{q}}, \mathfrak{q}_{\mathfrak{q}}) = (A_{\mathfrak{p}}, \mathfrak{p}_{\mathfrak{p}})$. Since A is Prüfer in B, the pair $(C_{\mathfrak{q}}, \mathfrak{q}_{\mathfrak{q}})$ is Manis in $(BC)_{\mathfrak{q}}$. $\square$

Corollary 5.11. *Let A be a subring of R and B, C be two R-overrings of A. If A is Prüfer in B and in C, then A is Prüfer in BC.*

This follows from theorems 10 and 6. □

Counterexample 5.12. If $A \subset B$ is a Prüfer extension and $A \subset C$ is a flat ring extension then C is not necessarily Prüfer in $B \otimes_A C$. Here is a simple example: Let A be a non trivial valuation ring of a field K. Then A is Prüfer in K, but the polynomial ring $A[T]$ in one variable T is not Prüfer in $K[T]$.

Indeed, let $\mathfrak{m}$ be the maximal ideal of A and let $M := \mathfrak{m} + TA[T]$, which is a maximal ideal of $C := A[T]$. In the extension $K[T]$ of C we have $C_{[M]} = C$, $M_{[M]} = M$, as is easily verified. The pair (C, M) is not Manis in $K[T]$. □

Remark 5.13. Let $A \subset R$ be a ring extension and $(B_i | i \in I)$ an upward directed family of R-overrings of A. Assume that A is Prüfer in every B_i. Then A is Prüfer in $B := \bigcup_{i \in I} B_i$.

Proof. Let C be an R-overring of A contained in B. We verify that A is weakly surjective in C and then will be done by Theorem 2. Now C is the union of the upward directed family of subrings $(C \cap B_i | i \in I)$. A is weakly surjective in $C \cap B_i$ for every $i \in I$. Thus A is weakly surjective in C (Remark 3.12). q.e.d.

We now have the means to establish a theory of "Prüfer hulls" analogous to the theory of weakly surjective hulls in §3.

Theorem 5.14. *Let $A \subset R$ be a ring extension. Then there exists a unique R-overring $\mathrm{Pr}(A, R)$ of A such that A is Prüfer in $\mathrm{Pr}(A, R)$, and $\mathrm{Pr}(A, R)$ contains every R-overring of A in which A is Prüfer.*

This follows from Corollary 11 and Remark 13 (cf. the proof of Prop. 3.11). □

Definition 2. We call $\mathrm{Pr}(A, R)$ the *Prüfer hull of A in R.*

Of course, $\mathrm{Pr}(A, R)$ is contained in the weakly surjective hull $M(A, R)$ of A in R, and $\mathrm{Pr}(A, R) = \mathrm{Pr}(A, C)$ for every R-overring C with $C \supset \mathrm{Pr}(A, R)$. Also $\mathrm{Pr}(A, R) = \mathrm{Pr}(B, R)$, if B is any R-overring of A contained in $\mathrm{Pr}(A, R)$.

Definition 3. For any ring A the *Prüfer hull $P(A)$ of A* is defined as the Prüfer hull of A in the complete quotient ring $Q(A)$ (cf. §3), $P(A) := \mathrm{Pr}(A, Q(A))$.

Of course, $P(A)$ is contained in the weakly surjective hull $M(A)$. The classical Prüfer rings (with zero divisors) are precisely the rings A with $\mathrm{Quot} A \subset P(A)$. If A' is a weakly surjective ring extension of A, contained in $M(A)$ without loss of generality, then
$A' \cdot P(A) \subset P(A')$ by Theorem 10 above.

Example 5.15. Let V be an affine algebraic variety over some real closed field k. The ring R of (k-valued, continuous) semialgebraic functions on $V(k)$ is "Prüfer closed", i.e. $P(R) = R$. This has been proved recently by Niels Schwartz [Sch_2] within the framework of his theory of real closed rings. His proof would take us here too far afield.

Let d be a natural number. In §6 we will see that R is Prüfer over the subring $A = k\big[\frac{1}{1+x^{2d}} | x \in R\big]$ generated by k and the elements $\frac{1}{1+x^{2d}}$, $x \in R$, cf. Th.6.8. Thus $R = P(A)$.

$\square$

§6 Examples of convenient ring extensions and relative Prüfer rings

In this section R is a ring and A is a subring of R. We are looking for handy criteria which guarantee that A is Manis or Prüfer in R, and we will discuss examples emanating from some of these criteria.

Theorem 6.1. *Assume that A is integrally closed in R. Assume further that for every $x \in R \setminus A$ there exists a monic polynomial $F(T) \in A[T]$ and a unimodular polynomial $G(T) \in A[T]$ (i.e. the ideal of A generated by all coefficients of $G(T)$ is A), such that $F(x) \in R^*$, $\deg G < \deg F$ and $G(x)/F(x) \in A$. Then A is Prüfer in R.*

Proof. We verify that for a given element x of R and a given maximal ideal $\mathfrak{m}$ of A there exists a polynomial $H(T) \in A[T] \setminus \mathfrak{m}[T]$ with $H(x) = 0$, and then will be done by Theorem 5.2.

If $x \in A$ we take $H(T) = T - x$. Now let $x \in R \setminus A$. We choose polynomials $F(T)$, $G(T)$ as indicated in the theorem. We put $b := G(x)/F(x) \in A$ and take $H(T) := bF(T) - G(T)$. Then $H(x) = 0$. If $b \in \mathfrak{m}$ then $H(T) \notin \mathfrak{m}[T]$, since $G(T)$ is unimodular. If $b \notin \mathfrak{m}$ then again $H(T) \notin \mathfrak{m}[T]$, since $\deg G < \deg F$. q.e.d.

Definition 1. We call a valuation v on R a *Prüfer-Manis valuation* (or *PM-valuation*, for short), if v is Manis and A_v is Prüfer in R. We call a subring B of R a *Prüfer-Manis subring* of R if $B = A_v$ for some Prüfer-Manis valuation v on R. We then also say that the ring B is *Prüfer-Manis* (or *PM*, for short) *in R*.

If A is Prüfer in R and B is an R-overring of A which is Manis in R, then, of course, B is PM in R. Thus the valuations which really matter in the theory of relative Prüfer rings are the PM-valuations and not just the Manis valuations. We defer a systematic theory of PM-subrings of R to later chapters, but now look for examples of such rings.

Theorem 6.2. *Assume that $A \neq R$ and the set $S := R \setminus A$ is multiplicatively closed. Assume further that for every $x \in R \setminus A$ there exists a monic polynomial $F(T) \in A[T]$ of degree ≥ 1 with $F(x) \in R^*$. Then A is PM in R.*

Proof. We verify that A is Prüfer in R and then will be done by Prop. 5.1.ii. We know from Theorem 2.1 that A is integrally closed in R. Let $x \in R \setminus A$ be given. We choose a polynomial $F(T) \in A[T]$ as indicated in the theorem. Certainly $F(x) \in R \setminus A$, since A is integrally closed in R. We conclude from the equation $1 = F(x) \cdot F(x)^{-1}$ that $1/F(x) \in A$, since otherwise we would get the contradiction $1 \in R \setminus A$. Now Theorem 1 tells us that A is Prüfer. $\square$

Definition 2. a) Let k be a subring of R. We say that R is *convenient over* k, if every R-overring A of k which has a multiplicatively closed complement $R \setminus A$ is PM in R.
b) We call the *ring* R *convenient*, if R is convenient over its prime ring $\mathbb{Z} \cdot 1$.

Example 1. Every field is a convenient ring. $\square$

The idea behind Definition 2 is that, as far as valuations are concerned, a convenient ring is nearly as "convenient" as a field. If R is only convenient over some subring k then at least this should be true for the (special) valuations v with $A_v \supset k$. In particular we expect that for a convenient ring extension $k \subset R$ we have a theory of R-Prüfer rings $A \supset k$ nearly as good as in the field case.

From Theorem 2 we extract

Scholium 6.3. Let k be a subring of R with the following property.
($*$) For every $x \in R \setminus k$ there exists some monic polynomial $F_x(T) \in k[T]$, $F_x \neq 1$, with $F_x(x) \in R^*$.
Then R is convenient over k. $\square$

We give some examples of ring extensions which are convenient and, up to the first and the last one, even fulfill condition ($*$).

Example 2 (Generalization of Example 1). If R has Krull dimension zero then R is convenient.

Proof. Let A be a subring of R with $A \neq R$ and $R \setminus A$ multiplicatively closed. We prove that A is Prüfer in R. Then it will follow from Prop. 5.1.ii that A is also Manis in R.

The ring A is integrally closed in R by Theorem 2.1.ii. Given an element $x \in R$ we prove that there exists a unimodular polynomial $F(T) \in A[T]$ with $F(x) = 0$. Then we will be done by Theorem 5.2.

If $x \in A$ take $F(T) = T - x$. Now let $x \in R \setminus A$. There exists some $n \in \mathbb{N}$ and $y \in R$ with $x^{n+1}y = x^n$, cf. [Huc, Th.3.5]. Then $(xy)^{n+1} = (xy)^n$. Since A is integrally closed in R, this implies $xy \in A$. Since $R \setminus A$ is closed under multiplication we conclude that $y \in A$. The polynomial $F(T) = yT^{n+1} - T^n$ fits our needs. $\square$

Example 3. Every ring R with $1 + \Sigma R^2 \subset R^*$ is convenient. Indeed, it suffices to know that $1 + R^2 \subset R^*$ in order to conclude that R is convenient. $\square$

Comment. This is the most important class of rings we have in mind for use in real algebra. Recall that for every ring A the localization $\Sigma^{-1}A$ with respect to the multiplicative set $\Sigma := 1 + \Sigma A^2$ is such a ring, and that A and $\Sigma^{-1}A$ have the same real spectrum. For many problems in real algebra we may replace A by $\Sigma^{-1}A$ and thus arrive at a convenient ring. {If A is not real, i.e. $-1 \in \Sigma A^2$, then $\Sigma^{-1}A$ is the null ring, but this does not bother us.}

Subexample 3 bis. If A is any ring and X is a pro-constructible subset of the real spectrum Sper$A$ then the ring $C(X, A)$ of abstract semialgebraic functions on X (cf. [Sch]) is convenient, since in this ring R we have $1 + \Sigma R^2 \subset R^*$. In general $C(X, A)$ has very many zero divisors.

Example 4. If more generally R is a ring such that, for every $x \in R$, there exists a natural number d with $1 + x^d \in R^*$, then R is convenient.

Such rings (with d even) seem to be important in the theory of orderings of higher level and higher real spectra (cf. e.g. [B_2], [B_3], [P], [BP], [Be]).

Example 5. Let A be an affine algebra over a field k which is not algebraically closed. Let $V(k)$ denote the set of rational points of the associated k-variety V. {We may identify $V(k) = \mathrm{Hom}_k(A, k)$.} Let U be a k-Zariski-open subset of $V(k)$. {In other words, U is open in the subspace topology of $V(k)$ in Spec$A$.} Let finally S be the multiplicative set consisting of all $a \in A$ with $a(p) \neq 0$ for every $p \in U$. Then $S^{-1}A$ is convenient over k.

Proof. We choose a monic polynomial $F(T) \in k[T]$, $F \neq 1$, in one variable T which has no zeros in k. Let $x \in S^{-1}A$ be given. Write $x = \frac{a}{s}$ with $a \in A$, $s \in S$, and $F(T) = T^d + c_1T^{d-1} + \cdots + c_d$. We have $F(x) = \frac{b}{s^d}$ with $b = a^d + c_1a^{d-1}s + \cdots + c_ds^d$. For every point $p \in U$ we have $\frac{b(p)}{s(p)^d} = F\big(\frac{a(p)}{s(p)}\big) \neq 0$, hence $b(p) \neq 0$. Thus $b \in S$ and $F(x)$ is a unit in $S^{-1}A$. $\square$

Definition 3. We call this ring $S^{-1}A$ the *ring of regular functions on U*.

If the field k is real closed and $U = V(k)$ then S is the set of divisors of the elements in $\Sigma := 1 + \Sigma A^2$, as is well known (e.g. [BCR, Cor. 4.4.5.], [KS, p.142]). Thus $S^{-1}A = \Sigma^{-1}A$, and we are back to Example 3.

Example 6. If R is a semi-local ring containing an infinite field k then R is convenient over k. Indeed, if R has the maximal ideals $\mathfrak{m}_1, \dots, \mathfrak{m}_r$, then for a given $x \in R$ we find some $\lambda \in k$ with $x - \lambda \notin \mathfrak{m}_i$ for $i = 1, \dots, r$, hence $x - \lambda \in R^*$. $\square$

If k is any ring and $(R_\alpha | \alpha \in I)$ is a direct system of k-algebras fulfilling condition $(*)$ from above then the same holds for the inductive limit $\varinjlim R_\alpha$. Thus Example 6 can be amplified to

Example 7. An inductive limit of semi-local k-algebras over some infinite field k is convenient over k. In particular, if R is a semi-local algebra over some infinite field k, the infinite Galois extensions of R (cf. e.g. [K], there called "coverings") are convenient over k. $\square$

Example 8. Assume that R is the total quotient ring of A, $R = \mathrm{Quot}A$. The ring A is called *additively regular* [Huc, p.32], if for every $x \in R$ there exists some $a \in A$

such that $x + a$ is a "regular element", i.e. a unit in R. Of course then condition $(*)$ is satisfied for $k := A$, and thus R is convenient over A. As Huckaba observes [Huc, p.32 f], if A is Noetherian or, more generally, if the set of zero divisors of A is a union of finitely many prime ideals, then A is additively regular [Huc, p.32 f]. □

Example 9. Assume again that $R = \mathrm{Quot}A$. The ring A is called a *Marot ring* [Huc, p.31], if each ideal of A which contains a non zero divisor is generated by a set of non zero divisors. Marot rings form a very broad class of rings. In particular, every additively regular ring is Marot [Huc, p.33 f]. If A is Marot then $R = \mathrm{Quot}A$ is convenient over A, cf. [Huc, Th.7.7 and Cor.7.8]. But now condition $(*)$ may be violated, as we can show by examples. □

As before R denotes a ring and A a subring of R. We return to the search for Prüfer subrings of R which are not necessarily Manis in R.

If R is a field then the intersection of finitely many valuation subrings of R is Prüfer in R, as is well known. Does the same hold if $k \subset R$ is a convenient extension and if all the valuation rings contain k? Or does this at least hold if the extension $k \subset R$ fulfills the stronger condition $(*)$ in 6.3? We can only prove the following result.

Theorem 6.4. *Let k be a subring of R with the following property.*
$(**)$ *For every $x \in R \setminus k$ there exists a monic polynomial $F_x(T) \in k[T]$, $F_x \neq 1$, with $F_x(x) \in R^*$ and constant term $F_x(0) \in k^*$.*

Let $v_1, \dots, v_n$ be valuations on R with $A_{v_i} \supset k$ for all i. Then the intersection A of the rings A_{v_i} is Prüfer in R.

Proof. A is integrally closed in R. Let $x \in R \setminus A$ be given. We prove that there exists a monic polynomial $H(T) \in k[T]$ of degree ≥ 1 with $H(x) \in R^*$ and $1/H(x) \in A$, and then will be done by Theorem 1.

For every index i with $1 \leq i \leq n$ we choose a monic polynomial $F_i(T) \in k[T]$ with $v_i(F_i(x)) > 0$, if such a polynomial exists. Otherwise we put $F_i(T) := 1$.
Let $G(T) := TF_1(T) \cdots F_n(T)$ and $y := G(x)$. Certainly $y \notin A$, since $x \notin A$ and A is integrally closed in R. A fortiori $y \notin k$. We claim that the polynomial $H(T) := F_y(G(T))$ fits our needs (with F_y as indicated in the theorem).

Certainly $H(T)$ is monic and $H(x) = F_y(y) \in R^*$. Given $i \in \{1, \dots, n\}$ we verify that $v_i(1/H(x)) \geq 0$, and then will be done.

Case 1. $v_i(x) < 0$. Now $v_i(H(x)) = (\deg H) \cdot v_i(x) < 0$, since $H(T)$ is monic and has coefficients in A_{v_i}. Thus $v_i(1/H(x)) > 0$.

Case 2. $v_i(x) \geq 0$. Then $v_i(H(x)) \geq 0$. Suppose that $v_i(H(x)) > 0$. Then $v_i(F_i(x)) > 0$, hence $v_i(y) > 0$. But $F_x(T) = T^d + c_1T^{d-1} + \cdots + c_d$ has constant term $c_d \in k^*$. Thus $H(x) = y^d + c_1y^{d-1} + \cdots + c_d$ has value $v_i(H(x)) = 0$. This is a contradiction. We conclude that $v_i(H(x)) = 0$, hence $v_i(1/H(x)) = 0$. □

Notice that, for k a subfield of a ring R, the previous condition $(*)$ (cf. 6.3) implies $(**)$. In particular $(**)$ holds in the examples 5 – 7 above. $(**)$ holds also in the examples 1, 3, 4 for k the prime ring in R.

Definition 4. Let $F(T) \in R[T]$ be a non-constant monic polynomial. Let v be a valuation on R. We call v an *F-valuation*, if $v(c) \geq 0$ for every coefficient c of F and $F(T)$ has no zero in the residue class field $\kappa(v) = \mathfrak{o}_v/\mathfrak{m}_v$. {Of course, this means that the image polynomial $\bar{F}(T) \in \kappa(v)[T]$ has no zero in $\kappa(v)$.}

Theorem 6.5. *Let $(v_i|i \in I)$ be a family of valuations on R. Assume that A is the intersection of the valuation rings A_{v_i} $(i \in I)$. Assume also that for each $x \in R \setminus A$ there exists a monic polynomial $F_x(T) \in A[T]$ of degree $d_x \geq 1$, such that $F_x(x) \in R^*$ and every v_i is an F_x-valuation. Then A is Prüfer in R.*

Proof. Each A_{v_i} is integrally closed in R. Thus A is integrally closed in R. By Theorem 1 we are done if we verify that $1/F_x(x) \in A$ for each $x \in R \setminus A$, i.e. $v_i(F_x(x)) \leq 0$ for each $x \in R\setminus A$ and $i \in I$. If $v_i(x) < 0$ then $v_i(F_x(x)) = d_x \cdot v_i(x) < 0$. If $v(x_i) \geq 0$ then $x \in A_{v_i}$, and $v_i(F_x(x)) = 0$ since v_i is an F_x-valuation. $\square$

Here we quote the seminal paper [R] by Peter Roquette, which in the case, that R is a field, bears close connection to Theorem 5. Roquette also obtained results on class groups which allow to conclude in important cases that A has trivial class group, hence is a Bezout ring. Our Theorem 5 generalizes the first part of [R, Theorem 1]. The second part, dealing with the class group of A, will be generalized in §7.

We now aim to criteria that A is Prüfer in R, which do not assume in advance that A is integrally closed in R. A prototype of the criteria to follow is a lemma of A. Dress, which states for R a field of characteristic not 2, that the subring of R generated by the elements $1/(1 + a^2)$ with $a \in F$, $a^2 \neq -1$, is Prüfer in R, cf. [D, Satz 2′], [KS, Chap III §12], [La, p.86].*)

Theorem 6.6. *Assume that for every $x \in R \setminus A$ there exists some monic polynomial $F(T) \in A[T]$ of degree ≥ 1 with $F(x) \in R^*$, $\frac{1}{F(x)} \in A$, $\frac{x}{F(x)} \in A$. Then A is Prüfer in R.*

Proof. Let B be an R-overring of A and $S := A \cap B^*$. We verify that $B = S^{-1}A$. Then we know that A is weakly surjective in every R-overring, and will be done by Theorem 5.2.
Of course, $S^{-1}A \subset B$. Let $x \in B \setminus A$ be given. We choose a polynomial $F(T)$ as indicated in the theorem. $s := \frac{1}{F(x)} \in A \subset B$. Also $F(x) \in B$, hence $s \in S$. By assumption $a := \frac{x}{F(x)} \in A$. Thus $x = \frac{a}{s} \in S^{-1}A$. $\square$

The following remark sheds additional light both on Theorem 6 and Theorem 1.

Remark 6.7. Assume that A is integrally closed in R (e.g. A is Prüfer in R). Let $x \in R$ and let $F(T) \in A[T]$ be a monic polynomial of degree $n \geq 1$ with $F(x) \in R^*$ and $\frac{1}{F(x)} \in A$. Then $\frac{x^r}{F(x)} \in A$ for $0 \leq r \leq n$.

Proof (cf. [Gi, p.154]). We proceed by induction on r. For $r = 0$ the assertion is trivial. Assume that $1 \leq r \leq n$ and that $\frac{x^s}{F(x)} \in A$ for $0 \leq s < r$.

---

*) Actually Dress made the slightly stronger assumption that -1 is not a square in F.

We write

$$F(T)^r = T^{nr} + \sum_{j=1}^{n} h_j(T) T^{(n-j)r}$$

with polynomials $h_j(T) \in A[T]$ of degree $< r$. The relation

$$\frac{1}{F(x)^{n-r}} = \frac{F(x)^r}{F(x)^n} = \big(\frac{x^r}{F(x)}\big)^n + \sum_{j=1}^{n} \frac{1}{F(x)^{j-1}} \frac{h_j(x)}{F(x)} \cdot \big(\frac{x^r}{F(x)}\big)^{n-j}$$

proves that $\frac{x^r}{F(x)}$ is integral over A, since by induction hypothesis $\frac{h_j(x)}{F(x)} \in A$ for every $j \in \{1, \ldots, n\}$. Thus $\frac{x^r}{F(x)} \in A$. □

Theorem 6.8. *Let k be a subring of R. (We will often take for k the prime ring in R.) Let $F(T) \in k[T]$ be a monic polynomial of degree $d \geq 1$. Assume that $d! \in R^*$ and that $F(x) \in R^*$ for every $x \in R$ with $F(x) \notin k$. The subring A of R generated by k, the element $1/d!$ and the set $\{1/F(x) | x \in R,\ F(x) \notin k\}$ is Prüfer in R.*

Proof. a) Let $B := \tilde{A}$, the integral closure of A in R. By Theorem 1 B is Prüfer in R. We now verify that for a given prime ideal $\mathfrak{q}$ of B and $\mathfrak{p} := \mathfrak{q} \cap A$ we have $A_{[\mathfrak{p}]} = B_{[\mathfrak{q}]}$. Since over every prime ideal $\mathfrak{p}$ of A there lies a prime ideal $\mathfrak{q}$ of B we then may conclude (Remark 5.5) that

$$B = \bigcap_{\mathfrak{q} \in \mathrm{Spec} B} B_{[\mathfrak{q}]} \subset \bigcap_{\mathfrak{p} \in \mathrm{Spec} A} A_{[\mathfrak{p}]} = A,$$

hence $A = B$, and we will be done.

b) We first prove that for any $x \in B_{[\mathfrak{q}]}$ we have $F(x) \in A_{[\mathfrak{p}]}$. Put $y := F(x) - 1$. Suppose $F(x) \notin A_{[\mathfrak{p}]}$, hence $y \notin A_{[\mathfrak{p}]}$. Clearly $F(x) \notin k$. By hypothesis $1 + y = F(x) \in R^*$ and $\frac{1}{1+y} \in A$. Also $\frac{y}{1+y} = 1 - \frac{1}{1+y} \in A$. Since $y \notin A_{[\mathfrak{p}]}$ we conclude that $\frac{1}{1+y} \in \mathfrak{p}$. On the other hand $\frac{y}{1+y} = y \cdot (\frac{1}{1+y}) \in (B_{[\mathfrak{q}]} \cdot \mathfrak{p}) \cap A \subset \mathfrak{q}_{[\mathfrak{q}]} \cap A = \mathfrak{p}$. We arrive at the contradiction $1 = \frac{1}{1+y} + \frac{y}{1+y} \in \mathfrak{p}$. Thus indeed $F(x) \in A_{[\mathfrak{p}]}$.

c) For $\ell = 0, 1, 2, \ldots$ we successively define polynomials $\Delta^\ell F(T)$ by

$$\Delta^0 F(T) := F(T), \quad \Delta^{\ell+1} F(T) := \Delta^\ell F(T+1) - \Delta^\ell F(T).$$

For every $x \in B_{[\mathfrak{q}]}$ we have $F(x) \in A_{[\mathfrak{p}]}$, thus also $\Delta^\ell F(x) \in A_{[\mathfrak{p}]}$ for any $\ell \in \mathbb{N}$. But $\Delta^{d-1} F(T) = d! T + c$ with $c \in k$. Thus $(d!)x \in A_{[\mathfrak{p}]}$ for every $x \in B_{[\mathfrak{q}]}$. Since $1/d! \in A \subset A_{[\mathfrak{p}]}$, we conclude that $A_{[\mathfrak{p}]} = B_{[\mathfrak{q}]}$, q.e.d.

Example 10. We denote the prime ring in R by $\mathbb{Z} \cdot 1$. Let $d \in \mathbb{N}$. Assume that $d! \in R^*$ and $1 + x^d \in R^*$ for all $x \in R$ with $x^d \notin \mathbb{Z} \cdot 1$. The subring A of R generated by $1/d!$ and the elements $1/(1 + x^d)$ with $x \in R$, $x^d \notin \mathbb{Z} \cdot 1$ is Prüfer in R. □

N.B. For $d = 2$ and R a field this example states a slight improvement of Dress's lemma cited above.

Remark. The condition $d! \in R^*$ cannot be omitted. For example, let $R := \mathbb{F}_2[T]/(1+T^2)$ with $\mathbb{F}_2$ the field consisting of 2 elements. Let A be the subring of R generated by the elements $1/(1+x^2)$ for all $x \in R$ with $x^2 \neq 1$. Then $A = \mathbb{F}_2$, and this not Prüfer in R, since $\mathbb{F}_2$ is not integrally closed in R.

As an illustration what has been done so far we return to Example 5. Thus let V be an affine variety over some field k which is not algebraically closed. Let U be a k-Zariski-open subset of $V(k)$, and let R be the ring of regular functions on U. We choose a monic polynomial $F(T) \in k[T]$, $F \neq 1$, which has no zeros in k.

Let B be any subring of R containing k (e.g. $B = k$). Let H_0 denote the subring $B[\frac{1}{F(x)} \mid x \in R]$ of R generated by B and the functions $\frac{1}{F(x)}$ for all $x \in R$. Let H denote the integral closure of H_0 in R.

Theorem 6.9. i) *H is an R-Prüferring.*

ii) *H is the set of all $x \in R$ such that $v(x) \geq 0$ for every Manis F-valuation v on R with $v(b) \geq 0$ for all $b \in B$.*

iii) *$H = B[\frac{x^i}{F(x)} \mid x \in R,\ 0 \leq i \leq 1]$.*

iv) *If the characteristic of k is zero or exceeds d, then $H = H_0$.*

Proof. H is an R-Prüferring by Theorem 1. Thus H is the intersection of the valuation rings A_v with v running through the set Ω of all Manis valuations on R with $A_v \supset H$.

Let v be a Manis valuation on R. Then $v \in \Omega$ iff $A_v \supset H_0$. This means that $A_v \supset B$ and $v\big(\frac{1}{F(x)}\big) \geq 0$ for every $x \in R$. If $x \notin A_v$ then $v(F(x)) < 0$, hence $v\big(\frac{1}{F(x)}\big) > 0$ automatically. Let $x \in A_v$. Then $v\big(\frac{1}{F(x)}\big) \geq 0$ iff $v(F(x)) = 0$ iff $\bar{F}(\bar{x}) \neq 0$ for $\bar{F}(T)$ the image of $F(T)$ in $\kappa(v)[T]$ and $\bar{x}$ the image of x in $\kappa(v)$. Thus Ω is the set of all Manis F-valuations v on R with $A_v \supset B$.

The ring $H' := B\big[\frac{x^i}{F(x)} \mid x \in R,\ 0 \leq i \leq 1\big]$ is Prüfer in R by Theorem 6. Every valuation $v \in \Omega$ has nonnegative values on H'. Thus $H_0 \subset H' \subset H$. Since H' is integrally closed in R, we have $H' = H$. If $d! \in k^*$, then we know from Theorem 8 that H_0 is Prüfer in R and conclude that $H_0 = H$. □

§7 Principal ideal results

We start out for a generalization of the second half of Roquette's theorem 1 in [R] mentioned in §6. We will rely on techniques developed by Alan Loper in the case of subrings of fields [Lo_1], [Lo_2].

In the following we fix a ring A and a monic polynomial $F(T) \in A[T]$ of degree $d \geq 1$.

Definition 1 (cf. [Lo_1]). Let $\varphi: A \to B$ be a ring extension of A. We call the polynomial F *unit valued in* B (abbreviated: uv in B), if $F(b) \in B^*$ for every $b \in B$. {Of course, $F(b) := F^\varphi(b)$ with $F^\varphi(T)$ the image polynomial of $F(T)$ in $B[T]$.}

More precisely we then should call F "uv with respect to φ", but in the following it will be always clear which homomorphism φ from A to B is taken.

N.B. If F is uv in some extension B of A different from the null ring then certainly $d \geq 2$.

Proposition 7.1 (cf. [Lo_1, Prop.1.14]). *Let $\mathfrak{m}$ be a maximal ideal of A. Then $F(T)$ is uv in $A_\mathfrak{m}$ iff $F(A) \subset A \setminus \mathfrak{m}$.*

Proof. If there exists some $a \in A$ with $F(a) \in \mathfrak{m}$, then certainly $F(T)$ is not uv in $A_\mathfrak{m}$. Assume now that $F(A) \subset A \setminus \mathfrak{m}$. Suppose that $F(T)$ is not uv in $A_\mathfrak{m}$. We have some $a \in A$, $s \in A \setminus \mathfrak{m}$ with $F\big(\frac{a}{s}\big) \in \mathfrak{m}A_\mathfrak{m}$. Since the ideal $\mathfrak{m}$ is maximal there exists some $t \in A$ with $st \equiv 1 \mod \mathfrak{m}$. Then in $A_\mathfrak{m}$

$$\frac{F(at)}{1} \equiv F\big(\frac{a}{s}\big) \equiv 0 \qquad \mod \mathfrak{m}A_\mathfrak{m},$$

hence $F(at) \in \mathfrak{m}$. This contradiction proves that $F(T)$ is uv in $A_\mathfrak{m}$. □

Corollary 7.2. *$F(T)$ is uv in A iff $F(T)$ is uv in $A_\mathfrak{m}$ for every maximal ideal $\mathfrak{m}$ of A.* □

We write $F(T) = T^d + c_1T^{d-1} + \cdots + c_d$ with $a_i \in A$, and introduce the homogenization $G(X,Y) \in A[X,Y]$ of F,

$$G(X,Y) := Y^dF\big(\frac{X}{Y}\big) = X^d + c_1X^{d-1}Y + \cdots + c_dY^d.$$

Proposition 7.3. *Let $\mathfrak{p}$ be a prime ideal of A. The following are equivalent.*

i) *F is uv in $A_\mathfrak{p}$.*
ii) *F is uv in $k(\mathfrak{p}) = Quot(A/\mathfrak{p})$, i.e. F has no zero in $k(\mathfrak{p})$.*
iii) *If $x, y \in A$ and $G(x,y) \in \mathfrak{p}$, then $y \in \mathfrak{p}$.*
iv) *If $x, y \in A$ and $G(x,y) \in \mathfrak{p}$, then $x \in \mathfrak{p}$ and $y \in \mathfrak{p}$.*

Proof. i) $\Leftrightarrow$ ii) is evident. iv) $\Rightarrow$ iii) is trivial, and iii) $\Rightarrow$ iv) is evident, since the form $G(X,Y)$ contains the term X^d.
i) $\Rightarrow$ iii): Let $x, y \in A$ and $G(x,y) \in \mathfrak{p}$. Suppose $y \notin \mathfrak{p}$. Then we have in $A_\mathfrak{p}$

$$F\big(\frac{x}{y}\big) = \frac{G(x,y)}{y^d} \in \mathfrak{p}A_\mathfrak{p}.$$

This contradicts the assumption that F is uv in $A_\mathfrak{p}$.
iii) $\Rightarrow$ i): Let $a \in A$, $s \in A \setminus \mathfrak{p}$ be given. Then $G(a,s) \in A \setminus \mathfrak{p}$. Thus

$$F\big(\frac{a}{s}\big) = \frac{G(a,s)}{s^d} \in A_\mathfrak{p}^*.$$ □

Proposition 7.4 (cf. [Lo_2, Cor.2.3]). *Assume that $(A, \mathfrak{p})$ is a Manis pair in some ring R. Let v denote a Manis valuation on R with $A_v = A$, $\mathfrak{p}_v = \mathfrak{p}$. The following are equivalent.*

i) *F is uv in $A_{\mathfrak{p}}$.*
ii) *v is an F-valuation.*
iii) *$v(G(x,y)) = d\min(v(x),v(y))$ for all $x,y \in R$.*

Proof. The equivalence i) $\Leftrightarrow$ ii) is clear from i) $\Leftrightarrow$ ii) in Proposition 3.
i) $\Rightarrow$ iii): Let $x,y \in R$ be given. The formula is a priori valid if $v(x) < v(y)$, since $G(X,Y)$ contains the term X^d. It is also valid if $v(x) = v(y) = \infty$. Assume now that $v(x) \geq v(y) \neq \infty$. We choose some $z \in R$ with $v(yz) = 0$. This is possible since v is Manis. Then $v(xz) \geq 0$. Thus $xz \in A$ and $yz \in A \setminus \mathfrak{p}$. We know from Prop. 3 that $G(xz,yz) = z^d G(x,y) \in A \setminus \mathfrak{p}$. Thus $v(G(x,y)) = -dv(z) = dv(y)$.
iii) $\Rightarrow$ i): Let $x,y \in A$ and $G(x,y) \in \mathfrak{p}$. Then the formula in iii) tells us that $x \in \mathfrak{p}$ and $y \in \mathfrak{p}$. Thus F is uv in $A_{\mathfrak{p}}$ by Proposition 3. □

We now study finitely generated A-submodules $\mathfrak{a}$ of R with $R\mathfrak{a} = R$. These submodules should be viewed as analogues of the finitely generated fractional ideals in the classical case that A is a domain and R its quotient field. We are looking for criteria that some power $\mathfrak{a}^d$ is a *principal module*, i.e. $\mathfrak{a}^d = Rb$ with some $b \in R^*$.

Definition 2. Let $(a_1,\dots,a_n)$ be a finite sequence in R. The *F-transform* of this sequence is the sequence $(b_1,\dots,b_n)$ in R defined inductively by

$$b_1 := a_1, \quad b_i := G(b_{i-1}, a_i^{d^{i-2}}) \qquad (i > 1).$$

In the following lemmas $(a_1,\dots,a_n)$ is a sequence in R and $(b_1,\dots,b_n)$ is its F-transform.

Lemma 7.5. *Assume that all $a_i \in A$. Let $\mathfrak{p}$ be a prime ideal of A such that F is uv in $A_{\mathfrak{p}}$. Then $Aa_1 + \dots + Aa_n \subset \mathfrak{p}$ iff $b_n \in \mathfrak{p}$.*

Proof. If $x,y \in A$ and $t \in \mathbb{N}$, then $Ax + Ay \subset \mathfrak{p}$ iff $Ax + Ay^t \subset \mathfrak{p}$. By Proposition 3 the latter is equivalent to $G(x,y^t) \in \mathfrak{p}$. The lemma follows from this by induction on n. □

Lemma 7.6 (cf. [Lo$_2$, Cor.2.4]). *Assume that A is the valuation ring A_v of a Manis valuation v on some ring R which is also an F-valuation. Then*

$$v(b_n) = d^{n-1}\min\{v(a_1),\dots,v(a_n)\}.$$

The proof goes by induction on n using the formula in Proposition 4.iii. □

Lemma 7.7. *Let $\mathfrak{a} := Aa_1 + \dots + Aa_n$. Assume that F is uv in R. Then*

$$R\mathfrak{a} = R \Longleftrightarrow b_n \in R^*.$$

Proof. $\Leftarrow$: This is evident since $b_n \in \mathfrak{a}$.
$\Rightarrow$: Suppose $b_n \notin R^*$. We choose a maximal ideal $\mathfrak{M}$ of R containing b_n. Our polynomial F is uv in R hence uv in $R_{\mathfrak{M}}$ by Corollary 2. Now Lemma 5, applied to

F as a polynomial over R, tells us that $Ra_1 + \cdots + Ra_n \subset \mathfrak{M}$. This contradicts the assumption $R\mathfrak{a} = R$. Thus $b_n \in R^*$. □

Now we are prepared to prove a generalization of the theorem by Roquette mentioned in §6.

THEOREM 7.8 (cf. [R, Th.1] for R a field). *Assume that S is a set of Manis valuations on a ring R and that $A = \bigcap\limits_{v \in S} A_v$. Assume further that there exists a monic polynomial $F(T) \in A[T]$ of degree $d \geq 1$ with the following two properties:*
(i) *$F(T)$ is uv in R.*
(ii) *Every $v \in S$ is an F-valuation.*

Then A is Prüfer in R. If $\mathfrak{a}$ is any finitely generated A-submodule of R with $R\mathfrak{a} = R$ then there exists some $t \in \mathbb{N}$ such that $\mathfrak{a}^{d^t}$ is principal. More precisely, if $a_1, \ldots, a_n$ is a system of generators of $\mathfrak{a}$ and $(b_1, \ldots, b_n)$ is the F-transform of the sequence $(a_1, \ldots, a_n)$, then

$$\mathfrak{a}^{d^{n-1}} = Ab_n.$$

PROOF. Theorem 6.5 tells us that A is Prüfer in R. Let $a_1, \ldots, a_n$ be a system of generators of $\mathfrak{a}$ and $(b_1, \ldots, b_n)$ the F-transform of $(a_1, \ldots, a_n)$. Lemma 7 tells us that $b_n \in R^*$.

It is evident that $b_n \in \mathfrak{a}^{d^{n-1}}$. The module $\mathfrak{a}^{d^{n-1}}$ is generated over A by the monomials $a_1^{e_1} \ldots a_n^{e_n}$ with $e_i \geq 0$, $e_1 + \cdots + e_n = d^{n-1}$. We now verify that

$$v(a_1^{e_1} \ldots a_n^{e_n}) \geq v(b_n) \tag{$*$}$$

for every such monomial and every $v \in S$. It then follows that $a_1^{e_1} \ldots a_n^{e_n}/b_n$ is an element of A_v for every $v \in S$, hence of A, and we conclude that $\mathfrak{a}^{d^{n-1}} = Ab_n$. The verification of $(*)$ is immediate by use of Lemma 6. Let $\gamma := \min\{v(a_1), \ldots, v(a_n)\}$. Then $v(a_1^{e_1} \ldots a_n^{e_n}) \geq (e_1 + \cdots + e_n)\gamma = d^{n-1}\gamma = v(b_n)$. □

In part II of the paper we will see that for A a Prüfer subring of a ring R the finitely generated A-submodules $\mathfrak{a}$ of R with $R\mathfrak{a} = R$ form an Abelian group. The quotient of this group by the subgroup of principal modules should be called the *class group of A in R*. Starting with Theorem 8 it is possible to get bounds on the torsion of the class group in good cases in much the same way as Roquette has explicated for R a field [R]. Here we only quote the following theorem which is an immediate consequence of Theorem 8.

THEOREM 7.9 (cf.[R, Th.2]). *Assume again that $A = \bigcap\limits_{v \in S} A_v$ for a set S of Manis valuations on some ring R. Assume further that there exist non-constant monic polynomials $F_1(T), \ldots, F_r(T)$ with coefficients in A $(r \geq 1)$, such that for every $j \in \{1, \ldots, r\}$ the following holds*
(1) *F_j is uv in R.*
(2) *Every $v \in S$ is an F_j-valuation.*

Let d denote the greatest common divisor of the degrees of $F_1, \ldots, F_r$. Then A is Prüfer in R, and for each finitely generated A-submodule $\mathfrak{a}$ of R with $R\mathfrak{a} = R$ there exists some $t \in \mathbb{N}$ such that $\mathfrak{a}^{d^t}$ is principal. □

Example 7.10. Let R be a ring such that $X^d + 1$ is uv in R for some (even) $d \in \mathbb{N}$ and $d!$ is a unit in R. Let A be a subring of R which contains $1/d!$ and the elements $1/(1 + x^d)$ for all $x \in R$. Then A is Prüfer in R by Example 10 in §6. For every finitely generated A-submodule $\mathfrak{a}$ of R with $\mathfrak{a}R = R$ there exists some $t \in \mathbb{N}$ with $\mathfrak{a}^{d^t}$ principal.

Proof. A is the intersection of the rings $A_{[\mathfrak{m}]}$ with $\mathfrak{m}$ running through the maximal ideals of A (Remark 5.5). These rings are Manis in R. The polynomial $X^d + 1$ is uv in $A_\mathfrak{m}$ for every $\mathfrak{m}$ (Cor.2), and thus the Manis valuations giving the rings $A_{[\mathfrak{m}]}$ are $(X^d + 1)$-valuations. Theorem 8 applies. □

In an important more special situation this result can be improved. Assume that $1 + \Sigma R^d \subset R^*$. A subring A of R containing the elements $1/(1 + q)$ with $q \in \Sigma R^d$ is Prüfer in R. If $\mathfrak{a} = Ax_1 + \cdots + Ax_n$ is a finitely generated submodule of R with $R\mathfrak{a} = R$, then $\mathfrak{a}^d = A(x_1^d + \cdots + x_n^d)$. This has been proved recently by E. Becker and V. Powers [BP, Cor. 5.11, Cor.5.13].

A slight expansion of the techniques used so far will give us a theorem containing this result of Becker and Powers as a special case, together with a proof which is rather different from the one in [BP].

Definition 3. Let $H(X_1, \ldots, X_n) \in A[X_1, \ldots, X_n]$ be a form, i.e. a homogeneous polynomial over A in $n \geq 2$ variables. Let $\varphi\colon A \to K$ be a homomorphism into a field K. We call H *isotropic over* K, if the image form $H^\varphi(X_1, \ldots, X_n) \in K[X_1, \ldots, X_n]$ is isotropic, i.e. has a non trivial zero in K^n, and we call H *anisotropic over* K otherwise.

In the following it will be always clear which homomorphism φ is under consideration. Thus the impreciseness in this definition will do no harm.

Theorem 7.11. *Let S be a set of Manis valuations on a ring R and $A := \bigcap_{v \in S} A_v$. Assume there is given a form $H(X_1, \ldots, X_n)$ over A in $n \geq 2$ variables of degree $d \geq 1$ with the following properties:*

i) *For every maximal ideal $\mathfrak{M}$ of R the form H is anisotropic over $R/\mathfrak{M}$.*
ii) *For every $v \in S$ the form H is anisotropic over $\kappa(v)$.*

Then A is Prüfer in R. If $\mathfrak{a}$ is an A-submodule of R generated by n elements $x_1, \ldots, x_n$ and $R\mathfrak{a} = R$ then $\mathfrak{a}^d = H(x_1, \ldots, x_n)A$.

Proof. a) We start with a proof of the second claim. Suppose that $H(x_1, \ldots, x_n)$ is not a unit in R. Then there exists a maximal ideal $\mathfrak{M}$ of R with $H(x_1, \ldots, x_n) \in \mathfrak{M}$. Since $Rx_1 + \cdots + Rx_n = R$ we conclude that H is isotropic over $R/\mathfrak{M}$, in contradiction to assumption (i) above. Thus $H(x_1, \ldots, x_n) \in R^*$.

b) Let $v \in S$ be given. We verify that

$$(*) \qquad v(H(x_1, \dots, x_n)) = d \min\{v(x_1), \dots, v(x_n)\}.$$

This is obvious if $v(x_i) = \infty$ for all $i \in \{1, \dots, n\}$. Assume now that $\gamma := \min\{v(x_1), \dots, v(x_n)\} < \infty$. We choose some $z \in R$ with $v(z) = -\gamma$, which is possible, since v is Manis. Then $v(zx_i) \geq 0$ for all $i \in \{1, \dots, n\}$ and $v(zx_i) = 0$ for at least one i. Since H is anisotropic over $\kappa(v)$ we conclude that $v(H(zx_1, \dots, zx_n)) = 0$, hence $v(H(x_1, \dots, x_n)) = -dv(z) = d\gamma$, as desired.
c) Now we see, as in the proof of Theorem 8, that

$$v(x_1^{e_1} \dots x_n^{e_n}) \geq v(H(x_1, \dots, x_n))$$

for any integers $e_i \geq 0$ with $e_1 + \dots + e_n = d$ and any $v \in S$, and we conclude that $x_1^{e_1} \dots x_n^{e_n}/H(x_1, \dots, x_n) \in A$. This proves that $\mathfrak{a}^d = H(x_1, \dots, x_n)A$.
d) Let

$$G(X, Y) := H(X, \dots, X, Y) = c_0 X^d + c_1 X^{d-1} Y + \dots + c_d Y^d.$$

$c_0 = H(1, \dots, 1, 0)$ is a unit in A, since the elements $1, \dots, 1, 0$ generate the ideal $\mathfrak{a} = A$ and $\mathfrak{a}^d = H(1, \dots, 1, 0)A$. We consider the monic polynomial

$$F(T) := c_0^{-1} G(T, 1) \in A[T].$$

F is uv in R, since $H(x, \dots, x, 1) \in R^*$ for every $x \in R$. If $v(x) \geq 0$ for some $v \in S$, then $v(H(x, \dots, x, 1)) = v(1) = 0$. Thus every $v \in S$ is an F-valuation. We conclude by Theorem 6.5 that A is Prüfer in R. □

REMARK. The multiplicative ideal theory to be developed in part II of this paper will give a more natural proof that A is Prüfer in R.

In order to exploit Theorem 11 in the real algebraic setting, we need an easy lemma.

LEMMA 7.12. *Let $H(X_1, \dots, X_n)$ be a form over a ring A of degree $d \geq 1$ in $n \geq 2$ variables. For each $i \in \{1, \dots, n\}$ we define*

$$F_i(T_1, \dots, T_{n-1}) := H(T_1, \dots, T_{i-1}, 1, T_i, \dots, T_{n-1}).$$

The following are equivalent

(1) *H is anisotropic over $A/\mathfrak{m}$ for every maximal ideal $\mathfrak{m}$ of A.*
(2) *$F_i(x_1, \dots, x_{n-1}) \in A^*$ for all $x_1, \dots, x_{n-1} \in A$ and $1 \leq i \leq n$.*

PROOF. (1) $\Longrightarrow$ (2): Let $x_1, \dots, x_{n-1} \in A$ and $i \in \{1, \dots, n\}$. Then $H(x_1, \dots, x_{i-1}, 1, x_i, \dots, x_{n-1}) \notin \mathfrak{m}$ for every maximal ideal $\mathfrak{m}$ of A. Thus $F_i(x_1, \dots, x_{n-1}) \in A^*$.
(2) $\Longrightarrow$ (1): Suppose there exists a maximal ideal $\mathfrak{m}$ of A such that H is isotropic over $A/\mathfrak{m}$. Then there exist elements $a_1, \dots, a_n \in A$ with $H(a_1, \dots, a_n) \in \mathfrak{m}$ but $a_i \notin \mathfrak{m}$ for some i. We choose an element $b_i \in A$ with $a_i b_i \equiv 1 \mod \mathfrak{m}$. We have $b_i^d H(a_1, \dots, a_n) = H(a_1 b_i, \dots, a_n b_i) \equiv F_i(a_1 b_i, \dots, a_{i-1} b_i, a_{i+1} b_i, \dots, a_n b_i) \mod \mathfrak{m}$. Thus, $F_i(a_1 b_i, \dots, a_{i-1} b_i, a_{i+1} b_i, \dots, a_n b_i) \in \mathfrak{m}$, a contradiction. □

Corollary 7.13 (cf. [BP]). *Let $d \in \mathbb{N}$ and let R be a ring with $1 + \Sigma R^{2d} \subset R^*$. Then the subring*

$$H := H_d(R) = \mathbb{Z}[\frac{1}{1+q} | q \in \Sigma R^{2d}]$$

is Prüfer in R. For each finitely generated H-submodule $\mathfrak{a} = Hx_1 + \cdots + Hx_n$ of R with $\mathfrak{a}R = R$ we have $\mathfrak{a}^{2d} = (x_1^{2d} + \cdots + x_n^{2d})H$.

Proof. Applying Theorem 6.8 with $F(T) = 1 + T^{2d}$ we see that H is Prüfer in R (cf. §6, Example 10). For every maximal ideal $\mathfrak{m}$ of H we choose a Manis valuation v on R with $A_v = H_{[\mathfrak{m}]}$, $\mathfrak{p}_v = \mathfrak{m}_{[\mathfrak{m}]}$. Let S denote the set of these valuations. Then $H = \bigcap_{v \in S} A_v$ (cf. 5.5). Now, if $v \in S$, $A_v = H_{[\mathfrak{m}]}$, then $H/\mathfrak{m} = H_{[\mathfrak{m}]}/\mathfrak{m}_{[\mathfrak{m}]}$, as is easily checked, and we learn from Proposition 1.7 that $\kappa(v)$ is the quotient field of $H/\mathfrak{m}$. Since $H/\mathfrak{m}$ is already a field, we have $\kappa(v) = H/\mathfrak{m}$. Let $n \geq 2$. Using Lemma 12 we see that the form $X_1^{2d} + \cdots + X_n^{2d}$ is anisotropic in $R/\mathfrak{M}$ for every maximal ideal $\mathfrak{M}$ of R, and also anisotropic in $H/\mathfrak{m}$ for every maximal ideal $\mathfrak{m}$ of H. Now Theorem 11 gives the second claim above. □

Becker and Powers have proved that $1 + \Sigma R^{2d} \subset R^*$ implies $1 + \Sigma R^2 \subset R^*$, and that then $H := H_d(R)$ coincides with $H_1(R)$ and the "real holomorphy ring" of R [BP, Prop.5.1 and Prop.5.7]. Thus, if $\mathfrak{a}$ is a finitely generated H-submodule of R with $R\mathfrak{a} = R$, then already $\mathfrak{a}^2$ is a principal submodule.

References

[A] T. Akiba, *Remarks on generalized rings of quotients, III*, J. Math. Kyoto Univ. 9-2 (1969), pp. 205–212.

[B] E. Becker, *Valuations and real places in the theory of formally real fields*, Lecture Notes in Math. 959, pp. 1–40, Springer-Verlag, 1982.

[B$_1$] E. Becker, *On the real spectrum of a ring and its applications to semialgebraic geometry*, Bull. Amer. Math. Soc. **15** (1986), pp. 19–60.

[B$_2$] E. Becker, *Partial orders on a field and valuation rings*, Comm. Algebra **7(18)** (1979), pp. 1933–1976.

[B$_3$] E. Becker, *Theory of real fields and sums of powers*, forthcoming book.

[BP] E. Becker and V. Powers, *Sums of powers in rings and the real holomorphy ring*, Preprint.

[Be] R. Berr, *Basic principles for a morphological theory of semi-algebraic morphisms*, Habilitationsschrift, Univ. Dortmund 1994.

[BCR] J. Bochnak, M. Coste, M.-F. Roy, *Géométrie algébrique réelle*, Ergeb. Math. Grenzgeb. 3. Folge, Band 12, Springer 1987.

[Bo] N. Bourbaki, *Algèbre commutative*, Chap.1–7, Hermann Paris, 1961 – 1965.

[BS] L. Bröcker, J.-H. Schinke, *On the L-adic spectrum*, Schriften Math. Inst. Univ. Münster, 2. Serie, Heft **40** (1986).

[Da] E.D. Davis, *Overrings of commutative rings III: Normal pairs.* Trans. Amer. Math. Soc. 182 (1973), pp. 175–185.

[D] A. Dress, *Lotschnittebenen mit halbierbarem rechten Winkel*, Arch. Math. **16** (1965), pp. 388–392.

[Eg] N. Eggert, *Rings whose overrings are integrally closed in their complete quotient ring*, J. Reine Angew. Math. **282** (1976), pp. 88–95.

[E] O. Endler, *Valuation theory*, Springer-Verlag, 1972.

[Gi] R. Gilmer, *Multiplicative ideal theory*, Marcel Dekker, New York, 1972.

[Gi_1] R. Gilmer, *Two constructions of Prüfer domains*, J. Reine Angew. Math. **239/240** (1970), pp. 153–162.

[Gr] J. Gräter, *Der allgemeine Approximationssatz für Manisbewertungen*, Mh. Math. **93** (1982), pp. 277–288.

[G_1] M. Griffin, *Prüfer rings with zero divisors*, J. Reine Angew. Math. **239/240** (1970), pp. 55–67.

[G_2] M. Griffin, *Valuations and Prüfer rings*, Canad. J. Math. **26** (1974), pp. 412–429.

[Ho] M. Hochster, *Prime ideal structure in commutative rings*, Trans. Amer. Math. Soc. **142** (1969), 43–60.

[Hu_1] R. Huber, *Bewertungsspektrum und rigide Geometrie*, Habilitationsschrift, Univ. Regensburg, 1990.

[Hu_2] R. Huber, *Continuous valuations*, Math. Z. **212** (1993), pp. 455–477.

[Hu_3] R. Huber, *Semirigide Funktionen*, Preprint Univ. Regensburg 1990.

[HK] R. Huber and M. Knebusch, *On valuation spectra*, Contemp. Math., **155** (1994), pp. 167–206.

[Huc] J. A. Huckaba, *Commutative rings with zero divisors*, Marcel Dekker, New York, 1988.

[K] M. Knebusch, *Real closures of commutative rings* I, J. Reine Angew. Math. **274/275**, pp. 61–89.

[KS] M. Knebusch and C. Scheiderer, *Einführung in die reelle Algebra*, Vieweg, 1989.

[Ko] S. Kochen, *Integer valued rational functions over the p-adic numbers: A p-adic analogue of the theory of real fields*, Proc. Symp. Pure Math., vol. XII, Number Theory (1976), pp. 57–73.

[La] T. Y. Lam, *Orderings, valuations and quadratic forms*, in "Regional Conference Series in Math." vol. 52, 1983.

[Lb] J. Lambek, *Lectures on rings and modules*, Waltham-Toronto-London, 1966.

[LM] M. Larsen and P. McCarthy, *Multiplicative theory of ideals*, Academic Press, New York and London, 1971.

[L] D. Lazard, *Autour de la platitude*, Bull. Soc. Math. France **97** (1969), pp. 81–128.

[Lo_1] A. Loper, *On rings without a certain divisibility property*, J. Number Theory **28** (1988), pp. 132–144.

[Lo_2] A. Loper, *On Prüfer non-D-rings*, J. Pure and Applied Algebra **96** (1994), pp. 271–278.

[M] M. E. Manis, *Valuations on a commutative ring*, Proc. Amer. Math. Soc. **20** (1969), pp. 193–198.

[Mar] M. Marshall, *Orderings and real places on commutative rings*, J. Algebra **140**, No. 2 (1991), pp. 484–501.

[Mat] H. Matsumura, *Commutative algebra*, Benjamin publishing company, 1970.

[P] V. Powers, *Valuations and higher level orders in commutative rings*, J. Algebra **172** (1995), pp. 255–272.

[Pt] M. Prechtel, *Universelle Vervollständigungen in der Kategorie der reell abgeschlossen Räume*, Dissertation der Univ. Regensburg, 1992.

[PR] A. Prestel and P. Roquette, *Formally p-adic fields*, Lecture Notes in Math. 1050, Springer-Verlag, 1984.

[Rh] C.P.L. Rhodes, *Relative Prüfer pairs of commutative rings*, Commun. in Algebra 19 (12), 3423–3445 (1991).

[Ri] F. Richman, *Generalized quotient rings*, Proc. Amer. Math. Soc. **16** (1965), pp. 794–799.

[R] P. Roquette, *Principal ideal theorems for holomorphy rings in fields*, J. Reine Angew. Math. **262/263** (1973), pp. 361–374.

[R_1] P. Roquette, *Bemerkungen zur Theorie der formal p-adischen Körper*, Beitr. z. Algebra und Geometrie **1** (1971), pp. 177–193.

[Sa] P. Samuel, *La notion de place dans un anneau*, Bull. Soc. Math. France **85** (1957), pp. 123–133.

[Sa_1] Séminaire d'algebre commutative dirigé par P. Samuel, *Les épimorphismes d'anneaux*, 1967/68.

[S] H.-W. Schülting, *Real holomorphy rings in real algebraic geometry*, Lecture Notes in Math. 959, pp. 433–442, Springer-Verlag, 1982.

[Sch] N. Schwartz, *The basic theory of real closed spaces*, Regensburger Math. Schriften **15**, 1987.

[Sch_1] N. Schwartz, *The basic theory of real closed spaces*, Memoirs Amer. Math. Soc. **397** (1989).

[Sch_2] N. Schwartz, *Letter to the authors*, 20.09.1994.

[St] H. H. Storrer, *Epimorphismen von kommutativen Ringen*, Comment. Math. Helv. **43** (1968), pp. 378–401.

[St_1] H. H. Storrer, *A characterization of Prüfer domains*, Canad. Bull. Math. 12 (1969), pp. 809–812.

Prof. Dr. Manfred Knebusch
Digen Zhang
Department of Mathematics
University of Regensburg
D-93040 Regensburg
Germany
Manfred.Knebusch@mathematik.uni-regensburg.de
Digen.Zhang@mathematik.uni-regensburg.de

Remarks on Quenching

Bernd Kawohl

Received: November 2 1995
Revised: April 29, 1996

Communicated by Bernold Fiedler

Abstract. Consider the parabolic problem

$$u_t - \text{div}\ (a(u, \nabla u)\nabla u) = -u^{-p} \tag{1}$$

for $t > 0$, $x \in \mathbb{R}^n$ under initial and boundary conditions $u = 1$, say. Since p is assumed positive, the right hand side becomes singular as $u \to 0$. When u reaches zero in finite or infinite time, one says that the solution quenches in finite or infinite time. This article gives a survey of results on this kind of problem and emphasizes those that have been obtained at the SFB 123 in Heidelberg. It is an updated version of an invited survey lecture at the International Congress of Nonlinear Analysts in Tampa, August 1992. To be specific, I shall cover existence and nonexistence of quenching points, asymptotic behaviour of the solutions in space and time near the quenching points, qualitative behaviour, application to mean curvature flow and phase transitions, reaction in porous medium flow etc..

The tools are variational methods and suitable maximum principles. Many of the results presented in this article were obtained with my coauthors Acker, Dziuk, Fila, Kersner and Levine, but related results will also be mentioned.

1991 Mathematics Subject Classification: 35K65, 35K57, 35K60, 35B05, 35B65

Model Problems

For the sake of simplicity I shall discuss four special cases of (1), namely:

$$u_t - \Delta u = -u^{-p}, \tag{A}$$

$$u_t - (\varphi(u_x))_x = -u^{-p}, \tag{B}$$

$$u_t - \frac{u_{xx}}{1+u_x^2} = -\frac{1}{u}, \tag{B'}$$

$$u_t - (u^m)_{xx} = -u^{-p}. \tag{C}$$

Note that for $n = 1$ case (A) is a special case of both (B) and (C). Equation (B') is a special case of (B), which has a significant application in the mathematical description of mean curvature flow of rotationally symmetric twodimensional surfaces in $\mathbb{R}^3$. To see this imagine the x-axis to be the axis of a revolutionary body whose surface at time t is described by $u(t,x)$. Then (see Figure 1) its inward velocity v is given by

$$v = \frac{u_t}{\sqrt{1+u_x^2}},$$

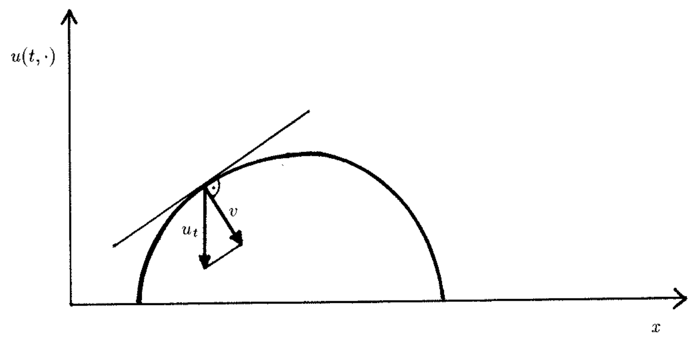

FIGURE 1: Derivation of (B')

while its principle curvatures are

$$\frac{u_{xx}}{\sqrt{1+u_x^2}^3} \qquad \text{in } x\text{−direction,}$$
$$-\frac{1}{u}\frac{1}{\sqrt{1+u_x^2}} \qquad \text{in } v\text{−direction.}$$

Therefore, after rescaling time by a factor of two, the mean curvature flow of our surface is described by

$$u_t = \frac{u_{xx}}{1+u_x^2} - \frac{1}{u} = (\arctan u_x)_x - \frac{1}{u},$$

and, incidentally, this is how one can see that (B') is a particular case of (B).

Chronologically ordered, my coauthors and I wrote the following papers on Problems (A), (B) and (C). Problem (A) was dealt with in [AK, KP, FK1, FK2] and [K], Problem (B') was studied in [DK,K], Problem (B) was treated in [FKL] and Problem (C) in [KK].

Quenching Occurs

Let us first consider Problem (A):

$$\begin{aligned} u_t - \Delta u &= -u^{-p}, \qquad x \in \Omega,\ t > 0, \\ u &= 1 \qquad\qquad \text{on the parabolic boundary} \end{aligned} \tag{2}$$

Here Ω is a bounded domain in $\mathbb{R}^n$. For $n = 1$ and $p = 1$ this problem was studied by Kawarada [Ka] and he stated the following result.

Theorem 1: *If there are no stationary solutions to* (2), *i.e. if Ω is too large, then*
i) u reaches zero in some point x_0 in finite time T.
ii) $u_t(t, x_0) \to -\infty$ as $t \nearrow T$.

Statement i) has been derived for numerous more general situations, e.g. by Acker and Walter, Levine and Montgomery or Lieberman to higher dimension, hyperbolic equations, nonlinear boundary conditions such as $(\partial u/\partial n) = -u^{-p}$ and so on. One can prove i) by energy methods or by comparison principles. The proof of ii) was wrong as stated by Kawarada. This was noted and corrected by Chan and Kwong, and by Acker and myself in 1987. Moreover, it was shown in [AK] that both statements of Theorem 1 hold for general $n \in \mathbb{N}$ and $p > 0$, provided Ω is a ball. In this case x_0 is uniquely determined and is the center of the ball.

In the same year I discovered why quenching and blow up problems have so much in common. In fact one can be transformed into the other, see [KP]. The fact that both classes of problems are amenable to similar techniques had puzzled people, e.g. Bandle and Stakgold [BS] or Friedman and Herrero [FH], who had studied equations like (A) with $p \in (-1, 0]$, a less singular case than ours. This observation was useful, because now one could try to mimic blow-up results like the ones of Friedman and McLeod for quenching problems. And in fact, using techniques from [FM], Deng and Levine were able to show in 1988 that both statements of Theorem 1 could be extended from balls to convex domains Ω. A year later Fila and I found the blow up rate of $|u_t|$.

Time Asymptotics

Let u be a solution of (2) and suppose that u quenches at $t = T,\ x = 0$. Then the following estimates are known.

There exists a constant $c \geq 0$ such that for $t < T$

$$c \leq \min_{x\in\Omega}\{u(t,x)\}(T-t)^{-1/(1+p)} \leq (1+p)^{1/(1+p)}. \tag{3}$$

Moreover, if Ω is convex, $c > 0$, see [FK1]. Relation (3) holds for general $n \in \mathbb{N}$, $p > 0$ and Ω.

For any positive constant C and for $t < T$, $|x| \leq C(T-t)^{1/2}$ we have

$$\lim_{t\to T} u(t,x)(T-t)^{-1/(1+p)} = (1+p)^{1/(1+p)} \qquad . \tag{4}$$

This result was first established for $n = 1$ and $p \geq 3$ by Guo in 1988, and subsequently generalized to $n = 1$ and $p \geq 1$ by Fila and Hulshof, and to general n, nonnegative p and Ω a ball by Fila, Hulshof and Quittner.

Time asymptotics of this nature have been extended to equations of type (B) and (B'), see [FKL] or [SS]. It is also possible to extend such results to equation (C), see [KK].

How does one get the upper bound in (3)? This one is easy. In a spatial minimum we have $\Delta u \geq 0$, so there $u_t \leq -u^{-p}$, or equivalently

$$\frac{1}{p+1}(u^{p+1})_t = u^p \, u_t \leq -1. \tag{5}$$

An integration of (5) from t to T yields $0 - u^{p+1}(t) \leq -(p+1)(T-t)$, i.e the desired upper bound for u. To derive the lower bound in (3) one shows

$$u_t + \delta u^{-p} \leq 0$$

for some $\delta > 0$ and for (t, x) in some subcylinder of $(0, T) \times \Omega$. Here the idea of proof is essentially due to [FM].

SPACE ASYMPTOTICS AT $t = T$

Consider equation (A) and suppose that $\Omega \subset \mathbb{R}^n$ is a ball with center in the origin. For simplicity, suppose that $u(0, x) \equiv 1$. Then the following inequalities were derived in 1989 by M.Fila and myself, see [FK1].

$$u(T, r) \leq \left[\frac{(p+1)^2}{2(1-p)}\right]^{1/(1+p)} \left(r^2\right)^{1/(1+p)} \qquad \text{for } 0 < p < 1, \tag{6}$$

$$u(t, r) \quad \geq \quad C_\varepsilon \left(r^2\right)^{\varepsilon + 1/(1+p)} \qquad \text{for } 0 < p, \tag{7}$$

for $t < T$. These inequalities tell us, that for $p < 1$ the function $u(T, \cdot)$ is of class C^1 at the origin, while for $p > 1$ it has a cusp-singularity and is merely Hölder continuous in the origin, see Figure 2.

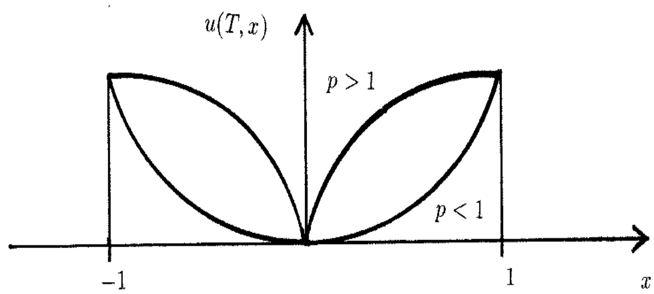

FIGURE 2: Shape of $u(\cdot, T)$

This distinction is consistent with the observation that $p < 1$ means less absorption than $p > 1$. Inequalities (6) and (7) are an ε apart. So the exact profile was still to be found. I had conjectured, but was unable to prove that for $p = 1$ the solution should develop a corner in the origin, like $u(r,T) \approx |r|$. I had been wrong, because in 1991 Filippas and Guo were able to find the exact asymptotics in the case $n = 1$ as follows

$$u(t,x) = \left[\frac{(p+1)^2}{8p}\right]^{1/(1+p)} \left(\frac{|x|^2}{|\ln|x||}\right)^{1/(1+p)} (1 + o(1)) \tag{8}$$

as $|x| \to 0$. This is definitely a sharper result. Again the method of proof relied on a corresponding blow up result, this time due to Herrero and Velazquez. In 1991 Fila, Levine and I generalized estimates (6) and (7) to equations (B) and (B'). In the context of equation (B') and for $p = 1$ one can interpret (7) as characterizing the rate at which the curvature of a rotational surface blows up. In fact, differential geometers like Huisken have found similar estimates for mean curvature flow in nonrotational settings as well.

Why is the assumption $p < 1$ made in (6)? To see this and to present another popular trick consider a solution u of the equation

$$u_t - \Delta u = -f(u)$$

and set

$$P(t,x) = \frac{1}{2}|\nabla u|^2 - F(u),$$

where $F'(u) = f(u)$. The letter P stands for L.Payne, who made this trick widely known, see [S]. A straightforward calculation shows that P satisfies the differential equation

$$P_t - \Delta P + b \cdot \nabla P \leq 0 \tag{9}$$

with $b = |\nabla u|^{-2}(2f(u)\nabla u - \nabla P)$ in $L^\infty_{loc}(\{(t,x) \mid 0 < t < T, |\nabla u(t,x)| \neq 0\})$. Now for $p < 1$ we have $P = -F(u) = -\frac{1}{1-p}u^{1-p} \leq 0$ in those points where $|\nabla u| = 0$. Thus, by the maximum principle, P must attain its maximum initially or on the lateral boundary of $[0,T] \times \Omega$. Since for convex Ω one can rule out that P attains its maximum on the boundary, and since $P(0,x) \leq 0$, we know $P(t,x) \leq 0$ or, in the case that Ω is a ball

$$u_r^2 \leq \frac{2}{1-p}u^{1-p}.$$

But now $u^{(p-1)/2}u_r \leq \sqrt{2/(1-p)}$ or

$$\frac{\partial}{\partial r}\left(u^{(p+1)/2}\right) = \left(\frac{1+p}{2}\right)u^{(p-1)/2}u_r \leq \left(\frac{1+p}{2}\right)\sqrt{2/(1-p)} = \frac{1+p}{\sqrt{2(1-p)}}.$$

An integration at $t = T$ yields

$$u^{(p+1)/2}(T,r) - u^{(p+1)/2}(T,0) \leq \left[\frac{1+p}{\sqrt{2(1-p)}}\right] r,$$

that is (6)

$$u(T,r) \leq r^{2/(1+p)} \left[\frac{1+p}{\sqrt{2(1-p)}} \right]^{2/(1+p)} .$$

LOCATION OF QUENCHING POINTS

Can one predict the points where a solution will quench? This question is related to the prediction of blow up points, and one of the early results on blow up stated single point blow up, see Weissler [W]. For the case of equation (A) and Ω being a ball, and under restrictions on the initial data, in 1987 Acker and I derived the inequalities

$$\begin{aligned} u_t &\leq 0 \\ ru_r = x \cdot \nabla u &\geq 0 \\ u_{rt} &\geq 0 \end{aligned}$$

in the parabolic time space cylinder $(0,T) \times \Omega$, and this implied that u quenches only in the origin, so one has a single point quenching result.

But more can sometimes be said for general initial data. In fact, for $n = 1$ and $p < 0$, Chen, Matano and Mimura have been able to derive finite point quenching results. They used lap-number type arguments to justify the occurrence of finitely many spatial oscillations of u after short time; and then they localized the above type of inequalities. This was nontrivial, because spatial minima of u can move in time.

Of course single point quenching results can also be shown for more general equations such as (B) or (B'), see [DK,FKL,AAG]. It is important to note though that in general it is necessary that $u_t \leq 0$ when u gets small.

LIFE AFTER QUENCHING

What happens after $t = T$ to a solution u of (1)? The answer depends on the notion of solution that we are willing to accept and on p. Suppose that the nonlinearity u^{-p} is regularized by the finite nonlinearity $u/(\varepsilon + u^{p+1})$. One can hope that then a classical global solution u_ε of (1) exists for every positive ε, that u_ε is decreasing in ε, and that it has a limit U as $\varepsilon \to 0$, which coincides with u for $t < T$. This hope has been replaced by a proof

a) in case of equation (A) and for $p < 1$ by D.Phillips [P], and
b) in case of equation (C) and for $p < m$, $m \geq 1$ in [KK].

In both cases there are regions in which $U = 0$, and in case a) the ω-limit set of U consists of equilibria or steady states, see [FLV,KK]. Moreover U is a global weak solution of

$$u_t - \Delta(u^m) = -u^{-p}\chi_{\{u>0\}},$$

for which uniqueness still appears to be open. So much for the case $0 < p < 1$.

If $p > 1$ and $n = 1$ I conjecture total quenching, that means I believe that $\lim_{\varepsilon \to 0} u_\varepsilon(x, T + \delta) = 0$ for every $x \in \Omega$ and every $\delta > 0$. The heuristic reason for this conjecture is the nonintegrability of u^{-p} as well as a corresponding total blow up result of Baras and Cohen. For $n \geq 2$ the situation is more complicated, see [FK1]. If Ω is convex, then

$$\int_\Omega u^{-q}\,dx \begin{cases} < \infty & \text{for } q < \frac{N}{2}(1+p) \\ = \infty & \text{for } q \geq \frac{N}{2}(1+p) \end{cases} \qquad \text{as } t \to T.$$

Another indication for complete quenching was kindly pointed out to the author by the referee. Using a transformation as in [KP], Galaktionov and Vazquez [GV1,GV2] converted the quenching problem to a blow-up problem. After deriving blow up results for the Cauchy problem and quasi-linear parabolic equations they were recently able to confirm my conjecture on total quenching for the Cauchy problem on $\mathbb{R}^n \times \mathbb{R}_+$ for equations of type (B) and (C).

If $p = 1$, little seems to be known for equations (A) and (C), but much is known for (B'). In fact, if $u(x,t)$ describes the radius of a compact rotational surface moving by mean curvature, then $u_x = \pm\infty$ on the boundary of the support of u. So near this boundary, the dependent and independent variable can be interchanged and the surface could also be described by a function $v(r)$, see Figure 3.

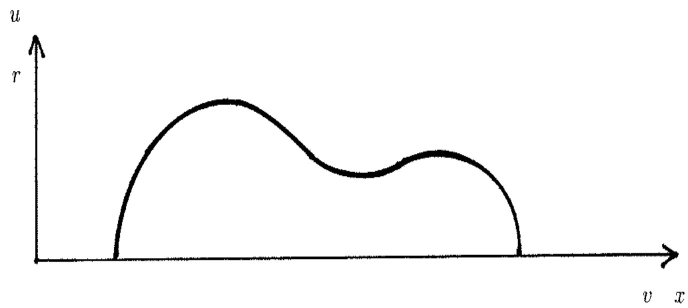

FIGURE 3: $u(t)$ and $v(r,t)$

The equation (B') for the horizontal graph

$$u_t = \frac{u_{xx}}{1+u_x^2} - \frac{1}{u} \tag{B'}$$

is then transformed into almost the same equation for the vertical graph, see [AAG],

$$v_t = \frac{v_{rr}}{1+v_r^2} + \frac{1}{r}v_r \quad . \tag{10}$$

So we have a parabolic equation, i.e. (B'), whose solution exhibits a hyperbolic phenomenon: finite speed of propagation of the free boundaries. This is reminiscent of phenomena described in [BD]. Nevertheless, equation (10) enables one to continue the analysis until the surface completely collapses. Eventually, it collapses into isolated points but u can stay non-concave in x until the time of collapse, see [AAG].

If one tries to apply the same transformation trick to equation (A), the outcome looks as follows

$$v_t = \frac{1}{v_r^2} v_{rr} + \frac{v_r}{r^p}. \tag{11}$$

Now (11) is totally different in nature from (10), because on the free boundary $r = 0$ we have $v_r = 0$, and so (11) reflects a very degenerate situation with "infinite" diffusion, while $v_r = 0$ causes no problems in the coefficients of (10). Again infinite diffusion seems to support the idea of total quenching mentioned above.

It is interesting to note that equation (11) can be rewritten in divergence form as

$$v_t = -\left(\frac{v_r}{|v_r|^2}\right)_r + \frac{v_r}{r^p},$$

and this in turn is equivalent to

$$v_t = -\mathrm{div}(|\nabla v|^{q-2} \nabla v) + \frac{v_r}{r^p} \quad \text{with } q = 0 \quad . \tag{12}$$

Now (12) looks like a backward heat equation with the Laplacian replaced by Δ_q for $q = 0$. Forward equations with Laplacian replaced by Δ_q and $q > 1$ are somewhat understood, see [EV], but (12) is far away from this situation. Therefore studying the operator Δ_q for $q = 0$ appears to be worthwhile and not just another academic exercise.

References

[AK] A.Acker & B.Kawohl: Remarks on quenching. Nonlinear Analysis, Theory, Methods and Applications, **13** (1989) p. 53–61

[AW] A.Acker & W.Walter: The quenching problem for nonlinear parabolic differential equations. In: *Ordinary and partial differential equations, Dundee 1976*, Eds. W.M.Everitt & B.D.Sleeman, Springer Lecture Notes in Math. **564** (1976) p. 1–12

[AAG] S.Altschuler, S.Angenent, Y.Giga: Mean curvature flow through singularities for surfaces of revolution. J. Geom. Anal. **5** (1995) p. 293–358

[BC] P.Baras & L.Cohen: Complete blow up after T_{max} for the solution of a semilinear heat equation. J. Funct. Anal. **71** (1987) p. 142–174

[BS] C.Bandle & I.Stakgold: The formation of the dead core in parabolic reaction diffusion problems. Trans. Amer. Math. Soc. **286** (1984) p. 275–293

[BD] M.Bertsch & R.Dal Passo: Hyperbolic phenomena in a strongly degenerate parabolic equation. Arch. Ration. Mech. Anal. **117** (1992) p. 349–387

[CK] C.Y.Chan & M.K.Kwong: Quenching phenomena for singular nonlinear parabolic equations. Nonlin. Anal. TMA **12** (1988) p. 1377–1383

[CMM] X.Chen, H.Matano & M.Mimura: Finite point extinction and continuity of interfaces in a nonlinear diffusion equation with strong absorption. J. reine angew. Math. **459** (1995) p. 1–36

[DK] G.Dziuk & B.Kawohl: On rotationally symmetric mean curvature flow, J. Differ. Equations, **93** (1991) p. 142–149

[EV] J.R. Esteban & J.L.Vazquez: Homogeneous diffusion in $\mathbb{R}$ with power-like non-linear diffusivity. Arch. Ration. Mech. Anal. **103** (1988) p. 39–80

[FH] M.Fila & J.Hulshof: A note on the quenching rate. Proc. Amer. Math. Soc. **112** (1991) p. 473–477

[FHQ] M.Fila, J.Hulshof & P.Quittner: The quenching problem on the N-dimensional ball. In: *Nonlinear Diffusion Equations and their Equilibrium States, 3.* Eds.: N.G.Lloyd, W.M.Ni, J.Serrin, L.A.Peletier, Birkhäuser, Boston, (1992) p. 183–196

[FLV] M.Fila, H.A.Levine & J.L.Vazquez: Stabilization of solutions of weakly singular quenching problems. Proc. Amer. Math. Soc., **119** (1993) p. 555–559

[FK1] M.Fila & B.Kawohl: Asymptotic analysis of quenching problems, Rocky Mt. J. Math. **22** (1992) p. 563–577.

[FK2] M.Fila & B.Kawohl: Is quenching in infinite time possible? Quarterly of Applied Math. **48** (1990) p. 531–534

[FKL] M.Fila, B.Kawohl & H.Levine: Quenching for quasilinear parabolic equations. Comm. Partial Differ. Equations **17** (1992) p. 593–614.

[FG] S.Filippas, J.S.Guo: Quenching profiles for one-dimensional semilinear heat equations. Quarterly of Applied Math. **51** (1993) p. 713–729

[FH] A.Friedman & M.Herrero: Extinction properties of semilinear heat equations with strong absorption. J.Math. Anal. Appl. **124** (1987) p. 530–547

[FM] A.Friedman & J.B.McLeod: Blow up of positive solutions of semilinear heat equations. Indiana Univ. Math. J. **35** (1985) p. 425–447

[GV1] V. Galaktionov & J.L.Vazquez: Necessary and sufficient conditions of complete blow-up and extinction for one-dimensional quasilinear heat equations. Arch. Ration. Mech. Anal. **129** (1995) p.225–244.

[GV2] V. Galaktionov & J.L.Vazquez: Continuation of blow-up solutions of nonlinear heat equations in several space dimensions, manuscript, 54 pp.

[G] J.Guo: On the quenching behaviour of the solution of a semilinear parabolic equation. J. Math. Anal. Appl. **151** (1990) p. 58–79.

[H] G.Huisken: Asymptotic behaviour for singularities of the mean curvature flow. J. Differ. Geometry **31** (1990) p. 285-299

[Ka] H.Kawarada: On solutions of initial boundary problem for $u_t = u_{xx} + 1/(1-u)$. Publ. Res. Inst. Math. Sci. **10** (1975) p. 729–736

[K] B.Kawohl: Remarks on quenching, blow up and dead cores. In: *Nonlinear Diffusion Equations and their Equilibrium States, 3.* Eds.: N.G.Lloyd, W.M.Ni, J.Serrin, L.A.Peletier, Birkhäuser, Boston, (1992) p. 275–286

[KK] B.Kawohl & R.Kersner: On degenerate diffusion with very strong absorption. Math. Methods Appl. Sci. **15** (1992) p. 469–477

[KP] B.Kawohl & L.Peletier: Observations on blow up and dead cores for nonlinear parabolic equations, Math. Zeitschr. **202** (1989) p. 207–217

[Le] H.Levine: Advances in quenching. in: *Nonlinear Diffusion Equations and their Equilibrium States, 3.* Eds.: N.G.Lloyd, W.M.Ni, J.Serrin, L.A.Peletier, Birkhäuser, Boston, (1992) p.319–346

[LM] H.A.Levine & J.T.Montgomery: Quenching of solutions of parabolic equations with nonlinear boundary conditions in several dimensions. J. reine angew. Math. **345** (1983) p. 23–28

[L] G.Lieberman: Quenching of solutions of evolution equations. Proc. Centre Math. Anal. Austral. Natl. Univ. **8** (1984), Miniconference on Nonlinear Analysis p. 151–157

[P] D.Phillips: Existence of solutions to a quenching problem. Appl. Anal. **24** (1987) p. 253-264

[SS] H.M.Soner & P.E.Souganidis: Singularities and uniqueness of cylindrically symmetric surfaces moving by mean curvature. Comm. PDE **18** (1993) p. 859–894

[S] R.Sperb: *Maximum principles and their applications,* Academic Press, New York 1981

[W] F.B.Weissler: Single point blow up for a semilinear initial value problem. J. Differ. Equations **55** (1984) p. 204-224

Bernd Kawohl
Mathematisches Institut
Universität zu Köln
D 50923 Köln
Germany
kawohl@mi.uni-koeln.de

On the Dimension of a Composition Algebra

Markus Rost

Received: December 23, 1995

Communicated by Ulf Rehmann

Abstract. The possible dimensions of a composition algebra are 1, 2, 4, or 8. We give a tensor categorical argument.

1991 Mathematics Subject Classification: Primary 17A75; Secondary 57M25.

I. Introduction

Let C be a composition algebra over a field of characteristic different from 2, let V be its pure subspace (consisting of the vectors orthogonal to 1) and let $d = \dim V$. We show that the following relation holds in the groundfield:

$$d(d-1)(d-3)(d-7) = 0.$$

This is not very surprising since the only possibilities for C are either the ground field, a separable quadratic extension, a quaternion algebra, or an octonion algebra. The proof of the relation given in this note seems to be different from former approaches (cf. [1], [2]). It works on a tensor categorical level. In characteristic 0 one recovers the determination of the possible dimensions of a composition algebra.

Our starting problem was to understand composition algebras from a tensor categorical point of view. Instead of composition algebras we looked at the equivalent notion of vector product algebras. These algebras can be obtained be rewriting the axioms of a composition algebra in terms of the pure vectors. Vector product algebras allow to use diagrammatic tensor calculus in a handy way. Using a graphical technique we found—just by playing around—a proof of the relation on $\dim V$. These notes contain alone the algebraic calculations which were extracted from the graph considerations. After these notes had been written, we noticed an identity in vector product algebras which perhaps makes the result less mysterious. So there is more to

say about the topic than explained in this text. We hope to come back to this at another place. Anyway, the text is completely self-contained and contains an argument on the possible dimensions.
Throughout the paper we assume char $\neq 2$.

ACKNOWLEDGEMENTS: I am indebted to B. Eckmann and T. A. Springer for useful comments. T. A. Springer suggested to use the relation (3.3) which reduced the amount of the calculations considerably. Moreover I thank the FIM at ETH Zürich for its hospitality.

II. COMPOSITION ALGEBRAS AND VECTOR PRODUCTS

We first recall a definition.

(1) COMPOSITION ALGEBRAS.
A composition algebra consists of a vector space C together with

(1.1) a nondegenerate symmetric bilinear form $\langle\, ,\rangle$ on C,

(1.2) a linear map $C \otimes C \to C, \quad x \otimes y \mapsto x \cdot y$,

(1.3) an element $0 \neq e \in C$,

such that (with $\mathrm{N}(x) = \langle x, x\rangle$)

(1.4) $e \cdot x = x \cdot e = x$,

(1.5) $\mathrm{N}(x \cdot y) = \mathrm{N}(x)\,\mathrm{N}(y)$.

For our purpose we have to consider the following algebraic structure.

(2) VECTOR PRODUCT ALGEBRAS.
A vector product algebra consists of a vector space V together with

(2.1) a nondegenerate symmetric bilinear form $\langle\, ,\rangle$ on V,

(2.2) a linear map $V \otimes V \to V, \quad x \otimes y \mapsto x \times y$,

such that

(2.3) $\langle x \times y, z\rangle$ is alternating in x, y, z,

(2.4) $(x \times y) \times x = \langle x, x\rangle y - \langle x, y\rangle x$.

The vector product $\times$ is anti-commutative, since (2.3) implies $x \times x = 0$. Therefore $x \times (y \times x) = (x \times y) \times x$. Hence the choice of the arrangement of the brackets in the lefthand side of (2.4) is not essential.

B. Eckmann has considered (continous) vector products in [B. Eckmann, Stetige Lösungen linearer Gleichungssysteme, Comment. Math. Helv. **15** (1942/43), 318–339], see also [B. Eckmann, Continous solutions of linear equations — An old problem, its history and its solution, Expo. Math. **9** (1991), 351–365]. He used the axioms

$$\langle x\times y, x\rangle = \langle x\times y, y\rangle = 0, \qquad N(x\times y) = \det \begin{vmatrix} \langle x,x\rangle & \langle x,y\rangle \\ \langle y,x\rangle & \langle y,y\rangle \end{vmatrix}.$$

They are perhaps more close to the intuitive idea of a vector product. Under presence of (2.1)–(2.2) they are easily seen to be equivalent to (2.3)–(2.4).

Vector product algebras and composition algebras are equivalent notions.

Namely, given a composition algebra C, let $V = \langle e\rangle^{\perp}$ and put

(i) $$x\times y = \frac{1}{2}(x\cdot y - y\cdot x).$$

Conversely, given a vector product algebra V, put $C = \langle e\rangle \perp V$ and define the product on C by

(ii) $$(ae+x)\cdot(be+y) = \big(ab - \langle x,y\rangle\big)e + ay + bx + x\times y.$$

The rewriting formulas (i) and (ii) identify composition algebras and vector product algebras on a “tensor categorical” level. This means that the composition rule (1.5) gives after polarization and decomposition with respect to $C = \langle e\rangle \perp V$ the same tensor equations as (2.3) and the polarization of (2.4).

This equivalence between composition algebras and vector product algebras seems to provide a convenient way to comprise some wellknown rules in composition algebras.

For the associator in C one finds

$$(x\cdot y)\cdot z - x\cdot(y\cdot z) = 2\big((x\times y)\times z - \langle x,z\rangle y + \langle y,z\rangle x\big)$$

for x, y, $z\in V$.

III. A Relation for the Contraction of $\langle\, ,\,\rangle$

Let V be a finite-dimensional vector product algebra and let $(e_i)_i$ be an orthonormal basis of V over some algebraic closure. Put

$$d = \sum_i \langle e_i, e_i\rangle.$$

(3) Proposition. *One has the relation*

$$d(d-1)(d-3)(d-7) = 0.$$

In the following we will tacitly apply (2.3) in the formulation

(2.3a) $\quad \langle x\times y, z\rangle = \langle x, y\times z\rangle,$

(2.3b) $\quad y\times x = -\,x\times y.$

The relation (2.4) will be used also in the following forms which are obtained by polarizing and from (2.3):

$$(x \times y) \times z + x \times (y \times z) = 2\langle x, z\rangle y - \langle x, y\rangle z - \langle z, y\rangle x, \tag{2.4a}$$

$$\begin{gathered}\langle x \times y, z \times t\rangle + \langle y \times z, t \times x\rangle = \\ 2\langle x, z\rangle\langle y, t\rangle - \langle x, y\rangle\langle z, t\rangle - \langle y, z\rangle\langle t, x\rangle.\end{gathered} \tag{2.4b}$$

Other relations to be used are

$$\sum_i e_i \times (v \times e_i) = \sum_i \langle e_i, e_i\rangle v - \sum_i \langle e_i, v\rangle e_i = dv - v = (d-1)v \tag{3.1}$$

and

$$\sum_{i,j} \langle e_i \times e_j, e_i \times e_j\rangle = \sum_{i,j} \big\langle e_i, e_j \times (e_i \times e_j)\big\rangle = (d-1)\sum_i \langle e_i, e_i\rangle = d(d-1). \tag{3.2}$$

To warm up, we first consider vector product algebras which correspond to associative composition algebras.

(4) Proposition. *Suppose that the following sharpening of* (2.4) *holds:*

$$(x \times y) \times z = \langle x, z\rangle y - \langle y, z\rangle x. \tag{4.1}$$

Then

$$d(d-1)(d-3) = 0.$$

Proof. Consider

$$A = \sum_{i,j,k} \big\langle e_i \times (e_k \times e_i), e_j \times (e_k \times e_j)\big\rangle.$$

By (3.1) we have

$$A = \sum_k (d-1)^2 \langle e_k, e_k\rangle = d(d-1)^2.$$

On the other hand, using (4.1) and (3.2) one finds

$$\begin{aligned} A &= \sum_{i,j,k} \Big\langle \big(e_i \times (e_k \times e_i)\big) \times e_j, e_k \times e_j \Big\rangle \\ &= \sum_{i,j,k} \Big\langle \langle e_i, e_j\rangle e_k \times e_i - \langle e_k \times e_i, e_j\rangle e_i, e_k \times e_j \Big\rangle \\ &= \sum_{i,k} \langle e_k \times e_i, e_k \times e_i\rangle - \sum_{i,j,k} \langle e_k \times e_i, e_j\rangle\langle e_i \times e_k, e_j\rangle \\ &= 2\sum_{i,k} \langle e_k \times e_i, e_k \times e_i\rangle = 2d(d-1). \end{aligned}$$

So

$$0 = A - A = d(d-1)(d-3). \qquad \square$$

Let us start with the proof of Proposition 3.

Put

$$h(u,v) = \sum_i (u \times e_i) \times (e_i \times v).$$

The following formula has been introduced by T. A. Springer.

$$h(u,v) = (d-4)\, u \times v. \tag{3.3}$$

To check it one uses (2.4a) with $x = u$, $y = e_i$ and $z = e_i \times v$ and finds

$$\begin{aligned} h(u,v) &= -\sum_i u \times \big(e_i \times (e_i \times v)\big) + 2\sum_i \langle u, e_i \times v\rangle e_i \\ &\quad - \sum_i \langle u, e_i\rangle e_i \times v - \sum_i \langle e_i \times v, e_i\rangle u \\ &= (d-1) u \times v + 2\sum_i \langle v \times u, e_i\rangle e_i \\ &\quad - u \times v - \sum_i \langle v, e_i \times e_i\rangle u \\ &= (d-1) u \times v - 2u \times v - u \times v - 0 = (d-4) u \times v. \end{aligned}$$

Formulas (3.3) and (3.2) make it easy to compute the sum

$$\begin{aligned} B &= \sum_{i,k} \big\langle h(e_i, e_k), h(e_k, e_i)\big\rangle \\ &= (d-4)^2 \sum_{i,k} \langle e_i \times e_k, e_k \times e_i\rangle = -d(d-1)(d-4)^2 \end{aligned}$$

We next compute B in a different way. One has

$$B = \sum_{i,j,k,l} \big\langle (e_i \times e_j) \times (e_j \times e_k), (e_k \times e_l) \times (e_l \times e_i)\big\rangle.$$

Applying (2.4b) shows

$$B + B' = 2C - D - D',$$

where

$$\begin{aligned} B' &= \sum_{i,j,k,l} \big\langle (e_j \times e_k) \times (e_k \times e_l), (e_l \times e_i) \times (e_i \times e_j)\big\rangle, \\ C &= \sum_{i,j,k,l} \langle e_i \times e_j, e_k \times e_l\rangle \langle e_j \times e_k, e_l \times e_i\rangle, \\ D &= \sum_{i,j,k,l} \langle e_i \times e_j, e_j \times e_k\rangle \langle e_k \times e_l, e_l \times e_i\rangle, \\ D' &= \sum_{i,j,k,l} \langle e_j \times e_k, e_k \times e_l\rangle \langle e_l \times e_i, e_i \times e_j\rangle. \end{aligned}$$

By reindexing one finds $B = B'$ and $D = D'$. Therefore

$$B = C - D.$$

We compute C and D:

$$\begin{aligned}
C &= \sum_{i,j,k,l} \big\langle e_i, e_j \times (e_k \times e_l)\big\rangle \big\langle (e_j \times e_k) \times e_l, e_i\big\rangle \\
&= \sum_{j,k,l} \big\langle e_j \times (e_k \times e_l), (e_j \times e_k) \times e_l\big\rangle \\
&= \sum_{j,k,l} \big\langle \big(e_j \times (e_k \times e_l)\big) \times (e_j \times e_k), e_l\big\rangle \\
&= -\sum_{k,l} \big\langle h(e_k \times e_l, e_k), e_l\big\rangle = -(d-4)\sum_{k,l}\big\langle (e_k \times e_l) \times e_k, e_l\big\rangle \\
&= -(d-1)(d-4)\sum_{l} \langle e_l, e_l\rangle = -d(d-1)(d-4),
\end{aligned}$$

$$\begin{aligned}
D &= \sum_{i,j,k,l} \big\langle e_i, e_j \times (e_j \times e_k)\big\rangle \big\langle (e_k \times e_l) \times e_l, e_i\big\rangle \\
&= \sum_{j,k,l} \big\langle e_j \times (e_j \times e_k), (e_k \times e_l) \times e_l\big\rangle \\
&= \sum_{k} (d-1)(d-1)\langle e_k, e_k\rangle = d(d-1)^2.
\end{aligned}$$

Hence

$$B = -d(d-1)(d-4) - d(d-1)^2 = -d(d-1)(2d-5).$$

Finally

$$\begin{aligned}
0 = B - B &= -d(d-1)(2d-5) + d(d-1)(d-4)^2 \\
&= d(d-1)(d^2 - 10d + 21) = d(d-1)(d-3)(d-7). \qquad \square
\end{aligned}$$

REFERENCES

[1] Hurwitz, A., Über die Komposition der quadratischen Formen, *Math. Ann.* 88, 1-25.

[2] Jacobson, N., Basic Algebra I, W. H. Freeman and Co., San Francisco, 1974.

Markus Rost
Fakultät für Mathematik
Universität Regensburg
Universitätsstraße 31
Germany
markus.rost@mathematik.uni-regensburg.de

Do Global Attractors Depend on Boundary Conditions?

Bernold Fiedler

Received: February 14, 1996

Communicated by Alfred K. Louis

Abstract. We consider global attractors of infinite dimensional dynamical systems given by dissipative partial differential equations

$$u_t = u_{xx} + f(x, u, u_x)$$

on the unit interval $0 < x < 1$ under separated, linear, dissipative boundary conditions. Global attractors are called orbit equivalent, if there exists a homeomorphism between them which maps orbits to orbits. The global attractor class is the set of all equivalence classes of global attractors arising for dissipative nonlinearities f. We show that the global attractor class does not depend on the choice of boundary conditions. In particular, Dirichlet and Neumann boundary conditions yield the same global attractor class.

The results are based on joint work with Carlos Rocha.

1 Attractor classes

Parabolic partial differential equations modelling reaction, diffusion, and drift are an important class of nonlinear infinite dimensional dynamical systems. Aside from applied motivation, much of the mathematical interest has centered on the dynamics of their finite dimensional global attractors. See for example [Hal88], [Lad91], [BV89], [Tem88], and the references there. The influence of boundary conditions has mainly been investigated in connection with stability of equilibria and shape of the underlying spatial domain, see for example [MM83], [Mat84].

Equations in one-dimensional domains have been studied in much detail, see for example [Cha74], [CI74], [Mat79], [Mat82], [Mat88], [Hen81], [Hen85], [Ang86], [Ang88], [BF88], [BF89], [AF88], [FMP89], [Nad90], [FP90]; mostly under Dirichlet or under Neumann boundary conditions, separately. In the present paper, we follow an approach developed more recently by [FR91], [Fie94], [FiRo94]. There the emphasis is on Neumann boundary conditions. Here, we indicate the necessary adaptations for general separated, linear, dissipative boundary conditions. Although the global attractors for a given nonlinearity will in general depend on our choice of boundary conditions, [HR87], we will show that the set of their orbit equivalence classes does not.

To be specific, we consider scalar equations

$$u_t = u_{xx} + f(x, u, u_x) \tag{1.1}$$

on the unit interval $0 \leq x \leq 1$. Fixing $0 \leq \tau_0, \tau_1 \leq 1$, we impose boundary conditions

$$(1 - \tau_\iota)u + \tau_\iota \partial_\nu u = 0 \tag{1.2}$$

at $x = 0, 1$. Here $\iota = 0,\ 1$; $\partial_\nu u = \pm u_x$ indicates the outward "normal" derivative with $+$ at $x = 1$, $-$ at $x = 0$, and subscripts t, x indicate partial derivatives of solutions $u = u(t, x)$. For the nonlinearities $f \in C^2$ we impose dissipation conditions

$$f(x, u, 0) \cdot u < 0, \tag{1.3}$$

for $|u| \geq C_1$ and, with continuous functions a, b as well as an exponent $\gamma < 2$

$$|f(x, u, p)| \leq a(u) + b(u)|p|^\gamma \tag{1.4}$$

at all arguments (x, u, p) of f. The estimators C_1, a, b, γ are allowed to depend on f. This setting is fixed throughout this paper.

The dissipation conditions (1.3), (1.4) guarantee the local semiflow of x-profiles of solutions $u(t, \cdot) \in X^\tau,\ t \geq 0$, to be globally defined and dissipative: any solution eventually remains in a fixed large ball $B \subseteq X^\tau$. See [Ama85], theorem 5.3 for a reference. In fact, we can choose B such that $|u| < C_1$ and $|p| < C_2$ on B. The state space X^τ is the Sobolev space H^2 intersected with boundary conditions (1.2), $\tau = (\tau_0, \tau_1)$.

By dissipativeness, equations (1.1), (1.2) possess a *global attractor* $\mathcal{A}_f^\tau \subseteq X^\tau$. This is the maximal compact invariant subset of X^τ or, here equivalently, the set of bounded solutions $u(t, \cdot),\ t \in \mathbb{R}$. Yes, including negative t. This global attractor is our principal object of study here. We call a global attractor $\mathcal{A}_f^\tau$ *orbit equivalent* to $\mathcal{A}_g^\sigma$,

$$\mathcal{A}_f^\tau \cong \mathcal{A}_g^\sigma, \tag{1.5}$$

if there exists a homeomorphism $H : \mathcal{A}_f^\tau \to \mathcal{A}_g^\sigma$ which maps orbits $\{u(t, \cdot) \,|\, t \in \mathbb{R}\}$ on $\mathcal{A}_f^\tau$ onto orbits in $\mathcal{A}_g^\sigma$. Obviously, $\cong$ is an equivalence relation and defines orbit equivalence classes of global attractors.

Let $\mathcal{E}_f^\tau$ denote the set of equilibrium solutions $u_t = 0$ of (1.1), (1.2). Clearly, $\mathcal{E}_f^\tau \subseteq \mathcal{A}_f^\tau$. We assume all equilibria to be hyperbolic: all eigenvalues of corresponding Sturm-Liouville eigenvalue problem are nonzero, for linearizations at equilibria. This is a generic nondegeneracy condition on f, for any given τ.

For given boundary conditions τ, we define the *attractor class* $\mathcal{A}(\tau)$ as the set of orbit equivalence classes of global attractors $\mathcal{A}_f^\tau$. Here $f \in C^2$ are assumed to be dissipative, as in (1.3), (1.4), with only hyperbolic equilibria.

Theorem 1.1 *In the above setting, the global attractor class $\mathcal{A}(\tau)$ does not depend on the boundary conditions (1.2) given by $\tau = (\tau_0, \tau_1) \in Q := [0,1]^2$. In other words, let $\tau, \sigma \in Q$. Then*

$$\mathcal{A}(\tau) = \mathcal{A}(\sigma). \tag{1.6}$$

Specifically, for any dissipative $f \in C^2$ with hyperbolic equilibria $\mathcal{E}_f^\tau$ there exists a dissipative $g \in C^2$, also with hyperbolic equilibria $\mathcal{E}_g^\sigma$, such that the respective global attractors $\mathcal{A}_f^\tau, \mathcal{A}_g^\sigma$ are orbit equivalent

$$\mathcal{A}_f^\tau \cong \mathcal{A}_g^\sigma, \tag{1.7}$$

in the sense of definition (1.5).

In section 2 we prove theorem 1.1. We conclude with a discussion of our result, in section 3.

For $\sigma, \tau \in (0,1]^2$, excluding the Dirichlet cases, the theorem is very easy to prove. We use a rescaling argument by Rafael Ortega. Let

$$u(x) = A(x)\, v(x) \tag{1.8}$$

with some smooth amplitude function $A > 0$ satisfying

$$\begin{array}{rcl} \tau_\iota A(\iota) & = & \sigma_\iota \\ (1-\tau_\iota)A(\iota) + \tau_\iota \delta_\nu A(\iota) & = & 1 - \sigma_\iota \end{array} \tag{1.9}$$

at $\iota = 0, 1$. Then the transformation (1.8) defines a linear isomorphism between the state spaces $u \in X^\tau$ and $v \in X^\sigma$ associated to boundary conditions τ and σ. Also, v satisfies an equation (1.1) with an appropriately rescaled dissipative nonlinearity g instead of f. Therefore $A_f^\tau \cong \mathcal{A}_g^\sigma$, by (1.8), (1.9), and $\mathcal{A}(\tau) = \mathcal{A}(\sigma)$ in the non-Dirichlet cases. (Strictly speaking, though, the transformation does not preserve the precise form (1.3) of our dissipation condition.) Our slightly more involved proof, given in section 2, will include even the Dirichlet case. In particular, the Neumann and the Dirichlet attractor classes will be shown to coincide. Note that all spaces X^τ, including the Dirichlet cases, are closed linear subspaces of $X = H^2$ depending continuously on the parameters τ; in particular all these spaces are isomorphic from an abstract view point.

We briefly outline the Morse-Smale structure behind our proof of theorem 1.1, in the remainder of the present section. Following [FiRo96], we first normalize f, for simplicity, such that

$$f(x, u, p) = -u \tag{1.10}$$

for $|x| \geq C_1$ or $|p| \geq C_2$. Such a normalization can be achieved without changing $\mathcal{A}_f^\tau$ or the flow on it, by dissipation conditions (1.3), (1.4). For $u \in X^\tau$ consider functionals

$$V(u) := \int_0^1 F(x, u, u_x) dx. \tag{1.11}$$

Following [Mat88] we observe that

$$\frac{d}{dt} V(u(t, \cdot)) = -\int_0^1 F_{pp} \cdot u_t^2 dx \tag{1.12}$$

along solutions $u(t,x)$ of (1.1), (1.2), if F satisfies

$$pF_{ppu} - fF_{ppp} + F_{ppx} = f_pF_{pp}, \tag{1.13}$$

for all (x,u,p) and obeys the boundary condition

$$F_p \cdot u_t = 0 \tag{1.14}$$

at $x = 0, 1$. By the standard method of characteristics, [Joh82], it is easy to find a solution $w = w(x,u,p)$ of the first order equation

$$pw_u - fw_p + w_x = f_p; \tag{1.15}$$

see also (2.1). In fact, normalization condition (1.10), guarantees global solvability of the characteristic equation of (1.15) which is studied in more detail in section 2 below. Solving then

$$F_{pp} = \exp(w) \tag{1.16}$$

we have solved (1.13). The boundary conditions (1.14) hold trivially in the Dirichlet case. We require $F_p = 0$, as an initial condition for (1.16) with respect to p, along the lines in (x,u,p)-space given by the boundary conditions (1.2), in all other cases.

By this construction, $F_{pp} = \exp(w)$ is positive. In particular the functional V becomes a Lyapunov functional on X^τ which decreases strictly along non-equilibrium orbits. With respect to the Riemannian metric on X^τ defined by F_{pp}, the semiflow (1.1), (1.2) is in fact *gradient*, or *Morse* with respect to V.

The functional V reveals that the global attractor $\mathcal{A}_f^\tau$ consists entirely of equilibria $\mathcal{E}_f^\tau$ and *heteroclinic* or *connecting* orbits. These orbits, by definition, limit onto (different) equilibria $\tilde{u}, u$ for $t \to +\infty$, $t \to -\infty$, respectively. They can be viewed as intersections of unstable and stable manifolds $W^u(u) \cap W^s(\tilde{u})$. Note that

$$\mathcal{A}_f^\tau = \mathcal{E}_f^\tau \cup \bigcup_{u \in \mathcal{E}_f^\tau} W^u(u). \tag{1.17}$$

Although this will not be very visible below, we emphasize the importance of *nodal properties* in our proof of theorem 1.1. Based on observations for linear equations, they imply that

$$t \mapsto z(u^1(t,\cdot) - u^2(t,\cdot)) \tag{1.18}$$

is nonincreasing along solutions $u^1(t,\cdot)$, $u^2(t,\cdot)$ of (1.1), (1.2). Here z, the zero number, denotes the number of strict sign changes of x-profiles. The zero number in (1.18) drops strictly whenever a multiple zero of the x-profile is encountered. Historically, the use of nodal properties dates back as far as [Stu36]. In [Mat82], their importance for infinite dimensional nonlinear dynamics was first realized. A comprehensive modern account of zero numbers is given in [Ang88].

The most striking consequence of nodal properties for our global attractors $\mathcal{A}_f^\tau$ is the *Morse-Smale property.* The Morse structure is generated by the Lyapunov functional V, as discussed above. By [Hen85], [Ang86], the intersections between stable and unstable manifolds which make up the global attractors $\mathcal{A}_f^\tau$ are automatically transverse, without further genericity assumption on f or τ:

$$W^u(u) \pitchfork W^s(\tilde{u}), \tag{1.19}$$

without any nondegeneracy assumptions except hyperbolicity of the equilibria $\mathcal{E}_f^\tau$.

Structural stability is the most important consequence of this Morse-Smale property. In fact, let g be C^2-near f and satisfy dissipation conditions (1.3), (1.4). Let also $\sigma \in Q = [0,1]^2$ be near τ. Then

$$\mathcal{A}_g^\sigma \cong \mathcal{A}_f^\tau, \tag{1.20}$$

as claimed in (1.7). For reference see [P69], [PS70], [PdM82], and for the infinite dimensional case [Oli92]. Since the argument is local in f, τ, this does not prove our theorem, of course.

2 Proof of theorem 1.1

If all equilibria $\mathcal{E}_f^\tau$ are hyperbolic, then the global attractor is Morse-Smale and therefore structurally stable, as we have seen at the end of the previous section. We give a geometric criterion for hyperbolicity of $\mathcal{E}_f^\tau$, in lemma 2.1. The criterion is based on a shooting approach to equilibria. In lemma 2.2, we relate global attractors for different boundary conditions, by an augmentation argument. Piecing these elements together, we finally prove theorem 1.1 by a homotopy argument which uses the Morse-Smale property.

Our geometric criterion for hyperbolicity is a slight adaptation of an argument in [Roc91]. Equilibria $u \in \mathcal{E}_f^\tau$ are solutions of

$$\begin{array}{rcl} \dot{u} &=& p \\ \dot{p} &=& -f(x,u,p) \\ \dot{x} &=& 1 \end{array} \tag{2.1}$$

which in addition satisfy the boundary conditions

$$\begin{array}{lll} l_0: & (1-\tau_0)u - \tau_0 p = 0, & \text{at } x = 0, \\ l_1: & (1-\tau_1)u + \tau_1 p = 0, & \text{at } x = 1. \end{array} \tag{2.2}$$

In passing we note that (2.1), together with $\dot{w} = f_p(x,u,p)$, are the equations of the characteristics of (1.15). For any real a, let $u(x,a),\ p(x,a)$ denote the solution of (2.1) with initial condition

$$\begin{array}{rcl} u(0,a) &:=& \tau_0 a \\ p(0,a) &:=& (1-\tau_0)a \end{array} \tag{2.3}$$

at $x = 0$. By normalization (1.10), these solutions are globally defined. Define the *shooting surface* $S_f^\tau \subseteq [0,1] \times \mathbb{R}^2$ as

$$S_f^\tau := \{(x,u,p) \mid u = u(x,a), p = p(x,a), a \in \mathbb{R}\}. \tag{2.4}$$

The sections $S_f^{\tau,x} \subseteq \mathbb{R}^2$ of S_f^τ for given x are called *shooting curves* . The shooting curves are planar C^1 Jordan curves, parametrized by the shooting parameter $a \in \mathbb{R}$. The set $\mathcal{E}_f^\tau$ of equilibria is given by precisely those values $a \in \mathbb{R}$ where the shooting curve $S_f^{\tau,x}$ at $x = 1$ intersects the line l_1 of boundary conditions at $x = 1$.

Lemma 2.1 *An equilibrium in $\mathcal{E}_f^\tau$ given by the shooting parameter a is hyperbolic if, and only if, the shooting curve $S_f^{\tau,x=1}$ intersects the target line l_1 transversely, at the intersection value a.*

Proof: Consider the equilibrium $u(x,a)$ corresponding to the intersection value a, or vice versa. The partial derivative $(u_a(x,a), p_a(x,a))$ is *the* nontrivial solution of the linearized equation which satisfies the homogeneous linear boundary condition l_0. (Well, we could take constant multiples instead.) Clearly, u is nonhyperbolic if, and only if, this partial derivative also satisfies the other boundary condition l_1, at $x = 1$. Reinterpreting geometrically, nonhyperbolicity is then equivalent to the tangent vector of the shooting curve $S_f^{\tau,x=1}: \ a \mapsto (u(1,a), p(1,a))$ being parallel to the line l_1. This is exactly nontransversality of intersection, and the lemma is proved. $\square$

Augmentation works as follows. We append new segments

$$\begin{array}{rcl} I_0 & := & [-\xi_0, 0) \\ I_1 & := & (1, 1+\xi_1] \end{array} \tag{2.5}$$

$\xi_0, \xi_1 > 0$, to the original x-interval $x \in [0,1]$. In the appended intervals, we define f by

$$f(x,u,p) := -\lambda^2(x)u, \tag{2.6}$$

where $\lambda^2(x) := \lambda_\iota^2 > 0$ is constant for $x \in I_\iota$. We will specify $\lambda = (\lambda_0, \lambda_1)$ below. In I_ι the shooting equation (2.1) becomes the hyperbolic linear equation

$$\begin{array}{rcl} \dot{u} & = & p \\ \dot{p} & = & \lambda_\iota^2 u \end{array} \tag{2.7}$$

In (u,p)-space, this linear equation induces a flow on the lines of boundary conditions

$$l(\tau_\iota): \quad (1-\tau_\iota)u \pm \tau_\iota p = 0. \tag{2.8}$$

Here $\iota = 1$ carries the plus-sign, whereas $\iota = 0$ requires a minus. The boundary condition parameters $\tau_\iota(x)$ are now considered to depend on $x \in I_\iota$, with their values at $x = \iota$ taken from the original boundary conditions (1.2). The flow induced by (2.7), (2.8) on τ_ι is

$$\dot{\tau}_\iota = \pm(\lambda_\iota^2 \tau_\iota^2 - (1-\tau_\iota)^2), \tag{2.9}$$

by direct calculation. This equation plays a central role in the proof of the following lemma.

For abbreviation, let $\mathcal{F}$ denote the set of dissipative nonlinearities $f, g \in C^2$ with hyperbolic equilibria as specified in theorem 1.1.

Lemma 2.2 *Let $f \in \mathcal{F}$ and consider arbitrary boundary conditions $\tau \in (0,1)^2$, in the interior of the closed unit square $Q = [0,1]^2$, and $\sigma \in Q$. Then there exists $g \in \mathcal{F}$ such that the global attractors $\mathcal{A}_f^\tau, \mathcal{A}_g^\sigma$ are orbit equivalent,*

$$\mathcal{A}_f^\tau \cong \mathcal{A}_g^\sigma \tag{2.10}$$

Proof: We will use augmentation (2.5)–(2.9) to construct g on an interval $x \in [-\xi_0, 1+\xi_1]$. Rescaling x by a scaling factor $s := 1/(1+\xi_0+\xi_1)$ and adjusting σ accordingly, first, we will not lose generality. As a main step, we will then construct a dissipative homotopy from (f,τ) to (g,σ), by augmentation, such that hyperbolicity of equilibria is preserved throughout the homotopy. In a final, third step we address the issue of C^2-regularization of our piecewise defined nonlinearities, by smoothing. By Morse-Smale structural stability, the homotopy which preserves hyperbolicity of equilibria then proves the lemma.

Rescaling x to $\tilde{x}$ by $x = s(\tilde{x}+\xi)$ transforms the x-interval $[0,1]$ to an $\tilde{x}$-interval $[-\xi, 1+\xi]$, if we choose $s = (1+2\xi)^{-1} \in (0,1)$. Simultaneously, boundary conditions $\sigma = (\sigma_0, \sigma_1)$ at $x = 0,1$ get transformed to boundary conditions $\tilde{\sigma} = (\tilde{\sigma}_0, \tilde{\sigma}_1)$, for $\tilde{u}(t,\tilde{x}) := u(t,x)$, which are given explicitly by

$$\tilde{\sigma}_\iota = \frac{s\sigma_\iota}{1-\sigma_\iota+s\sigma_\iota}. \tag{2.11}$$

Note that $\tilde{\sigma}_\iota = 0,1$ for $\sigma_\iota = 0,1$, respectively.

We consider the case $0 \le \sigma_\iota < 1$ first. Fix $\xi = \bar{\xi} > 0$ large enough or, equivalently, $s = (1+2\bar{\xi})^{-1}$ small enough, such that in particular

$$0 \le \tilde{\sigma}_\iota < \tau_\iota < 1. \tag{2.12}$$

Now consider the τ_ι flow (2.9) in an equation which is augmented according to (2.6). We choose $\lambda_\iota > 0, \iota = 0,1$, such that the time which the τ_ι flow (2.9) takes from τ_ι to $\tilde{\sigma}_\iota$ coincides with the large prescribed value $\bar{\xi}$:

$$\tilde{\sigma}_\iota = \tau_\iota(\pm\bar{\xi}) \tag{2.13}$$

for the initial values $\tau_\iota(0) = \tau_\iota$ at $x = \iota$. Indeed this can be achieved by choosing $\lambda_\iota > 0$ such that the unique equilibrium

$$\tau_\iota^* = (1+\lambda_\iota)^{-1} \tag{2.14}$$

of (2.9) in $(0,1)$ is slightly above $\tau_\iota < 1$.

In the remaining Neumann case $\tilde{\sigma}_\iota = \sigma_\iota = 1$, we simply choose $\tau_\iota^* > 0$ slightly below $\tau_\iota > 0$, and (2.13) remains valid.

We describe our homotopy of attractors in terms of changing the boundaries $x = -\xi, 1+\xi$, simultaneously, from their original value $\xi = 0$ to their final values $\xi = \bar{\xi}$. On these larger x-intervals the nonlinearity f is augmented to f^ξ by (2.6). The boundary conditions $\tau = \tau(\xi)$ are adjusted, according to (2.9), in parallel with the homotopy parameter ξ. Note that by a rescaling of x with factor $s = 1/(1+2\xi)$, this induces a homotopy of global attractors for rescaled nonlinearities in the class $\mathcal{F}$. Clearly, dissipativeness is preserved. In view of Morse-Smale structural stability, it therefore only remains to prove that hyperbolicity of equilibria is preserved throughout the homotopy.

Hyperbolicity of equilibria follows from lemma 2.1. Indeed, transversality of the shooting curve $S_{f^\xi}^{\tau(\xi),x}$, at $x = 1+\xi$, to the line $l(\tau_1(\xi))$ follows in three steps, using (2.5)–(2.9). First, in $I_0 = [-\xi, 0)$, the initial line $l(\tilde{\sigma}_0) = l(\tau_0(\xi)) = S_{f^\xi}^{\tau(\xi),-\xi}$ gets mapped diffeomorphically to the line $l(\tau_0(0)) = l(\tau_0) = l_0$. Second, in $x \in [0,1]$, we obtain the original f-shooting curve

$$S_f^{\tau,1} \pitchfork l(\tau_1). \tag{2.15}$$

Here we use hyperbolicity of $\mathcal{E}_f^\tau$ and lemma 2.1. Third, in $I_1 = (1, 1+\xi]$, the line $l_1 = l(\tau_1) = l(\tau_1(0))$ and the f-shooting curve $S_f^{\tau,1}$ get mapped onto

$$S_{f^\xi}^{\tau(\xi),1+\xi} \quad \pitchfork l(\tilde{\sigma}_1), \tag{2.16}$$

by the shooting diffeomorphism. Transversality is inherited from (2.15). A final application of lemma 2.1 proves that hyperbolicity of equilibria is preserved during our homotopy $0 \leq \xi \leq \bar{\xi}$. Of course, rescaling of x does not affect hyperbolicity.

Smoothing the discontinuities of our augmentation of f, at $x = 0, 1$, we obtain a C^2-augmentation. Making the x-intervals, where smoothing acts, small enough, we can guarantee transversality (2.16) to hold throughout our homotopy $0 \leq \xi \leq \bar{\xi}$. In particular, all f^ξ are Morse-Smale. Defining g as (the rescaled version of) $f^{\bar{\xi}}$, structural stability of Morse-Smale systems finally implies

$$\mathcal{A}_f^\tau \cong \mathcal{A}_g^\sigma. \tag{2.17}$$

In (2.17) we have used that rescaling does not change the orbit type of the global attractor and, simultaneously, transforms $\tilde{\sigma} = \tau(\bar{\xi})$ to the boundary condition σ by (2.11). This proves the lemma. □

In the previous lemma we have shown $\mathcal{A}(\tau) = \mathcal{A}(\sigma)$, for attractor classes with $\tau, \sigma \in (0,1)^2$. (The transformation (1.8) would even allow for $\tau, \sigma \in (0,1]^2$.) To complete the proof of theorem 1.1, anyway, it remains to address the case of τ or σ in the boundary ∂Q of the square $Q = [0,1]$. If $g \in \mathcal{F}, \sigma \in \partial Q$, then local structural stability of Morse Smale systems shows that for $f := g$ and any $\tau \in (0,1)^2$ close to σ we have orbit equivalence $\mathcal{A}_g^\sigma \cong \mathcal{A}_f^\tau$. Therefore $\mathcal{A}(\sigma) \subseteq \mathcal{A}(\tau)$. To complete the proof of theorem 1.1 it remains to show that, conversely,

$$\mathcal{A}(\sigma) \supseteq \mathcal{A}(\tau), \tag{2.18}$$

for some $\tau \in (0,1)^2$. By lemma 2.2, claim (2.18) actually holds for all $\tau \in (0,1)^2, \sigma \in Q$. This completes our proof of theorem 1.1. □

3 Discussion

We begin our discussion with remarks on x-dependent diffusion and on another attempt of simplifying our proof, by transformation of x. We then indicate why periodic boundary conditions $x \in S^1$ produce a class of Morse-Smale attractors quite different from the class $\mathcal{A}(\tau)$ of separated boundary conditions $\tau = (\tau_0, \tau_1) \in Q = [0,1]^2$. We conclude with a few comments on global attractors in the case of higher space dimension, dim $x > 1$, and the case of systems, dim $u > 1$.

Transforming x to $y = \eta(x) \in [0,1]$ in (1.1), (1.2) and denoting $v(y) := u(x)$ yields an equation

$$D(y)^{-1}\, v_t = v_{yy} + g(y, v, v_y) \tag{3.1}$$

with transformed boundary conditions

$$(1-\sigma_\iota)v \pm \sigma_\iota\, v_y = 0 \tag{3.2}$$

at $y = \iota = 0, 1$. Explicitly, we have

$$\begin{array}{rcl} D(y) & = & (\eta_x(x))^2 > 0, \\ \sigma_\iota & = & (1 + \frac{1-\tau_\iota}{\tau_\iota}\eta_x)^{-1}. \end{array} \tag{3.3}$$

Given $\tau \in (0,1)^2$ we can clearly reach all $\sigma \in (0,1)^2$ by a proper choice of the function η. The standard linear homotopy from $D(y)^{-1}v_t$ to v_t, in (3.1), is a Morse-Smale homotopy of attractors which does not change the equilibria. Indeed, the shooting surface never changes, during the homotopy, because $D(y)^{-1}$ only multiplies the time derivative. Therefore we conclude $\mathcal{A}_g^\sigma \cong \mathcal{A}_f^\tau$, as stated in lemma 2.2.

A main disadvantage of this rather simple argument is the fact that Neumann as well as Dirichlet boundary conditions $\tau_\iota = 0, 1$ remain unchanged by the transformation η; see (3.3). It is the case $\sigma \in \partial Q$, where we really seem to need the augmentation in lemma 2.2.

Of course we could have discussed orbit equivalence of attractors in the class of pairs (f, D), allowing for space–dependent diffusion from the very start. Fixing $D \equiv 1$, though, provides a stronger statement in theorem 1.1. Parenthetically we note that introducing $D > 0$ does not produce any additional global attractors, by the above arguments. As we have argued in the discussion section of [FiRo96], we do not expect additional global attractors to arise, even in fully nonlinear, uniformly parabolic, dissipative cases.

Passing to higher-dimensional domains $x \in \Omega \subseteq \mathbb{R}^d$, with $\partial\Omega$ smooth and bounded, we may again consider dissipative scalar equations

$$u_t = \Delta u + f(x, u, \nabla u) \tag{3.4}$$

on Ω, under mixed boundary conditions

$$(1-\tau)u + \tau\partial_\nu u = 0. \tag{3.5}$$

Now $\tau = \tau(x) \in [0,1]$ is a given function on $\partial\Omega$. A transformation $u(x) = A(x)v(x)$ is still feasible, normalizing $\tau \in (0,1]$ to become a uniform Neumann condition $\sigma \equiv 1$ for v; see (1.8), (1.9). But we have lost variational structure, nodal properties, and Morse-Smale when passing to (3.4), (3.5). Essentially arbitrary finite-dimensional flows occur in (3.4), see [Pol95]. Even if we assume the global attractor $\mathcal{A}_f^\tau$ to be structurally stable, there is no reason to believe that its orbit equivalence class is determined by the equilibria, alone.

To include the Dirichlet cases, it is tempting to try and augment Ω, by attaching a collar outside $\partial\Omega$, such that boundary conditions on the enlarged region differ from the original ones. A structurally stable attractor $\mathcal{A}_f^{\sigma,\Omega}$ should still be recovered in the enlarged region Ω'. If Ω is starshaped with respect to the origin, a homothety

$\Omega' = s\Omega$ by a scaling factor $s > 1$ comes to mind. In the annular region $A = \Omega' \setminus \Omega$ we can determine an eigenfunction for a positive eigenvalue λ of the Laplacian Δ with boundary conditions

$$\begin{aligned} (1-\tau)u - \tau\partial_{\nu'}u &= 0 \quad \text{on} \quad \partial\Omega \\ (1-\sigma)u + \sigma\partial_{\nu'}u &= 0 \quad \text{on} \quad \partial\Omega'. \end{aligned} \tag{3.6}$$

Here ν' denotes the outward normal of A; on $\partial\Omega$ we have $\nu' = -\nu$. In the one-dimensional case, this eigenfunction was the crucial shooting augmentation in the "annulus" $A = I_0 \cup I_1$. The task remains open to augment the PDE (3.4) in A in such a "singular way" that the original attractor $\mathcal{A}_f^{\tau,\Omega}$ is recovered on Ω' with new boundary conditions σ. For systems, $u \in \mathbb{R}^k$, a similar problem arises. Even in the case of one-dimensional x, though, is is not yet clear how to properly recover $\mathcal{A}_f^\tau$ on the enlarged interval Ω' then.

Jacobi systems are the spatially discrete ODE analogue to our scalar PDE (1.1), (1.2) in one space dimension; see [FO88]. Specifically, Jacobi systems have the tri-diagonal nonlinear form

$$\dot{u}_i = f_i(u_{i-1}, u_i, u_{i+1}), \tag{3.7}$$

$i = 0, \ldots, n$, with strictly positive partial derivatives of the nonlinearities f_i with respect to the off-diagonal entries u_{i-1}, u_{i+1}. For convenience we impose linear boundary conditions in the following form

$$\begin{aligned} (1+\tau_0)u_{-1} - 2\tau_0 u_0 &= 0 \\ (1+\tau_1)u_{n+1} - 2\tau_1 u_n &= 0. \end{aligned} \tag{3.8}$$

System (3.7), (3.8) may, but need not, arise by finite difference semidiscretization in space of (1.1), (1.2). Then $\tau_\iota = 1$, $\iota = 0, 1$ corresponds to Neumann boundary conditions, as before, and $\tau_\iota = 0$ are Dirichlet conditions

$$u_{-1} = u_{n+1} = 0. \tag{3.9}$$

Only boundary conditions $0 \le \tau_\iota \le 1$ arise by discretization of dissipatively admissible PDE boundary conditions. Note, however, that the choice $\tau_\iota = -1$ again corresponds to Dirichlet boundary conditions

$$u_0 = u_n = 0, \tag{3.10}$$

at least formally.

For $\tau_\iota \neq -1$, the state space of our system (3.7), (3.8) is $u = (u_0, \cdots, u_n) \in X = \mathbb{R}^{n+1}$. A natural dissipation condition is

$$u_i \cdot f_i(u_i, u_i, u_i) < 0 \tag{3.11}$$

for all $i = 0, \cdots, n$, provided $|u_i| \ge C$. Here C is a large constant. Under boundary conditions (3.8) with

$$|\tau_0|, |\tau_1| \le 1, \tag{3.12}$$

condition (3.11) ensures that $\|u\| := \max |u_i|$ decreases to level C or below, eventually. If τ_0 or τ_1 violate condition (3.12), then $\max |u_i|$ may grow indefinitely on the boundary, in spite of dissipation condition (3.11). Therefore we restrict attention to the region (3.12).

For Neumann condition $\tau_\iota = 1$, it was argued in [FiRo96], theorem 8.2, that the attractor class $\mathcal{A}^{dis}(\tau)$ for Jacobi systems (3.7), (3.8) coincides with the PDE attractor class $\mathcal{A}^{con}(\tau) := \mathcal{A}(\tau)$ of our present theorem 1.1. Here $\mathcal{A}^{dis}(\tau)$ ranges over all Jacobi systems of any dimension n. For $0 \le \tau_\iota \le 1$, we expect similar arguments to provide

$$\mathcal{A}^{dis}(\tau) = \mathcal{A}^{con}(\tau) \tag{3.13}$$

to be τ-independent, by theorem 1.1.

Augmentation to $i \in \{-2, -1, \cdots, n+1, n+2\}$ also seems a viable approach to discrete attractor classes. Consider the new left boundary u_{-2}, u_{-1}, for example. Comparing a new boundary condition $\sigma = (\sigma_0, \sigma_1)$ with the old τ-condition (3.8), at the left end, we obtain

$$\begin{array}{rcl} (1+\tau_0)u_{-1} - 2\tau_0 u_0 & = & 0 \\ (1+\sigma_0)u_{-2} - 2\sigma_0 u_{-1} & = & 0. \end{array} \tag{3.14}$$

Adding the two equations with real coefficients $-\beta, \alpha$ we obtain the right hand side of an augmentation

$$\dot{u}_{-1} = \alpha(1+\sigma_0)u_{-2} - (2\alpha\sigma_0 + \beta(1+\tau_0))u_{-1} + 2\beta\tau_0 u_0. \tag{3.15}$$

This augmentation is Jacobi and dissipative, for $|\sigma_\iota|, |\tau_\iota| \le 1$, if

$$\begin{array}{rcl} \alpha & > & 0 \\ \beta\tau_0 & > & 0 \\ \alpha(1-\sigma_0) & < & \beta(1-\tau_0) \end{array} \tag{3.16}$$

Note that equilibrium shooting, $\dot{u}_{-1} \equiv 0$, maps the σ_0 boundary condition to the τ_0 condition under our choice (3.15) of augmentation.

Let $\mathcal{A}^n(\tau)$ denote the attractor class for Jacobi systems (3.7), (3.8), this time with fixed dimension $n+1$. In view of theorem 1.1 and (3.13) it seems natural to ask whether $\mathcal{A}^n(\tau)$ can be independent of τ, at least for $0 \le \tau_\iota \le 1$. More daringly: let $\mathcal{A}_n^{con}$ denote the set of attractor classes in $\mathcal{A}^{con}$ of dimension at most $n+1$. Is it true, for $0 \le \tau_\iota \le 1$ and at least for large n, that

$$\mathcal{A}^n(\tau) = \mathcal{A}_n^{con}? \tag{3.17}$$

In particular $\mathcal{A}^n(\tau)$ would not depend on τ, of course.

Transforming the boundary value at $i = -1$ by $s\tilde{u}_{-1} := u_{-1}$ requires $0 < s < 1$ to remain in the class of dissipative Jacobi systems where (3.11) holds. For $0 < |\tau_0| \le |\sigma_0| \le 1$ of equal sign we obtain an embedding

$$\mathcal{A}^n(\tau_0, \tau_1) \subseteq \mathcal{A}^n(\sigma_0, \sigma_1), \tag{3.18}$$

which does not quite answer our question. Dissipative Jacobi augmentation (3.15), (3.16) does not provide an answer, either. Some modest conclusions are

$$\begin{array}{c} \mathcal{A}^n(\tau_0, \tau_1) \subseteq \mathcal{A}^{n+1}(1, \tau_1) \\ \mathcal{A}^n(\tau_0, \tau_1) \subseteq \mathcal{A}^{n+1}(\sigma_0, \tau_1). \end{array} \tag{3.19}$$

Again $|\tau_\iota|, |\sigma_\iota| \leq 1$. In addition we require $\tau_0 \neq 0$ and, in the second equation, $\tau_0 \neq 1$. Aside from these constraints, τ and σ are arbitrary. Replacing i by $n - i$ we also observe symmetry for all τ,

$$\mathcal{A}^n(\tau_0, \tau_1) = \mathcal{A}^n(\tau_1, \tau_0). \tag{3.20}$$

For example, this implies Neumann embedding

$$\mathcal{A}^n(\tau_0, \tau_1) \subseteq \mathcal{A}^{n+2}(1, 1), \tag{3.21}$$

for $\tau_0, \tau_1 \neq 1$. Similarly, for $\tau_0, \tau_1 \neq 0, 1$ and all σ we obtain

$$\mathcal{A}^n(\tau) \subseteq \mathcal{A}^{n+2}(\sigma) \tag{3.22}$$

from (3.19), (3.20).

Note that independence of $\mathcal{A}^{n+1}(\sigma)$ from σ might break down, at least for $\sigma_0 \searrow -1$. In that case, the boundary condition (3.14) collapses to $u_{-1} = 0$, formally. This is equivalent to the Dirichlet attractor class $\mathcal{A}^n(0, \sigma_1)$ of Jacobi systems in one lower dimension.

As a final remark, we emphasize that periodic boundary conditions $x \in S^1$ generate sets $\mathcal{A}^{con}(per)$, $\mathcal{A}^n(per)$ of attractor classes which are much richer than their colleagues $\mathcal{A}^{con}(\tau) = \mathcal{A}^{con}(sep)$ living in separated boundary conditions. In fact, the Neumann class can be shown to be contained in the periodic class $\tau = (1, 1)$, by reflection through the boundary and smoothing:

$$\mathcal{A}^{con}(sep) \subset \mathcal{A}^{con}(per), \tag{3.23}$$

again by theorem 1.1. As remarked in [AF88], even for nonlinearities $f = f(u, p)$ independent of x, time periodic rotating waves can arise in $\mathcal{A}^{con}(per)$, which simply do not possess any counterpart in the gradient case $\mathcal{A}^{con}(sep)$. In particular, Lyapunov functionals like V fail. A similar remark applies to the spatially discrete case $\mathcal{A}^n(per)$ of cyclic Jacobi systems $i \pmod{(n+1)}$. Since reflection through the boundary for Neumann condition yields only an embedding

$$\mathcal{A}^n(1, 1) \not\subseteq \mathcal{A}^{2n+1}(per),$$

the characterization of attractor classes in the case of periodic boundary conditions remains wide open.

REFERENCES

[Ama85] H. Amann. Global existence for semilinear parabolic systems. *J. Reine Angew. Math.*, 360: 47–83, (1985).

[AF88] S. Angenent and B. Fiedler. The dynamics of rotating waves in scalar reaction diffusion equations. *Trans. Amer. Math. Soc.*, 307: 545–568, (1988).

[Ang86] S. Angenent. The Morse-Smale property for a semi-linear parabolic equation. *J. Diff. Eq.*, 62: 427–442, (1986).

[Ang88] S. Angenent. The zero set of a solution of a parabolic equation. *J. reine angew. Math.*, 390: 79–96, (1988).

[BF88] P. Brunovský and B. Fiedler. Connecting orbits in scalar reaction diffusion equations. *Dynamics Reported*, 1: 57–89, (1988).

[BF89] P. Brunovský and B. Fiedler. Connecting orbits in scalar reaction diffusion equations II: The complete solution. *J. Diff. Eq.*, 81: 106–135, (1989).

[BV89] A.V. Babin and M.I. Vishik. *Attractors in Evolutionary Equations.* Nauka, Moscow 1989.

[Cha74] N. Chafee. A stability analysis for a semilinear parabolic partial equation. *J. Differential Equations*, 15: 522–540, (1974).

[CI74] N. Chafee and E. Infante. A bifurcation problem for a nonlinear parabolic equation. *J. Applicable Analysis*, 4: 17–37, (1974).

[Fie94] B. Fiedler. Global attractors of one-dimensional parabolic equations: sixteen examples. *Tatra Mountains Math. Publ.*, 4: 67–92, (1994).

[FMP89] B. Fiedler and J. Mallet-Paret. Connections between Morse sets for delay-differential equations. *J. reine angew. Math.*, 397: 23–41, (1989).

[FO88] G. Fusco and W.M. Oliva. Jacobi matrices and transversality. *Proc. Royal Soc. Edinburgh* A, 109: 231–243, (1988).

[FP90] B. Fiedler and P. Poláčik. Complicated dynamics of scalar reaction diffusion equations with a nonlocal term. *Proc. Royal Soc. Edinburgh* A, 115: 167–192, (1990).

[FR91] G. Fusco and C. Rocha. A permutation related to the dynamics of a scalar parabolic PDE. *J. Diff. Eq.*, 91: 75–94, (1991).

[FiRo94] B. Fiedler and C. Rocha. Heteroclinic orbits of semilinear parabolic equations. Preprint, to appear in J. Diff. Eqs., (1994).

[FiRo96] B. Fiedler and C. Rocha. Orbit equivalence of global attractors of semilinear parabolic differential equations. Preprint, (1996).

[Hal88] J.K. Hale. *Asymptotic Behavior of Dissipative Systems.* Math. Surv. 25. AMS Publications, Providence, 1988.

[Hen81] D. Henry. *Geometric Theory of Semilinear Parabolic Equations.* Lect. Notes Math. 840. Springer-Verlag, New York, 1981.

[Hen85] D. Henry. Some infinite dimensional Morse-Smale systems defined by parabolic differential equations. *J. Diff. Eq.*, 59: 165–205, (1985).

[HR87] J. Hale and C. Rocha. Interaction of diffusion and boundary conditions. *Nonlinear Analysis*: 11: 633–649, (1987).

[Joh82] F. John. *Partial Differential Equations.* Springer-Verlag, New York, 1982.

[Lad91] O.A. Ladyzhenskaya. *Attractors for Semi-groups and Evolution Equations.* Cambridge University Press, 1991.

[Mat79] H. Matano. Asymptotic behavior and stability solutions of semilinear diffusion equations. *Publ. Res. Inst. Math. Sc. Kyoto Univ.*, 15: 401–454, (1979).

[Mat82] H. Matano. Nonincrease of the lap-number of a solution for a one-dimensional semi-linear parabolic equation. *J. Fac. Sci. Univ. Tokyo Sec. IA*, 29: 401–441, (1982).

[Mat84] H. Matano. Existence of nontrivial unstable sets for equilibriums of strongly order-preserving systems. *J. Fac. Sci. Univ. Tokyo Sec. IA-Math.*, 3: 645–673, (1984).

[Mat88] H. Matano. Asymptotic behavior of solutions of semilinear heat equations on S^1. In *Nonlinear Diffusion Equations and their Equilibrium States II.* Springer Verlag, New York, 1988.

[MM83] H. Matano and M. Mimura. Pattern formation in competition-diffusion systems in nonconvex domains. *Publ. Res. Inst. Math. Sci.*, 19: 1049–1079, (1983).

[Nad90] N.S. Nadirashvili. On the dynamics of nonlinear parabolic equations. *Soviet Math. Dokl.*, 40: 636–639, (1990).

[Oli92] W.M. Oliva. Stability of Morse-Smale maps. Preprint, (1992).

[P69] J. Palis Jr. On Morse-Smale dynamical systems. *Topology*, 8: 385–404, (1969).

[PdM82] J. Palis Jr. and W. de Melo. *Geometric Theory of Dynamical Systems.* Springer Verlag, New York, 1982.

[Pol95] P. Poláčik. High-dimensional ω-limit sets and chaos in scalar parabolic equations. *J. Diff. Eq.*, 119: 24–53, (1995).

[PS70] J. Palis and S. Smale. Structural stability theorems. In *Global Analysis.* Proc. Symp. in Pure Math. vol. XIV. AMS, Providence, 1970.

[Roc91] C. Rocha. Properties of the attractor of a scalar parabolic PDE. *J. Dyn. Diff. Eq.*, 3: 575–591, (1991).

[Stu36] C. Sturm. Sur une classe d'équations à différences partielles. *J. Math. Pure Appl.*, 1: 373–444, (1836).

[Tem88] R. Teman. *Infinite Dimensional Dynamical Systems in Mechanics and Physics.* Springer-Verlag, New York, 1988.

Bernold Fiedler
Institut für Mathematik I
Freie Universität Berlin
Arnimallee 2-6
D-14195 Berlin
Germany
fiedler@math.fu-berlin.de

Maximal Indexes of Tits Algebras

A. S. Merkurjev[1]

Received: April 16, 1996

Communicated by Ulf Rehmann

Abstract. Let G be a split simply connected semisimple algebraic group over a field F and let C be the center of G. It is proved that the maximal index of the Tits algebras of all inner forms of G_L over all field extensions L/F corresponding to a given character χ of C equals the greatest common divisor of the dimensions of all representations of G which are given by the multiplication by χ being restricted to C. An application to the discriminant algebra of an algebra with an involution of the second kind is given.

1991 Mathematics Subject Classification: Primary 20G15.

Let G be an adjoint semisimple algebraic group defined over a field F, let $\pi : \widetilde{G} \to G$ be the universal covering and let $C = \ker(\pi)$ denote the center of $\widetilde{G}$. In [13] Tits has constructed a homomorphism

$$\beta : C^*(F) \to \mathrm{Br}(F)$$

where $C^*(F)$ is the group of characters of C defined over F and $\mathrm{Br}(F)$ is the Brauer group of F. For any character $\chi \in C^*(F)$ one can choose a central simple algebra A (called the *Tits algebra*), representing the class $\beta(\chi) \in \mathrm{Br}(F)$, in such a way that there is a group homomorphism

$$\widetilde{G} \to \mathbf{GL}_1(A)$$

restricting to the character χ on the center C and inducing an irreducible representation over a separable closure F_{sep} of the field F. It follows from the representation theory of semisimple algebraic groups that the index $\mathrm{ind}(A)$ of the algebra A divides the dimension of any irreducible representation $\rho : \widetilde{G}^q \to \mathbf{GL}(V)$ of a quasisplit inner form $\widetilde{G}^q$ of $\widetilde{G}$ such that the restriction of ρ to the center C^q of $\widetilde{G}^q$ is given by the multiplication by χ (we identify the Galois modules of the character groups C^* and C^{q*}). Therefore, if we denote by $n_\chi(G)$ the greatest common divisor of the dimensions of all such representations, then $\mathrm{ind}(A)$ divides $n_\chi(G)$. The numbers $n_\chi(G)$ depend only on the class of the inner forms of G, i.e. on the Dynkin diagram $D = \mathrm{Dyn}(G_{\mathrm{sep}})$, and the action of the absolute Galois group of F on $\mathrm{Aut}(D)$. In particular, if G is of

[1]I would like to thank the Université de Franche-Comté at Besançon and the Alexander von Humboldt-Stiftung for financial support.

inner type, then the numbers $n_\chi(G)$ depend only on the isomorphism class of G over F_{sep} and were computed in [5].

It was proved in [5], case by case, that, for a group G of inner type, the maximal possible index of the Tits algebra A corresponding to χ reaches its upper bound $n_\chi(G)$. More precisely, there is a field extension E/F and an inner form G' of the group $G \times_F E$ over E such that for any character of the center of the universal covering of G', defined over E, the index of the Tits algebra A corresponding to χ equals $n_\chi(G) = n_\chi(G')$.

We give here a uniform proof of this statement for all adjoint semisimple algebraic groups G (not necessarily of inner type). The field E appears as a function field of a "classifying variety" Y for the corresponding adjoint quasisplit group G^q.

The universal property of the variety Y asserts that any inner form of G over an arbitrary field extension L/F arises from some L-point of Y. Hence, the Tits algebras over the function field $E = F(Y)$ are generic ones, and, therefore, are of maximal index. It follows that, if the index of the Tits algebra A corresponding to χ reaches the upper bound $n_\chi(G)$ over some field extension, then it does so over $F(Y)$.

In the first part of the paper we define, for a group scheme $\mathcal{G}$, the dual group scheme $\mathcal{G}'$ with respect to a $\mathcal{G}$-torsor. This construction is a slight generalization of the corollary of Prop. 34 in [10]. For an adjoint semisimple algebraic group G over a field F we construct a classifying variety Y over F such that the scheme $\mathcal{G}'$, dual to $G \times_F Y$ with respect to a certain torsor, represents the algebraic family of all inner forms of G.

In section 4 we define Tits algebras and give a list of all Tits algebras for all absolutely simple groups of classical types.

The main result is formulated in section 5. The rest of the paper is devoted to the proof of the theorem. In the last section we give an application of the theorem in the case of groups of outer type A_{2n-1} which was not covered in [5].

All the group schemes considered in the paper are assumed to be flat affine of finite type over a Noetherian separated base scheme Y.

For a field F we denote by F_{sep} a separable closure and by Γ the absolute Galois group $\mathrm{Gal}(F_{\mathrm{sep}}/F)$. The split 1-dimensional torus $\mathrm{Spec}F[t, t^{-1}]$ is denoted by $\mathbb{G}_m$.

1. Dual group scheme with respect to a torsor

Let $\mathcal{G}$ be a group scheme over a scheme Y, and let $\pi : X \to Y$ be a (left) $\mathcal{G}$-torsor [7]. Denote by $\mathrm{Aut}_{\mathcal{G}}(X)$ the group of all $\mathcal{G}$-automorphisms of X over Y. If $X = \mathcal{G}$ is a trivial torsor, then the map $\mathcal{G}(Y) \to \mathrm{Aut}_{\mathcal{G}}(X)$ given by the rule $g \mapsto (g' \mapsto g' \cdot g^{-1})$ is clearly a group isomorphism.

Consider the sheaf of groups in the flat topology Y_{fl} on Y:

$$S(Z) = \mathrm{Aut}_{\mathcal{G}\times_Y Z}(X \times_Y Z).$$

Proposition 1.1. *The sheaf S is represented by a group scheme over Y.*

Proof. Since $\pi : X \to Y$ is faithfully flat, it is sufficient to prove that the restriction of S on X is represented by a group scheme (by faithfully flat descent, [7, Th.2.23]). But over X the torsor $\pi \times_Y \mathrm{id} : X \times_Y X \to X$ is trivial, hence for any scheme Z over X we have a canonical isomorphism $S(Z) \xrightarrow{\sim} \mathcal{G}(Z)$ and therefore the restriction of S on X is represented by $\mathcal{G} \times_Y X$. □

We denote by $\mathcal{G}'$ the group scheme over Y representing S and call $\mathcal{G}'$ *the group scheme dual to* $\mathcal{G}$ *with respect to the* $\mathcal{G}$*-torsor* $\pi : X \to Y$. It follows from the proof of proposition 1.1 that the group schemes $\mathcal{G}' \times_Y X$ and $\mathcal{G} \times_Y X$ are isomorphic over X.

By definition the group scheme $\mathcal{G}'$ acts on X over Y.

Proposition 1.2. *The morphism* $\pi : X \to Y$ *is a* $\mathcal{G}'$*-torsor.*

Proof. By faithfully flat descent we may assume that $X = \mathcal{G}$ is a trivial $\mathcal{G}$-torsor. Then the action of $\mathcal{G}' \xrightarrow{\sim} \mathcal{G}$ on X clearly leads to the structure of a trivial $\mathcal{G}'$-torsor on X. □

There is a natural bijection of the set of isomorphism classes of $\mathcal{G}$-torsors $\pi : X \to Y$ over Y and the set $H^1_{\mathrm{fl}}(Y, \mathcal{G})$ (see [7]).

Let $f : \mathcal{G} \to \mathcal{G}_1$ be a morphism of group schemes over Y, and let $\pi : X \to Y$ be a $\mathcal{G}$-torsor. A $\mathcal{G}_1$-torsor $\pi_1 : X_1 \to Y$ representing the image of the class of π under the map

$$H^1_{\mathrm{fl}}(Y, \mathcal{G}) \to H^1_{\mathrm{fl}}(Y, \mathcal{G}')$$

is called *the image of the* $\mathcal{G}$*-torsor* $\pi : X \to Y$ *under* f. Let $\mathcal{G}'$ (resp. $\mathcal{G}'_1$) be the group scheme dual to $\mathcal{G}$ (resp. $\mathcal{G}_1$) with respect to the torsor $\pi : X \to Y$ (resp. $\pi_1 : X_1 \to Y$). The natural group homomorphism

$$\mathrm{Aut}_{\mathcal{G}}(X) \to \mathrm{Aut}_{\mathcal{G}_1}(X_1)$$

induces a group scheme homomorphism $f' : \mathcal{G}' \to \mathcal{G}'_1$ over Y of the dual group schemes.

2. **PGL**-torsors

Let $p : \mathcal{V} \to Y$ be a vector bundle over Y and $\mathcal{E} = \mathrm{End}_Y(\mathcal{V})$ (viewed as a vector bundle over Y). Consider the group scheme $\mathcal{G} = \mathbf{PGL}(\mathcal{V})$ over Y. Let $\pi : X \to Y$ be a $\mathcal{G}$-torsor. The group scheme $\mathcal{G}$ acts on $\mathcal{E}$ and on X over Y, hence on $\mathcal{E} \times_Y X$. Denote by $\mathrm{Sec}_{\mathcal{G}}(\mathcal{E})$ the $\Gamma(Y, \mathcal{O}_Y)$-algebra of $\mathcal{G}$-invariant sections $X \to \mathcal{E} \times_Y X$ of the vector bundle $\mathcal{E} \times_Y X \to X$. Consider the sheaf T of algebras on Y_{fl}:

$$T(Z) = \mathrm{Sec}_{\mathcal{G} \times_Y Z}(\mathcal{E} \times_Y Z).$$

Proposition 2.1. *The sheaf* T *is represented by the total space of an Azumaya algebra over* Y.

Proof. By faithfully flat descent we may assume that $X = \mathcal{G}$ is a trivial torsor. Then for any scheme Z over Y we have $T(Z) = \mathrm{Mor}_Y(Z, \mathcal{E})$, hence T is represented by $\mathcal{E}$ which is the total space of the associated locally free sheaf $\mathcal{E}nd_Y(\mathcal{V})$ of Azumaya algebras. □

We call an Azumaya algebra $\mathcal{A}$ over Y whose total space represents T *the algebra associated to the* $\mathcal{G}$*-torsor* $\pi : X \to Y$. It follows from the proof of proposition 2.1 that the $\mathcal{O}_X$-algebra $\pi^*\mathcal{A}$ is isomorphic to $\pi^*(\mathcal{E}nd_Y(\mathcal{V}))$.

Consider the sheaf of sets on Y_{fl}:

$$U(Z) = \mathrm{Iso}_{\mathcal{O}_Z-alg}(\lambda^*\mathcal{A}, \lambda^*\mathcal{E}nd_Y(\mathcal{V}))$$

for any $\lambda : Z \to Y$. The group $\mathcal{G}(Z)$ acts naturally on $U(Z)$ making U a $\mathcal{G}$-torsor.

Proposition 2.2. *The sheaf* U *is represented by the* $\mathcal{G}$*-torsor* $\pi : X \to Y$.

Proof. A morphism $\mu : Z \to X$ over Y defines a trivialization of the torsor $X \times_Y Z \to Z$ and, hence, an isomorphism of $\mathcal{O}_Z$-algebras $(\pi\mu)^*\mathcal{A}$ and $(\pi\mu)^*(\mathcal{E}nd_Y(\mathcal{V}))$. Therefore, we get a map $X(Z) \to U(Z)$ which gives rise to a map of sheaves

$$\mathrm{Mor}_Y(*, X) \to U.$$

To prove that this map is a bijection, by faithfully flat descent one may assume that X is a trivial torsor. By the Skolem-Noether theorem in this case the statement is clear. □

Remark 2.3. Proposition 2.2 shows how to reconstruct the $\mathcal{G}$-torsor $\pi : X \to Y$ out of the algebra $\mathcal{A}$. Thus, we have a bijection between the set of isomorphism classes of Azumaya algebras $\mathcal{A}$ over Y such that $\pi^*\mathcal{A} \xrightarrow{\sim} \pi^*(\mathcal{E}nd_Y(\mathcal{V}))$ and the set of isomorphism classes of $\mathbf{PGL}(\mathcal{V})$-torsors over Y.

The group scheme $\mathbf{PGL}_1(\mathcal{A})$ over Y acts naturally on the sheaf U, and the action commutes with that of $\mathcal{G}$. Hence, we have a group scheme homomorphism $\mathbf{PGL}_1(\mathcal{A}) \to \mathcal{G}'$ where $\mathcal{G}'$ is the group scheme dual to $\mathcal{G}$ with respect to the $\mathcal{G}$-torsor $\pi : X \to Y$. To prove that this homomorphism is an isomorphism, by faithfully flat descent, one may consider the split situation in which our statement is clear. Hence,

$$\mathcal{G}' = \mathbf{PGL}(\mathcal{V})' = \mathbf{PGL}_1(\mathcal{A}).$$

Now let $\mathcal{G}$ be an arbitrary group scheme over Y, let $\pi : X \to Y$ a $\mathcal{G}$-torsor and let

$$f : \mathcal{G} \to \mathbf{PGL}(\mathcal{V})$$

be a projective representation over Y, where $\mathcal{V}$ is a vector bundle over Y. Denote by $\mathcal{A}$ the Azumaya algebra on Y associated to the $\mathbf{PGL}(\mathcal{V})$-torsor, which is equal to the image of π under f. We call $\mathcal{A}$ *the algebra associated to the $\mathcal{G}$-torsor π and the projective representation f*. There is a natural group scheme homomorphism

$$f' : \mathcal{G}' \to \mathbf{PGL}_1(\mathcal{A})$$

where $\mathcal{G}'$ is the group scheme dual to $\mathcal{G}$ with respect to π.

3. Inner forms

Let G be a semisimple algebraic algebraic group defined over a field F with center $Z(G)$. Denote by $\overline{G}$ the corresponding adjoint group $G/Z(G)$. An algebraic group G' over F is called *a twisted form of G* if $G'_{\mathrm{sep}} \simeq G_{\mathrm{sep}}$. The set of isomorphism classes of twisted forms of G is in 1–1 correspondence with the set $H^1(F, \mathrm{Aut}(G_{\mathrm{sep}}))$ ([10]). The natural homomorphism

$$\overline{G}(F_{\mathrm{sep}}) \to \mathrm{Aut}(G_{\mathrm{sep}}), \qquad \bar{g} \mapsto (g' \mapsto gg'g^{-1})$$

induces the map

$$\alpha : H^1(F, \overline{G}(F_{\mathrm{sep}})) \to H^1(F, \mathrm{Aut}(G_{\mathrm{sep}})).$$

A twisted form G' of the group G is called an *inner form of G* if the cocycle corresponding to G' belongs to the image of α. The group G is called of *inner type* if G is an inner form of a split group.

Assume now that G is an adjoint group, i.e. $\overline{G} = G$. Let X be a G-torsor over F. It corresponds to some element $\xi \in H^1(F, G(F_{\mathrm{sep}}))$ ([10]). It is straightforward to check that the group G', dual to G with respect to the torsor X, corresponds to $\alpha(\xi) \in H^1(F, \mathrm{Aut}(G_{\mathrm{sep}}))$.

We have proved

Proposition 3.1. *Let G and G' be adjoint semisimple algebraic groups over a field F. Then G' is an inner form of G iff there is a G-torsor X over F such that G' is the dual group with respect to the G-torsor X.* $\square$

Remark 3.2. The second condition of proposition 3.1 can be taken as the definition of an inner form of an adjoint group (in order to avoid referring to cocycles).

4. Tits algebras

Let G be an adjoint semisimple algebraic group defined over a field F, let $\widetilde{G} \to G$ be the universal covering, and C the kernel of the covering. It is known that C, being the center of $\widetilde{G}$, is a closed subscheme of $\widetilde{G}$ of multiplicative type (not necessarily reduced) ([2],[12]). Denote by C^* the finite Γ-module $\mathrm{Hom}(C_{\mathrm{sep}}, \mathbb{G}_m)$ of characters.

The group G is an inner form of some quasisplit group defined over F ([1],[12]). By proposition 3.1, there exists a G-torsor X over F such that the group G', dual to G with respect to X, is quasisplit. The choice of a point of X over F_{sep} defines an isomorphism $G_{\mathrm{sep}} \xrightarrow{\sim} G'_{\mathrm{sep}}$ which is uniquely determined up to conjugation. This isomorphism extends uniquely to an isomorphism $\widetilde{G}_{\mathrm{sep}} \xrightarrow{\sim} \widetilde{G}'_{\mathrm{sep}}$ where $\widetilde{G}'$ is the universal covering of G' over F ([12]). Hence, we obtain an isomorphism of the centers $\varphi : C_{\mathrm{sep}} \xrightarrow{\sim} C'_{\mathrm{sep}}$. One can easily see that this isomorphism is defined over F (hence, induces an isomorphism of Γ-modules $\varphi^* : C^* \xrightarrow{\sim} C'^*$) and depends only on the choice of the G-torsor X (which is not unique in general) but not on the point of X over F_{sep}.

Denote by B a Borel subgroup in $\widetilde{G}'$ defined over F, by T a maximal torus in B defined over F and by Λ the subgroup in T^* generated by roots of $\widetilde{G}$ relative to T. The restriction map induces the natural isomorphism of Γ-modules

$$T^*/\Lambda \xrightarrow{\sim} C^*.$$

There is a partial ordering on T^*: we write $\alpha > \beta$ for α, $\beta \in T^*$ if $\alpha - \beta$ is a sum of roots of B. In each coset of $\widetilde{T}^*/\Lambda$ there is a unique minimal element with respect to this ordering called the *minimal weight*.

Choose a character $\chi \in C^*$ defined over F and put $\chi' = \varphi^*(\chi) \in C'^*$. By the representation theory of quasisplit semisimple groups (see [13]) there is an irreducible representation $\tilde{\rho} : \widetilde{G}' \to \mathbf{GL}(V)$ such that the restriction of $\tilde{\rho}$ to C' is given by multiplication by χ'. Consider a central simple F-algebra A associated to the G-torsor X and the projective representation $\rho : G' \to \mathbf{PGL}(V)$ induced by $\tilde{\rho}$ (section 2). The algebra A is called the *Tits algebra of the group G corresponding to the representation ρ*. Its class in $\mathrm{Br}(F)$ depends only on the choice of character $\chi \in C^*$ ([13]) and is called *the Tits class of the group G corresponding to* χ. By construction, the index of A divides $\dim V$. Denote by $n_\chi(G)$ the greatest common divisor of the numbers $\dim V$ for all representations $\tilde{\rho} : \widetilde{G}' \to \mathbf{GL}(V)$ such that the restriction of $\tilde{\rho}$ on C' is given by multiplication by χ'. We have observed that $\mathrm{ind}\, A$ divides $n_\chi(G)$ (see [5]). If $\chi = 0$, then $n_\chi(G) = 1$.

Let $\chi \in C^*(F) \simeq (T^*/\Lambda)^\Gamma$ and $\mu \in T^*$ be the minimal weight in the coset χ. Let $\tilde{\rho} : \widetilde{G}' \to \mathbf{GL}(V)$ be a representation (unique up to an isomorphism) with the highest weight μ called the *minimal representation.* The Tits algebra corresponding to ρ is called the *minimal Tits algebra of* G and is denoted by A_χ. The algebra A_χ is the canonical representative of the Tits class corresponding to χ. For example, if $\chi = 0$, then $A_\chi = F$. Any Tits algebra is Brauer equivalent to a minimal one.

Remark 4.1. The isomorphism $\varphi : C_{\text{sep}} \overset{\sim}{\to} C'_{\text{sep}}$ depends on the choice of the G-torsor X. Another choice of X changes φ by an automorphism of C' induced by an (outer) automorphism of G', but clearly does not change the numbers $n_\chi(G)$.

Remark 4.2. By definition, the numbers $n_\chi(G)$ depend only on the quasisplit inner form of G and hence do not change if we replace G by any inner form of it. In turn, the class of inner forms of G is uniquely determined by the isomorphism class of G over F_{sep} and the action of Γ on the group of outer automorphisms

$$\mathrm{Out}(G_{\text{sep}}) = \mathrm{Aut}(G_{\text{sep}})/\,\mathrm{Int}(G_{\text{sep}}) = \mathrm{Aut}(\mathrm{Dyn}(G_{\text{sep}}))$$

of the group G_{sep}. If we change F by a field extension E/F such that F is separably closed in E, the numbers $n_\chi(G)$ do not change. If G is a group of inner type (i.e. G' is a split group, or equivalently, Γ acts trivially on $\mathrm{Out}(G_{\text{sep}})$) then the numbers $n_\chi(G)$ depend only on the isomorphism class of G over F_{sep} and are computed in [5].

We would like to classify the Tits classes of all adjoint semisimple algebraic groups. A Tits algebra of the product of adjoint semisimple groups is the tensor product of the Tits algebras of factors. Since any adjoint semisimple group is the product of the groups $G_1 = R_{L/F}(G)$ where G is an absolutely simple adjoint group over a finite separable field extension L/F ([12]), it suffices to describe the Tits algebras of G_1. If $\widetilde{G} \to G$ is the universal covering of G with kernel C, then

$$\widetilde{G}_1 = R_{L/F}(\widetilde{G}) \to R_{L/F}(G) = G_1$$

is the universal covering of G_1 with kernel $C_1 = R_{L/F}(C)$.

Let $F \subset L \subset F_{\text{sep}}$, $\Gamma_0 = \mathrm{Gal}(F_{\text{sep}}/L) \subset \Gamma$. We have a canonical isomorphism

$$\theta : C^*(L) = (C^*)^{\Gamma_0} \overset{\sim}{\to} (C_1^*)^\Gamma = C_1^*(F),$$

and for any $\chi_0 \in C^*(L)$ the Tits algebra A_χ with $\chi = \theta(\chi_0)$ for the group G_1 equals the corestriction in the extension L/F of the Tits algebra A_{χ_0} of G ([13]). Hence, it is sufficient to classify the Tits classes of absolutely simple adjoint groups.

Below is the list of minimal Tits algebras and numbers $n_\chi(G)$ for absolutely simple adjoint groups. We use the notation and the computations from [4] and [5].

4.1. Type A_n. An adjoint simple algebraic group of the type A_n, defined over F, is isomorphic to the projective unitary group $G = \mathbf{PGU}(B,\tau)$, where B is an Azumaya algebra of degree $n+1$ over an étale quadratic extension L/F with an involution τ of the second kind trivial on F. Its universal covering is the special unitary group $\widetilde{G} = \mathbf{SU}(B,\tau)$

Assume first that L splits, i.e. $L \simeq F \times F$. In this case $B \simeq A \times A^{op}$ with the switch involution τ where A is a central simple algebra of degree $n+1$ over F, where

$\widetilde{G} = \mathbf{SL}_1(A)$ and $G = \mathbf{PGL}_1(A)$. Then $C = \boldsymbol{\mu}_{n+1}$, and $C^* = \mathbb{Z}/(n+1)\mathbb{Z}$ with the trivial Γ-action. For any $i = 0, 1, \ldots, n$, consider the natural representation

$$\rho_i : \widetilde{G} \to \mathbf{GL}_1(\lambda^i A)$$

where $\lambda^i A$ are *external powers* of A (see [4]). In the split case, ρ_i is the i-th external power representation known as a minimal representation. Hence, $\lambda^i A$ for $i = 0, 1, \ldots, n$ are minimal Tits algebras of G. If $\chi = i + (n+1)\mathbb{Z} \in C^*$, then $n_\chi = (n+1)/gcd(i, n+1)$.

Now let $\widetilde{G} = \mathbf{SU}(B, \tau)$, where B is a central simple algebra of degree $n+1$ with a unitary involution over a quadratic separable field extension L/F. The group Γ acts on $C^* = \mathbb{Z}/(n+1)\mathbb{Z}$ by $x \mapsto -x$ through $\mathrm{Gal}(L/F)$. The only non-trivial element in $C^*(F)$ is $\chi = \frac{n+1}{2} + (n+1)\mathbb{Z}$ (when n is odd). There is a natural homomorphism

$$\rho : \widetilde{G} \to \mathbf{GL}_1(D(B, \tau)),$$

where $D(B, \tau)$ is the discriminant algebra (see section 10). In the split case ρ is the external $\frac{n+1}{2}$-power representation. Hence, the algebra $D(B, \tau)$ is the minimal Tits algebra for the group G corresponding to χ. The number n_χ equals 2 if $(n+1)$ is a 2-power and equals 4 otherwise (see section 10).

4.2. Type B_n. An adjoint simple algebraic group of type B_n, defined over F, is isomorphic to the special orthogonal group $G = \mathbf{O}^+(V, q)$ where (V, q) is a non-degenerate quadratic form of dimension $2n+1$. Its universal covering is the spinor group $\widetilde{G} = \mathbf{Spin}(V, q)$. Then $C = \boldsymbol{\mu}_2$, $C^* = \mathbb{Z}/2\mathbb{Z} = \{0, \chi\}$. The embedding

$$\widetilde{G} \hookrightarrow \mathbf{GL}_1(C_0(V, q)),$$

where $C_0(V, q)$ is the even Clifford algebra of (V, q), is, in the split case, the spinor representation known as a minimal representation. Hence, the even Clifford algebra $C_0(V, q)$ is the minimal Tits algebra A_χ. The number n_χ equals 2^n.

4.3. Type C_n. An adjoint simple algebraic group of type C_n, defined over F, is isomorphic to the group of projective similitudes $G = \mathbf{PGSp}(A, \sigma)$, where A is a central simple algebra of degree $2n$ with a symplectic involution σ. Its universal covering is the symplectic group $\widetilde{G} = \mathbf{Sp}(A, \sigma)$. Then $C = \boldsymbol{\mu}_2$ and $C^* = \mathbb{Z}/2\mathbb{Z} = \{0, \chi\}$. The embedding

$$\widetilde{G} \hookrightarrow \mathbf{GL}_1(A)$$

is, in the split case, a minimal representation. Hence, A is the minimal Tits algebra A_χ. The number n_χ is the largest 2-power which divides $2n$.

4.4. Type D_n. An adjoint simple algebraic group of type D_n, defined over F (of non-trialitarian type if $n = 4$), is isomorphic to the group of proper projective similitudes $G = \mathbf{PGO}^+(A, \sigma, f)$ where A is a central simple, algebra of degree $2n$ with an orthogonal pair (σ, f) (see [4]). Its universal covering is the spinor group $\widetilde{G} = \mathbf{Spin}(A, \sigma, f)$. Then $C^* = \{0, \chi, \chi^+, \chi^-\}$ where χ factors through the special orthogonal group $\mathbf{O}^+(A, \sigma, f)$. The composition

$$\mathbf{Spin}(A, \sigma, f) \to \mathbf{O}^+(A, \sigma, f) \hookrightarrow \mathbf{GL}_1(A)$$

is, in the split case, the standard minimal representation. Hence, A is the minimal Tits algebra A_χ. The number n_χ equals the largest 2-power which divides $2n$.

Assume that the discriminant of σ is trivial (i.e. the center Z of the Clifford algebra $C(A,\sigma,f)$ splits). The group Γ acts trivially on C^*. The natural compositions

$$\mathbf{Spin}(A,\sigma,f) \hookrightarrow \mathbf{GL}_1(C(A,\sigma,f)) \to \mathbf{GL}_1(C^{\pm}(A,\sigma,f))$$

where $C^{\pm}(A,\sigma,f)$ are simple components of $C(A,\sigma,f)$, are, in the split case, semispinor minimal representations. Hence, $C^{\pm}(A,\sigma,f)$ are minimal Tits algebras $A_{\chi^{\pm}}$. The numbers $n_{\chi^{\pm}}$ equal 2^{n-1}.

If the discriminant of σ is not trivial, then Γ interchanges χ^+ and χ^- and χ is the only nontrivial Γ-invariant character.

4.5. Exceptional types.

4.5.1. *Trialitarian type D_4.* The image of the map $\Gamma \to \mathrm{Aut}(C^*)$ contains a subgroup of order 3. It implies that $C^*(F) = 0$ and there are no nontrivial characters and Tits algebras.

4.5.2. *Type E_6.* In this case $C^* \simeq \mathbb{Z}/3\mathbb{Z}$ and for a nontrivial character $\chi \in C^*(F)$ one has $n_\chi = 27$.

4.5.3. *Type E_7.* In this case $C^* \simeq \mathbb{Z}/2\mathbb{Z}$ and for a nontrivial character $\chi \in C^*(F)$ one has $n_\chi = 8$.

4.5.4. *Types E_8, F_4 and G_2.* In these cases $C^* = 0$ and there are no nontrivial characters and Tits algebras.

5. The classifying variety of a group

Let G be an adjoint semisimple algebraic group over a field F and Y be a scheme over F. Consider the group scheme $\mathcal{G} = G \times_F Y$ over Y, and an arbitrary $\mathcal{G}$-torsor $\pi : X \to Y$. Denote by $\mathcal{G}'$ the dual scheme with respect to this torsor. For any rational point $y \in Y(F)$ the fiber $\mathcal{G}'_y$ of $\mathcal{G}'$ over y is dual to $\mathcal{G}_y = G$ with respect to the G-torsor $\pi_y : X_y \to \operatorname{Spec} F$. Hence, by proposition 3.1, an algebraic group $\mathcal{G}'_y$ is an inner form of G. So, we can view the scheme $\mathcal{G}'$ as the algebraic family of inner forms of G.

Now we take a specific scheme Y. Let $G \hookrightarrow \mathbf{GL}_n$ be any faithful representation over F. Consider the homogeneous variety $Y = \mathbf{GL}_n / G$ and the canonical G-torsor $\pi : \mathbf{GL}_n \to Y$. The variety Y is called the *classifying variety of* G. The universal property of Y asserts that any inner form of G is a member of the algebraic family $\mathcal{G}'$ over Y:

Proposition 5.1. *For any inner form G' of G over F there exists a rational point $y \in Y(F)$ such that $G' \simeq \mathcal{G}'_y$ over F.*

Proof. This follows from Hilbert's Theorem 90 and the exact sequence of pointed sets ([7],[10])

$$Y(F) \to H^1(F, G(F_{\mathrm{sep}})) \to H^1(F, \mathbf{GL}_n(F_{\mathrm{sep}}))$$

induced by the exact sequence

$$1 \to G(F_{\mathrm{sep}}) \to \mathbf{GL}_n(F_{\mathrm{sep}}) \to Y(F_{\mathrm{sep}}) \to 1. \quad \square$$

Now let G_1 be any adjoint semisimple algebraic group over F, and G be its quasisplit inner form. Consider the classifying variety $Y = \mathbf{GL}_n / G$ and the group scheme $\mathcal{G}'$ dual to $\mathcal{G} = G \times_F Y$ with respect to the $\mathcal{G}$-torsor $\pi : \mathbf{GL}_n \to Y$. By proposition 5.1, we have $G_1 \simeq \mathcal{G}'_y$ for some $y \in Y(F)$. Let $\xi \in Y$ be the generic point. The generic fiber $\mathcal{G}'_\xi$ is an adjoint semisimple algebraic group over the function field $F(Y)$. The $\mathcal{G}$-torsor π enables us to identify the character module C^* of the center of the universal coverings of the groups G_1, G and $\mathcal{G}'_\xi$.

Now we formulate the main result.

Theorem 5.2. *For any character $\chi \in C^*(F)$, the index of the Tits class of the group $\mathcal{G}'_\xi$ corresponding to χ equals $n_\chi(G_1) = n_\chi(G) = n_\chi(\mathcal{G}'_\xi)$.*

Corollary 5.3. *For any adjoint semisimple algebraic group G_1 over a field F there exists a field extension E/F and an inner form G_2 of the group $G_1 \otimes_F E$ over E such that F is separably closed in E and for any character χ of the center of the universal covering of G_2, with χ defined over E, the index of the Tits class of the group G_2 corresponding to χ equals $n_\chi(G_1) = n_\chi(G_2)$.* □

6. G-modules

Let $\mathcal{G}$ be a group scheme over a scheme Y. Assume that $\mathcal{G}$ acts on a scheme X over Y. The morphism of the $\mathcal{G}$-action on X we denote by

$$\theta : \mathcal{G} \times_Y X \to X.$$

A *$\mathcal{G}$-module $\mathcal{F}$ on X* is a quasicoherent $\mathcal{O}_X$-module $\mathcal{F}$ together with an isomorphism of $\mathcal{O}_{\mathcal{G}\times_Y X}$-modules

$$\varphi : \theta^*\mathcal{F} \xrightarrow{\sim} p_2^*\mathcal{F}$$

(where $p_2 : \mathcal{G} \times_Y X \to X$ is the projection), satisfying the cocycle condition

$$p_{23}^*(\varphi) \circ (\mathrm{id} \times \theta)^*(\varphi) = (m \times \mathrm{id})^*(\varphi)$$

where $m : \mathcal{G} \times_Y \mathcal{G} \to \mathcal{G}$ is the multiplication.

Giving a $\mathcal{G}$-module structure on a quasicoherent $\mathcal{O}_X$-module $\mathcal{F}$ is equivalent to giving, naturally in Y-schemes Z, a homomorphism of the group $\mathcal{G}(Z)$ into the automorphism group of the pair $(X \times_Y Z, \mathcal{F} \otimes_Y Z)$ ([8],[11]).

Assume that $\mathcal{G}$ acts on an Azumaya algebra $\mathcal{B}$ over X, i.e. the structure of $\mathcal{G}$-module $\mathcal{B}$ is given by an $\mathcal{O}_{\mathcal{G}\times_Y X}$-algebra isomorphism

$$\psi : \theta^*\mathcal{B} \xrightarrow{\sim} p_2^*\mathcal{B}.$$

Denote by $\underline{\underline{M}}(\mathcal{G}, X, \mathcal{B})$ the abelian category of $\mathcal{G}$-modules $\mathcal{F}$ on X, which are also left $\mathcal{B}$-modules and coherent $\mathcal{O}_X$-modules, such that the following diagram commutes:

$$\begin{array}{ccc} \theta^*\mathcal{B} \otimes \theta^*\mathcal{F} & \longrightarrow & \theta^*\mathcal{F} \\ {\scriptstyle \psi\otimes\varphi}\downarrow & & \downarrow{\scriptstyle \varphi} \\ p_2^*\mathcal{B} \otimes p_2^*\mathcal{F} & \longrightarrow & p_2^*\mathcal{F}, \end{array}$$

where the horizontal maps are given by the action of $\mathcal{B}$ on $\mathcal{F}$. Morphisms in the category are morphisms of $\mathcal{B}$- and $\mathcal{G}$-modules.

If the algebra $\mathcal{B}$ is trivial, i.e. $\mathcal{B} = \mathcal{O}_X$, then the category is simply denoted by $\underline{\underline{M}}(\mathcal{G}, X)$.

Let $\mathcal{A}$ be an Azumaya algebra on Y. Consider the Azumaya algebra $\mathcal{B} = \pi^*\mathcal{A}$ on X, where $\pi : X \to Y$ is the structure morphism, and the category $\underline{\underline{M}}(Y, \mathcal{A})$ of left $\mathcal{A}$-modules which are coherent $\mathcal{O}_Y$-modules. For $\mathcal{M} \in \underline{\underline{M}}(Y, \mathcal{A})$ the $\mathcal{O}_X$-module

$\mathcal{F}=\pi^*\mathcal{M}$ has a natural structure of a $\mathcal{B}$-module. Since $\pi\theta=\pi p_2$, it follows that we also have a natural $\mathcal{G}$-module structure on $\mathcal{F}$ given by the isomorphisms

$$\varphi:\theta^*\mathcal{F}\simeq(\pi\theta)^*\mathcal{M}=(\pi p_2)^*\mathcal{M}\simeq p_2^*\mathcal{F}.$$

Thus, we have obtained a functor

$$\pi^*:\underline{\underline{M}}(Y,\mathcal{A})\to\mathcal{M}(\mathcal{G},X,\mathcal{B}),\qquad \mathcal{M}\mapsto(\pi^*\mathcal{M},\varphi).$$

Proposition 6.1. *If $\pi:X\to Y$ is a $\mathcal{G}$-torsor then π^* is an equivalence of categories.*

Proof. Under the isomorphisms

$$\mathcal{G}\times_Y X\xrightarrow{\sim}X\times_Y X,\qquad (g,x)\mapsto(gx,x)$$

$$\mathcal{G}\times_Y\mathcal{G}\times_Y X\xrightarrow{\sim}X\times_Y X\times_Y X,\qquad (g_1,g_2,x)\mapsto(g_1g_2x,g_2x,x)$$

the action morphism θ is identified with the first projection $p_1:X\times_Y X\to X$ and morphisms $m\times\mathrm{id}$, $\mathrm{id}\times\theta$ are identified with the projections $p_{13},p_{12}:X\times_Y X\times_Y X\to X\times_Y X$. Hence, the isomorphism φ giving a $\mathcal{G}$-module structure on an $\mathcal{O}_X$-module $\mathcal{F}$ can be identified with descent data, i.e. with an isomorphism

$$\psi:p_1^*\mathcal{F}\xrightarrow{\sim}p_2^*\mathcal{F}$$

of $\mathcal{O}_{X\times_Y X}$-modules satisfying the usual cocycle condition

$$(p_{23}^*\psi)\circ(p_{12}^*\psi)=p_{13}^*\psi.$$

The statement follows now by faithfully flat descent ([7, Prop.2.22]). □

7. Modules under groups of multiplicative type

Let C be a diagonalizable group scheme over a field F, and let $C^*=\mathrm{Hom}(C,\mathbb{G}_m)$ be the character group. It is known that $C=\mathrm{Spec}F[C^*]$, where $F[C^*]$ is the group algebra of C^* over F, and the comorphism

$$\overline{m}:F[C^*]\to F[C^*]\otimes_F F[C^*]$$

of the multiplication is given by the formula $\overline{m}(\chi)=\chi\otimes\chi$ ([2]).

To introduce an action of C on an affine scheme $X=\mathrm{Spec}A$ over F is the same as to give a C^*-graded structure on the F-algebra A ([3]):

$$A=\coprod_{\chi\in C^*}A_\chi.$$

The comorphism of the action of C on X,

$$\overline{\theta}:A\to F[C^*]\otimes_F A$$

is given by the formula

$$\overline{\theta}(\sum_{\chi\in C^*}a_\chi)=\sum_{\chi\in C^*}(\chi\otimes a_\chi).$$

The trivial action corresponds to the trivial graded structure: $A_\chi=0$ for $\chi\neq 0$.

Let M be an A-module. A C-module structure of the associated $\mathcal{O}_X$-module $\mathcal{F}=\widetilde{M}$ is given by an isomorphism of $F[C^*]\otimes_F A$-modules

$$\overline{\varphi}:(F[C^*]\otimes_F A)\otimes_{A,\overline{\theta}}M\xrightarrow{\sim}(F[C^*]\otimes_F A)\otimes_{A,\overline{p_2}}M,$$

satisfying the cocycle condition. Let $\overline{\varphi}(1 \otimes 1 \otimes m) = \sum_{\chi \in C^*}(\chi \otimes 1 \otimes m_\chi)$, where $m_\chi, m \in M$. Since $(e \otimes \mathrm{id})^* \varphi = \mathrm{id}$ ([11]), where $e : \mathrm{Spec} F \to G$ is the group unit, it follows that

$$m = \sum_{\chi \in C^*} m_\chi. \tag{$*$}$$

It is easy to check that the cocycle condition implies that $(m_\chi)_\rho$ equals m_χ if $\chi = \rho$ and equals 0 if $\chi \neq \rho$. Hence, the equality $(*)$ gives rise to the direct sum decomposition

$$M = \coprod_{\chi \in C^*} M_\chi$$

making M a C^*-graded A-module. Therefore, the category $\underline{\underline{M}}(C, X)$ is equivalent to the category of finitely generated C^*-graded modules.

Let an algebraic group G over a field F act on an affine scheme X over F and on an Azumaya algebra $\mathcal{B}$ on X. Assume that a closed central group subscheme $C \subset G$ of multiplicative type acts trivially on X and $\mathcal{B}$. Denote by C^* the Γ-module of characters $\mathrm{Hom}(C_{\mathrm{sep}}, \mathbb{G}_m)$. Since the group C_{sep} is diagonalizable, it follows that for any $\mathcal{F} \in \underline{\underline{M}}(G, X, \mathcal{B})$ we have a decomposition

$$\mathcal{F}_{\mathrm{sep}} = \coprod_{\chi \in C^*} (\mathcal{F}_{\mathrm{sep}})^\chi \tag{$**$}$$

into a direct sum of G_{sep}-submodules $(\mathcal{F}_{\mathrm{sep}})^\chi$ on X_{sep} (since C is central and acts trivially on X and $\mathcal{B}$).

Choose any Γ-invariant character $\chi \in C^*$ (defined over F). Clearly, $(\mathcal{F}_{\mathrm{sep}})^\chi$ and its direct complement in $(**)$ are defined over F, hence we have a canonical decomposition

$$\mathcal{F} = \mathcal{F}^\chi \oplus \mathcal{F}_\chi$$

into a direct sum of G-submodules on X. In other words, these submodules are uniquely determined by the property that $c - \chi(c)$ is trivial on $\mathcal{F}^\chi$ and invertible on $\mathcal{F}_\chi$ for all c in C.

Consider the full subcategories $\underline{\underline{M}}^\chi(G, X, \mathcal{B})$ and $\underline{\underline{M}}_\chi(G, X, \mathcal{B})$ in $\underline{\underline{M}}(G, X, \mathcal{B})$ consisting of all G-modules $\mathcal{F}$ such that $\mathcal{F} = \mathcal{F}^\chi$ and $\mathcal{F} = \mathcal{F}_\chi$ respectively. It is clear that

$$\underline{\underline{M}}(G, X, \mathcal{B}) \simeq \underline{\underline{M}}^\chi(G, X, \mathcal{B}) \times \underline{\underline{M}}_\chi(G, X, \mathcal{B}).$$

If $\chi = 0$ is the trivial character then the category $\underline{\underline{M}}^\chi(G, X, \mathcal{B})$ is equivalent to the category $\underline{\underline{M}}(G/C, X, \mathcal{B})$.

8. Equivariant algebraic K-theory

The K-groups of the category $\underline{\underline{M}}(\mathcal{G}, X, \mathcal{B})$ (see section 6) we denote by $K_*(\mathcal{G}, X, \mathcal{B})$. These groups are clearly contravariant with respect to flat $\mathcal{G}$-morphisms in X. If $\mathcal{B} = \mathcal{O}_X$ is the trivial algebra we simply write $K_*(\mathcal{G}, X)$.

Let G be an algebraic group over F acting on a scheme X over F. We will need the following particular cases of the localization theorem [11, th. 2.7] and the homotopy invariance theorem [11, cor. 4.2] in equivariant algebraic K-theory.

Proposition 8.1. *Let $U \subset X$ be an open G-equivariant subscheme. Then the restriction homomorphism $K_0(G, X) \to K_0(G, U)$ is surjective.* □

Proposition 8.2. *Assume that G acts linearly on an affine space $\mathbb{A}^n_F$ over F. Then the structure morphism $p : \mathbb{A}^n_F \to \operatorname{Spec} F$ induces an isomorphism*

$$p^* : K_*(G, \operatorname{Spec} F) \xrightarrow{\sim} K_*(G, \mathbb{A}^n_F). \quad \square$$

The category $\underline{\underline{M}}(G, \operatorname{Spec} F)$ is equivalent to the category of finite dimensional representations of G over F. The group $K_0(G, \operatorname{Spec} F)$ we denote by $R(G)$.

Assume that G acts on an Azumaya algebra $\mathcal{B}$ over X and contains a closed central subscheme C over F of multiplicative type, acting trivially on X and $\mathcal{B}$. For $\chi \in C^*(F)$ the K-groups of the category $\underline{\underline{M}}^\chi(G, X, \mathcal{B})$ we denote by $K_*^\chi(G, X, \mathcal{B})$. Since $K_*^\chi(G, X, \mathcal{B})$ is a canonical direct summand of $K_*(G, X, \mathcal{B})$ (section 7), it follows that the statements of propositions 8.1 and 8.2 still hold if we replace K_* by K_*^χ.

The group $K_0^\chi(G, \operatorname{Spec} F)$ we simply denote by $R^\chi(G)$. It is generated by the classes of all representations $\rho : G \to \mathbf{GL}(V)$ such that the restriction of ρ to C is given by χ.

9. Proof of the theorem

Let G_1 be an adjoint semisimple group over a field F, let G be the quasisplit inner form of G_1 with universal covering $\widetilde{G} \to G$, and let C be the kernel of the covering.

Choose a faithful representation $G \hookrightarrow \mathbf{GL}_n$ over F and consider the classifying variety $Y = \mathbf{GL}_n / G$ over F and the group scheme $\mathcal{G}'$ over Y dual to $\mathcal{G} = G \times_F Y$ with respect to the $\mathcal{G}$-torsor $\pi : \mathbf{GL}_n \to Y$. Let ξ be the generic point of Y. The $\mathcal{G}$-torsor π enables us to identify the character modules C^* and C'^*, where C' is the kernel of the universal covering of $\mathcal{G}'_\xi$. Choose a character $\chi' \in C'^*$ defined over $F(Y)$ and denote by $\chi \in C^*$ the corresponding character over F.

Consider a representation $\tilde{\rho} : \widetilde{G} \to \mathbf{GL}(V)$ such that the restriction of $\tilde{\rho}$ to C is given by χ. Consider also the Azumaya algebra $\mathcal{A}$ on Y associated to the $\mathcal{G}$-torsor π and the projective representation $\rho : G \to \mathbf{PGL}(V)$ induced by $\tilde{\rho}$ (section 2). We know that there is an isomorphism of G-algebras

$$\pi^*(\mathcal{A}) \simeq \pi^*(\mathcal{E}nd(V \times_F Y))$$

on $\mathbf{GL}_n$ (section 2) and that $\mathcal{A}_\xi$ is the Tits algebra corresponding to the character χ' (section 4). We have to show that $\operatorname{ind} \mathcal{A}_\xi = n_\chi(G)$.

Consider the homomorphism

$$\delta : K_0(\mathcal{A}_\xi^{op}) \to \mathbb{Z},$$

taking an A_ξ^{op}-module M to $\dim_{F(Y)} M$. It is easy to see that

$$\operatorname{im}(\delta) = \operatorname{ind} \mathcal{A}_\xi \cdot \deg \mathcal{A} \cdot \mathbb{Z}.$$

Consider also the homomorphism $\gamma : R^\chi(\widetilde{G}) \to \mathbb{Z}$, taking a representation space U to $\dim_F U$. It is clear that $\operatorname{im}(\gamma) = n_\chi(G) \cdot \mathbb{Z}$. For the proof of the theorem it is sufficient to find a surjective homomorphism

$$\alpha : R^\chi(\widetilde{G}) \to K_0(\mathcal{A}_\xi^{op})$$

such that the composition $\delta \circ \alpha$ equals $\deg \mathcal{A} \cdot \gamma = \dim V \cdot \gamma$. The homomorphism α will be found as a composite of seven epimorphisms $\alpha_1, \alpha_2, \ldots, \alpha_7$.

Consider $\mathbf{GL}_n$ as an open subvariety of the affine space $\mathbb{A} = \mathbb{A}_F^{n^2}$ of all $n \times n$-matrices over F on which the group G (and hence $\widetilde{G}$) acts linearly. The open embedding $\mathbf{GL}_n \hookrightarrow \mathbb{A}$ is clearly $\widetilde{G}$-equivariant. By proposition 8.2 (see also a remark at the end of section 8) the structure morphism $\mathbb{A} \to \operatorname{Spec} F$ induces an isomorphism

$$\alpha_1 : R^\chi(\widetilde{G}) = K_0^\chi(\widetilde{G}, \operatorname{Spec} F) \xrightarrow{\sim} K_0^\chi(\widetilde{G}, \mathbb{A}).$$

By proposition 8.1, the restriction homomorphism

$$\alpha_2 : K_0^\chi(\widetilde{G}, \mathbb{A}) \to K_0^\chi(\widetilde{G}, \mathbf{GL}_n)$$

is surjective.

Denote by $\mathcal{B}$ the algebra $\pi^*\mathcal{E}nd(V \times_F Y) = \mathcal{O}_{\mathbf{GL}_n} \otimes_F \operatorname{End} V$ on $\mathbf{GL}_n$. The group $\widetilde{G}$ clearly acts on $\mathcal{B}$. Consider two functors

$$\underline{\underline{M}}^\chi(\widetilde{G}, \mathbf{GL}_n) \underset{v}{\overset{u}{\rightleftarrows}} \underline{\underline{M}}^0(\widetilde{G}, \mathbf{GL}_n, \mathcal{B}^{op}),$$

$$u(\mathcal{F}) = V^* \otimes_F \mathcal{F}, \qquad v(\mathcal{M}) = V \otimes_{\operatorname{End} V^*} \mathcal{M},$$

where V^* is the F-vector space dual to V. The canonical isomorphisms $V \otimes_{\operatorname{End} V^*} V^* \simeq F$ and $V^* \otimes_F V \simeq \operatorname{End} V^*$ show that u and v are mutually inverse equivalences of categories. Hence, the functor u induces an isomorphism

$$\alpha_3 : K_0^\chi(\widetilde{G}, \mathbf{GL}_n) \xrightarrow{\sim} K_0^0(\widetilde{G}, \mathbf{GL}_n, \mathcal{B}^{op}).$$

Since the center C of $\widetilde{G}$ acts trivially on $\mathbf{GL}_n$ and $\mathcal{B}$, it follows that the categories $\underline{\underline{M}}^0(\widetilde{G}, \mathbf{GL}_n, \mathcal{B}^{op})$ and $\underline{\underline{M}}(G, \mathbf{GL}_n, \mathcal{B}^{op})$ are equivalent. Hence, we have an isomorphism

$$\alpha_4 : K_0^0(\widetilde{G}, \mathbf{GL}_n, \mathcal{B}^{op}) \xrightarrow{\sim} K_0(G, \mathbf{GL}_n, \mathcal{B}^{op}).$$

The isomorphism $G \times_F X \simeq \mathcal{G} \times_Y X$ shows that the categories $\underline{\underline{M}}(G, \mathbf{GL}_n, \mathcal{B}^{op})$ and $\underline{\underline{M}}(\mathcal{G}, \mathbf{GL}_n, \mathcal{B}^{op})$ are equivalent. Hence, we have an isomorphism

$$\alpha_5 : K_0(G, \mathbf{GL}_n, \mathcal{B}^{op}) \xrightarrow{\sim} K_0(\mathcal{G}, \mathbf{GL}_n, \mathcal{B}^{op}).$$

Since $\pi : \mathbf{GL}_n \to Y$ is a $\mathcal{G}$-torsor and $\mathcal{B} \simeq \pi^*\mathcal{A}$, it follows from proposition 6.1 that the functor

$$\pi^* : \underline{\underline{M}}(Y, \mathcal{A}^{op}) \to \underline{\underline{M}}(\mathcal{G}, \mathbf{GL}_n, \mathcal{B}^{op})$$

is an equivalence of categories. Hence, π^* induces an isomorphism

$$\alpha_6 : K_0(\mathcal{G}, \mathbf{GL}_n, \mathcal{B}^{op}) \xrightarrow{\sim} K_0(Y, \mathcal{A}^{op}).$$

By localization (Proposition 8.1), the functor

$$\underline{\underline{M}}(Y, \mathcal{A}^{op}) \to \underline{\underline{M}}(\mathcal{A}_\xi^{op}), \quad \mathcal{F} \mapsto \text{stalk of } \mathcal{F} \text{ at the generic point } \xi$$

induces an epimorphism

$$\alpha_7 : K_0(Y, \mathcal{A}^{op}) \to K_0(\mathcal{A}_\xi^{op}).$$

It can be easily checked that the composition $\alpha = \alpha_7 \circ \alpha_6 \circ \cdots \circ \alpha_1$ takes the class of a representation space U of the group $\widetilde{G}$ to the generic stalk $\mathcal{F}_\xi$ where

$$\pi^*\mathcal{F} = V^* \otimes_F U \otimes_F \mathcal{O}_{\mathbf{GL}_n}$$

and hence satisfies the desired condition. $\square$

10. Examples

Let L/F be a Galois quadratic field extension, $\Pi = \mathrm{Gal}(L/F)$, and let B be a central simple algebra over L of degree $2n$ with involution τ of the second kind trivial on F.

Consider the special unitary group $\widetilde{G} = \mathbf{SU}(B,\tau)$ over F. The group $\widetilde{G}(F)$ of F-points of $\widetilde{G}$ consists of all elements $b \in B^\times$ such that $\tau(b)\cdot b = 1$ and $\mathrm{Nrd}(b) = 1$ where Nrd is the reduced norm homomorphism. The Galois group Γ acts on $C^* \simeq \mathbb{Z}/2n\mathbb{Z}$ through its factor group $\Pi = \{1,\pi\}$ by $\pi(k + 2n\mathbb{Z}) = -k + 2n\mathbb{Z}$ (see section 4). The Tits algebra corresponding to the only nontrivial character $\chi = n + 2n\mathbb{Z} \in C^*(F)$ can be constructed as follows (see [4],[5]).

Consider the Severi-Brauer variety X over L corresponding to the algebra B and the canonical locally free sheaf J of rank $2n$ on X, so $B = \mathrm{End}_X(J)$ [9]. The canonical nondegenerate bilinear form on the n^{th}-exterior power of J

$$\Lambda^n J \otimes \Lambda^n J \to \Lambda^{2n} J \simeq \mathcal{O}_X$$

induces in the usual way an involution σ of the first kind on the algebra $\lambda^n B = \mathrm{End}_{\mathcal{O}_X}(\lambda^n J)$ over L. One can check that the involutions σ and $\tau' = \lambda^n\tau$ on $\lambda^n B$ commute. Therefore, the set $\{x \in \lambda^n B : \sigma(x) = \tau'(x)\}$ is a central simple algebra over F. We denote this algebra by $D(B,\tau)$ and call it *the discriminant algebra of* (B,τ) ([4]). It is the Tits algebra corresponding to the character χ.

The discriminant algebra enjoys the following properties:

1. The degree of $D(B,\tau)$ equals $\binom{2n}{n}$.

2. The restriction of σ to $D(B,\tau)$ is an involution of the first kind. In particular, the exponent of $D(B,\tau)$ divides 2.

3. $D(B,\tau) \otimes_F L \simeq \lambda^n B \sim B^{\otimes n}$. Since $\exp(B^{\otimes n})$ divides 2, it follows that $\mathrm{ind}(B^{\otimes n})$ also divides 2, and hence $\mathrm{ind}\, D(B,\tau)$ divides 4.

Let $\widetilde{G}'$ be the quasisplit inner form of $\widetilde{G}$. It is the special unitary group of the hyperbolic hermitian form over the quadratic extension L/F ([12]). Since $\widetilde{G}'_{\mathrm{sep}} \simeq \mathbf{SL}_{2n}(F_{\mathrm{sep}})$ it follows that

$$R(\widetilde{G}'_{\mathrm{sep}}) \simeq \mathbb{Z}[t_1, t_2, \ldots, t_{2n-1}]$$

where t_i is the class of the i^{th}-exterior power of the standard representation of $\mathbf{SL}_{2n}$. This ring is $C^* = \mathbb{Z}/2n\mathbb{Z}$-graded, the degree of t_i being equal to $i \pmod{2n}$. The rank map $R(\widetilde{G}'_{\mathrm{sep}}) \to \mathbb{Z}$ takes t_i to $\binom{2n}{i}$. The action of the Galois group Π on $R(\widetilde{G}'_{\mathrm{sep}})$ is given by $\pi(t_i) = t_{2n-i}$. We have also ([13]):

$$R(\widetilde{G}') \simeq \mathbb{Z}[t_1, t_2, \ldots, t_{2n-1}]^{\Pi}.$$

Using this description of the ring $R(\widetilde{G}')$ and the fact that the image of the map $R^\chi(\widetilde{G}') \to \mathbb{Z}$, taking a representation space U of the group $\widetilde{G}'$ to $\dim_F U$, equals $n_\chi \cdot \mathbb{Z}$, one can easily compute the number $n_\chi(G)$ for $G = \widetilde{G}/C$ (see [6]): $n_\chi(G)$ is equal to 2 if n is a 2–power and equals 4 otherwise. Hence, the corollary of the theorem gives in this case the following

Proposition 10.1. *For any Galois quadratic field extension L/F and $n \in \mathbb{N}$ there is a field extension E/F and a central simple algebra B of degree $2n$ over $E \otimes_F L$ with involution τ of the second kind trivial on E such that* $\mathrm{ind}\, D(B,\tau) = 2$ *if n is a 2-power and equals 4 otherwise.* □

References

[1] A. Borel, J. Tits. Groupes réductifs. *Publ. Math. I. H. E. S.* 27 (1965), 55–151.

[2] M. Demazure, P. Gabriel. *Groupes Algébriques Linéares.* Masson, Paris, 1970.

[3] R. G. Kempf, F. Knudsen, D. Mumford, B. Saint-Donat. *Toroidal imbeddings I.* Lecture Notes in Math. 339, Springer-Verlag, Berlin, 1973.

[4] M.-A. Knus, A. S. Merkurjev, M. Rost, J.-P. Tignol. *Book of involutions,* in preparation.

[5] A. S. Merkurjev, I. A. Panin, A. Wadsworth. Index reduction formulas for twisted flag varieties. Preprint, 1995.

[6] A. S. Merkurjev, I. A. Panin, A. Wadsworth. Index reduction formulas for twisted flag varieties II, Preprint, 1996.

[7] J. S. Milne. *Étale Cohomology.* Princeton Univ. Press, Princeton, N. J., 1980.

[8] D. Mumford, J. Fogarty. *Geometric Invariant Theory.* Springer-Verlag, 1982.

[9] D. Quillen. *Higher Algebraic K-theory.* Lecture Notes in Math. 341, Springer-Verlag, Berlin (1973), 85–147.

[10] J.-P. Serre. *Cohomologie Galoisienne,* 5^{eme} *édition.* Lecture Notes in Math. 5, Springer-Verlag, Berlin, 1994.

[11] R. Thomason. Algebraic K-theory of group scheme action, *in: "Algebr. topol. and algebr. K-theory (ed. W.Browder), Proc. conf. Princeton, oct.24–28, 1983, Princeton, N.J., 1987*, 539–563.

[12] J. Tits. Classification of algebraic semisimple groups, *in: Algebraic Groups and discontinuous Subgroups, (eds. A.Borel and G.D.Mostow), Proc. Symp. Pure Math., Vol. 9, 1966*, 33–62

[13] J. Tits. Représentations linéaires irréductibles d'un groupe réductif sur un corps quelconque. *J. reine angew. Math.* 247 (1971), 196–220.

A. S. Merkurjev
Sankt-Petersburg State University
and
Fakultät für Mathematik
Universität Bielefeld
Postfach 100131
D 33501 Bielefeld
Germany
merkurev@mathematik.uni-bielefeld.de

Boolean Localization, in Practice

J. F. Jardine

Received: June 13, 1996
Revised: September 11, 1996

Communicated by Ulf Rehmann

Abstract. A new proof of the existence of the standard closed model structure for the category of simplicial presheaves on an arbitrary Grothendieck site is given. This proof uses the principle of Boolean localization.

1991 Mathematics Subject Classification: Primary 55U10, secondary 18F20, 55U35.

Introduction

This paper is an exposition on the use of the topos theoretic principle of Boolean localization in demonstrating the existence of closed model structures for the categories of simplicial sheaves and presheaves on a Grothendieck site $\mathcal{C}$.

Explicitly, a closed model category is a category $\mathcal{M}$ equipped with three classes of maps, called cofibrations, fibrations and weak equivalences, such that the following list of axioms is satisfied:

CM1: $\mathcal{M}$ is closed under all finite limits and colimits.

CM2: Suppose that the following diagram commutes in $\mathcal{M}$:

$$\begin{tikzcd} X \arrow[rr, "g"] \arrow[dr, "h"'] & & Y \arrow[dl, "f"] \\ & Z. & \end{tikzcd}$$

If any two of f, g and h are weak equivalences, then so is the third.

CM3: If f is a retract of g and g is a weak equivalence, fibration or cofibration, then so is f.

CM4: Suppose that we are given a commutative solid arrow diagram

$$\begin{tikzcd} U \arrow[r] \arrow[d, "i"'] & X \arrow[d, "p"] \\ V \arrow[r] \arrow[ur, dashed] & Y \end{tikzcd}$$

where i is a cofibration and p is a fibration. Then the dotted arrow exists, making the diagram commute, if either i or p is also a weak equivalence.

CM5: Any map $f : X \to Y$ may be factored:

(a) $f = p \cdot i$ where p is a fibration and i is a trivial cofibration, and

(b) $f = q \cdot j$ where q is a trivial fibration and j is a cofibration.

Here, and as usual, one says that a map is a trivial cofibration (respectively trivial fibration) if it is both a cofibration (respectively fibration) and a weak equivalence.

The fundamental example of a closed model category is the category $\mathbf{S}$ of simplicial sets [11], [12], [2]: the cofibrations of $\mathbf{S}$ are the monomorphisms, the weak equivalences are the maps which induce isomorphisms in all possible homotopy groups of associated realizations, and the fibrations are the Kan fibrations. Recall that a Kan fibration is a map $q : X \to Y$ of simplicial sets which has the "right lifting property" with respect to all inclusions $\Lambda^n_k \subset \Delta^n$ of horns in simplices. Here, the k^{th} horn Λ^n_k is the subcomplex obtained from the boundary $\partial\Delta^n$ of the standard n-simplex by deleting the k^{th} face from its list of generators.

This paper addresses the various flavours of homotopy theory that arise from contravariant simplicial set-valued diagrams, or presheaves of simplicial sets, defined on small categories equipped with Grothedieck topologies. The list of all possible Grothendieck topologies includes the option of having no topology at all, so the theory includes that of ordinary small diagrams of simplicial sets.

There are both local and global homotopy theories for simplicial presheaves. The local theory is a theory of local weak equivalences and local fibrations. In particular, if one is working in a context so civilized as the category of simplicial presheaves on the category of open subsets of a topological space X, then a map (ie. natural transformation) $f : Y \to Z$ is a local fibration if each of the induced maps $f_x : Y_x \to Z_x$, $x \in X$, in stalks is a Kan fibration of simplicial sets. Similarly, a local weak equivalence in this case is a map which induces weak equivalences in all stalks. One uses the same notion of local weak equivalence in the global theory (so that the two theories induce equivalent homotopy categories), along with cofibrations, or monomorphisms of simplicial presheaves, and then global fibrations are defined by a lifting property. There is a difference between the two theories: the Eilenberg-Mac Lane objects $K(A, n)$ associated to sheaves of abelian groups A are certainly locally fibrant, but almost never globally fibrant. A globally fibrant model of $K(A, n)$ is most properly thought of as a type of injective resolution of the abelian sheaf A, up to a degree shift.

The main results of this paper (Theorems 18, 27) together assert that the cofibrations, local weak equivalences and global fibrations determine closed model structures on the categories of simplicial presheaves and simplicial sheaves on an arbitrary Grothendieck site, and that the homotopy categories associated to simplicial presheaves and sheaves on any such site are equivalent. In all of this, one of the main technical difficulties is to arrange for a definition of local weak equivalence which specializes to the stalkwise notion in cases where the underlying topos has enough points. Historically, this was done for simplicial presheaves in a somewhat ad hoc way [4], by using sheaves of homotopy groups for associated presheaves of Kan complexes. Here, one finds an alternative definition of local weak equivalence and proofs of the main results which are based on the method of Boolean localization. The proof in the simplicial sheaf case is roughly what Joyal had in mind in his letter to Grothendieck [7] of 1984, except that it's been somewhat reverse engineered so that the relationship between sheaves of homotopy groups and weak equivalences comes out only after the fact.

Stated bluntly, the Boolean localization principle asserts that every Grothendieck topos can be faithfully imbedded in a topos that satisfies the axiom of choice. The applicability of Boolean localization in homotopy theory was first noticed by Van Osdol [14] in the 1970's, in his proof of what was then called the Illusie conjecture [3], but the descriptions of the underlying topos theory in the literature remained fragmentary until the appearance of the Mac Lane-Moerdijk book [9] in 1992. Even so, the principle as stated in [9] has to be reinterpreted somewhat to achieve the form that is used in this paper. This is done in the first section below. This reinterpretation is trivial for a topos theorist, but quite opaque to almost everybody else.

The reader who is familiar with the "Simplicial presheaves" paper [4] will notice minor technical improvements here and there, particularly in the statement and proof of Lemma 12, and in the proof of Lemma 14, along with a more aggressive use of Kan's Ex^{∞} functor throughout. The basic thrust of using a transfinite small object argument to prove the factorization axiom CM5 survives, and the local fibration concept continues to be an essential building block of the theory.

The idea appearing in the third section, that homotopy groups should really be fibred group objects, is due to Joyal as far as I can tell. Such objects, combinatorially defined, are exactly the right kind of thing to feed to a Boolean localization functor. They also have other uses: in particular, fibred homotopy group objects appear implicitly (the π_*-Kan condition) in the proof of the Bousfield-Friedlander theorem [1], [2] that recognizes homotopy cartesian diagrams of bisimplicial sets. One can also express the theory of long exact sequences for fibrations in these terms.

The writeup that follows assumes that the reader knows the basic exactness properties of a topos, and is familiar with the nuts and bolts of the associated sheaf construction. In this connection, there is one notational oddity: I use the notation L^2F to denote the associated sheaf of a presheaf F. There is some precedent for this in the literature – see [13], for example. The notation is used in order to avoid the repeated appearance of some rather ugly very wide tildes. It is also assumed that the reader is familiar with the ordinary homotopy theory of simplicial sets [10], [2].

1. Boolean localization.

Suppose that $\mathcal{C}$ is an arbitrary small Grothendieck site, and let $\mathcal{E}$ denote the sheaf category $Shv(\mathcal{C})$ on the site $\mathcal{C}$. A *Boolean localization* of $\mathcal{E}$ is a complete Boolean algebra $\mathcal{B}$ and a geometric topos morphism $\wp : Shv(\mathcal{B}) \to \mathcal{E}$, such that the inverse image functor $\wp^* : \mathcal{E} \to Shv(\mathcal{B})$ is faithful.

The definition is a bit of a mouthful. A complete Boolean algebra $\mathcal{B}$ can be characterized as a poset having at least a terminal object 1 and an initial object 0 such that $0 \neq 1$. Furthermore, $\mathcal{B}$ is required to have all limits (meets) and all colimits (joins), such that

(1) $\mathcal{B}$ is complemented in the sense that every element x has a complement $\neg x$ satisfying
$$x \vee \neg x = 1 \quad \text{and} \quad x \wedge \neg x = 0,$$
and

(2) $\mathcal{B}$ satisfies the distributive law

$$x \wedge (y \vee z) = (x \wedge y) \vee (x \wedge z).$$

The word "complete" refers to the fact that $\mathcal{B}$ is required to have *all* meets as opposed to all finite meets. Complete Boolean algebras also satisfy the infinite distributive law:

$$x \wedge (\bigvee_{i \in I} y_i) = \bigvee_{i \in I} (x \wedge y_i)$$

(see [9, p.51,114]). Finally, in $\mathcal{B}$ a family of subobjects $y_i \leq x$ of x is said to be covering if $\bigvee_{i \in I} y_i = x$. The infinite distributive law guarantees that the covering families of $\mathcal{B}$ satisfy the axioms for a pretopology, and hence give rise to a category of sheaves $Shv(\mathcal{B})$.

Boolean localizations exist for all Grothendieck toposes $\mathcal{E}$: this is a major theorem of topos theory (Mac Lane and Moerdijk call it Barr's Theorem [9, p.513], but a result of Diaconescu plays a major part – see [9, p.511]). It's also important to know, so we don't leave the realm of small sites, that the construction doesn't blow up: if the cardinality of the set of morphisms of the underlying site $\mathcal{C}$ is bounded by some infinite cardinal β, then $|\mathcal{B}| < \beta$.

Boolean localization is a vast generalization of what it means for a topos to have enough points. Specifically, the topos $\mathcal{E}$ has enough points if there is a collection $x_i : \mathbf{Sets} \to \mathcal{E}$ of geometric morphisms such that two maps $f, g : F \to G$ of $\mathcal{E}$ coincide if and only if $x_i^* f = x_i^* g$ for all $i \in I$. The set category $\mathbf{Sets}$ is equivalent to the sheaf category $Shv(\{0,1\})$ on the Boolean algebra $\{0,1\}$; more generally, the product category $\prod_{i \in I} \mathbf{Sets}$ is equivalent to $Shv(\mathcal{P}(I))$ where $\mathcal{P}(I)$ is the complete Boolean algebra determined by the set of all subsets of the set I. Finally, any collection of points $x_i : \mathbf{Sets} \to \mathcal{E}$ determines a geometric morphism $x : Shv(\mathcal{P}(I)) \to \mathcal{E}$ which is a Boolean localization for $\mathcal{E}$ if the collection of points is big enough. In other words, the topos $\mathcal{E}$ has enough points if and only if there is a Boolean localization of the form $Shv(\mathcal{P}(I)) \to \mathcal{E}$ for some set I.

We shall discuss the homotopy theoretic consequences of the existence of Boolean localizations here, and defer to the Mac Lane-Moerdijk text for its proof. The applications depend explicitly on the fact that the topos $Shv(\mathcal{B})$ satisfies the axiom of choice in the sense that every epimorphism in $Shv(\mathcal{B})$ has a section; we begin by giving an explicit proof of this result (Proposition 2).

Lemma 1. *Suppose that F is a sheaf (of sets) on a complete Boolean algebra $\mathcal{B}$. Then the category $Sub(F)$ of subobjects of F is a complete Boolean algebra.*

Proof: The category $Sub(F)$ has all meets and joins, and satisfies the infinite distributive law, by an argument on the presheaf level. Given $G \in Sub(F)$, define

$$\neg G = \bigvee_{H \wedge G = \emptyset} H.$$

It's clear that $G \wedge \neg G = \emptyset$; the interesting bit is to show that $G \vee \neg G = F$.

First of all, we show that every subobject $G \leq hom(\ ,B)$ of a representable sheaf is representable. In effect,

$$G = \varinjlim_{\phi: hom(\ ,A) \to G} hom(\ ,A),$$

and the category of morphisms $\phi : hom(\ ,A) \to G$ is small, since it can be identified with a subcategory of subobjects of B in the Boolean algebra $\mathcal{B}$. There is an isomorphism

$$\bigvee_{\phi: hom(\ ,B) \to G} B \cong \varinjlim_{\phi: hom(\ ,B) \to G} B$$

in $\mathcal{B}$, and so G is represented by the object

$$\bigvee_{\phi: hom(\ ,B) \to G} B.$$

It follows that $Sub(hom(\ ,B))$ is a complete Boolean algebra. Every subobject $F \leq hom(\ ,B)$ is represented by a subobject $A \leq B$ of $\mathcal{B}$, and $\neg A$ in $Sub(B)$ is the subobject $(\neg A) \wedge B$. Observe that $(\neg A) \wedge B$ is terminal among all subobjects of B which miss A, so that $hom(\ ,(\neg A) \wedge B) = \neg hom(\ ,B)$ in the category of subobjects of $hom(\ ,B)$.

It's certainly the case, in general, that $G \vee \neg G \leq F$ in the category of subobjects of the sheaf F. Take a sheaf morphism $\phi : hom(\ ,A) \to F$, and form the pullback diagram

$$\begin{tikzcd} \phi^{-1}(G) \arrow[r] \arrow[d] & G \arrow[d] \\ hom(\ ,A) \arrow[r, "\phi"'] & F \end{tikzcd}$$

Then there is an induced diagram

$$\begin{tikzcd} \phi^{-1}(G) \vee \neg\phi^{-1}(G) \arrow[r] \arrow[d, "\cong"'] & G \vee \neg G \arrow[d] \\ hom(\ ,A) \arrow[r, "\phi"'] & F \end{tikzcd}$$

Such diagrams exist for all such maps ϕ, and F is a colimit of representables, so that the morphism $G \vee \neg G \leq F$ has a section, and is therefore an isomorphism. □

Proposition 2. *Suppose that $\mathcal{B}$ is a complete Boolean algebra. Then every epimorphism in the sheaf category $Shv(\mathcal{B})$ has a section.*

Proof: Suppose that $\pi : F \to G$ is an epimorphism of $Shv(\mathcal{B})$. Sheaf epimorphisms are defined by the existence of partial lifts along covering families, so by looking at the terminal object, one finds an object $A \in \mathcal{B}$ such that $A \neq 0$ and there is a lifting diagram

$$\begin{tikzcd} & F \arrow[d, "\pi"] \\ hom(\ ,A) \arrow[ur] \arrow[r] & G. \end{tikzcd}$$

Observe that the map $hom(\ ,A) \to G$ defines $hom(\ ,A)$ as a subobject of G, since $hom(\ ,A)$ is a subobject of the terminal sheaf $*$. It follows that the set of all partial lifts

$$\begin{tikzcd} & F \arrow[d, "\pi"] \\ N \arrow[ur, "s"] \arrow[r, hook] & G \end{tikzcd}$$

defined on subobjects N of G is non-empty. This set has maximal elements, by Zorn's Lemma.

Suppose that

$$\begin{tikzcd} & F \arrow[d, "\pi"] \\ M \arrow[ur, "s"] \arrow[r, hook] & G \end{tikzcd}$$

is such a maximal element, and suppose that $M \neq G$. Then M has a non-empty complement $\neg M$ in G, and we can form the pullback diagram

$$\begin{tikzcd} \pi^{-1}(\neg M) \arrow[r] \arrow[d, "\pi_*"'] & F \arrow[d, "\pi"] \\ \neg M \arrow[r] & G \end{tikzcd}$$

Then the map π_* is an epimorphism, and so there is a diagram

$$\begin{tikzcd} & \pi^{-1}(\neg M) \arrow[d, "\pi_*"] \\ hom(\ ,C) \arrow[ur, "s'"] \arrow[r, hook] & \neg M \end{tikzcd}$$

for some representable subobject $hom(\ ,C)$ of $\neg M$ with $C \neq 0$. Finally, $hom(\ ,C) \wedge M = \phi$, so that $M \neq hom(\ ,C) \vee M$, and there is a lift

$$\begin{tikzcd} & F \arrow[d, "\pi"] \\ hom(\ ,C) \vee M \arrow[ur, "s' \vee s"] \arrow[r, hook] & G, \end{tikzcd}$$

contradicting the maximality of the lifting s. □

Generally, a map $f : X \to Y$ of presheaves on a Grothendieck site $\mathcal{C}$ is said to be a *local epimorphism* if for all sections $y \in Y(U)$, $U \in \mathcal{C}$, there is a covering sieve $R \subset hom(\ ,U)$ and elements $x_\phi \in X(V)$ for each morphism $\phi : V \to U$ in R, such that y lifts to X along R in the sense that $\phi^*(y) = f(x_\phi)$ in $Y(V)$ for all $\phi \in R$, as in the picture

$$\begin{tikzcd}
X(U) \arrow[r, "\phi^*"] \arrow[d, "f"'] & X(V) \arrow[d, "f"] & x_\phi \arrow[dd] \\
Y(U) \arrow[r, "\phi^*"'] & Y(V) & \\
y \arrow[rr] & & \phi^*(y)
\end{tikzcd}$$

In cases where there is an adequate notion of stalk, local epimorphisms are stalkwise epimorphisms: the point is that all sections should be "liftable" up to local refinement.

Examples of local epimorphisms of presheaves include all sheaf epimorphisms and the associated sheaf map $\eta : X \to L^2X$. It's easy to show that local epimorphisms are closed under composition and that a map $f : X \to Y$ is a local epimorphism if and only if the induced map $f_* : L^2X \to L^2Y$ is an epimorphism of sheaves.

There is a dual notion of local monomorphism: a map $g : A \to B$ of presheaves is a *local monomorphism* if for all $x, y \in A(U)$, $U \in \mathcal{C}$, $g(x) = g(y)$ implies that there is a covering sieve $R \subset hom(\ ,U)$ such that $\phi^*(x) = \phi^*(y) \in A(V)$ for all maps $\phi : V \to U$ in R. Again, the associated sheaf map $\eta : X \to L^2X$ is a local monomorphism, local monomorphisms are closed under composition, and a map g is a local monomorphism if and only if the induced map g_* of associated sheaves is a monomorphism of sheaves.

Now suppose that $\wp : Shv(\mathcal{B}) \to \mathcal{E}$ is a fixed Boolean localization, where $\mathcal{E} = Shv(\mathcal{C})$. This means, in particular, that the inverse image functor $\wp^* : \mathcal{E} \to Shv(\mathcal{B})$ is faithful. The functor $\wp^*$ also preserves finite limits and all colimits – this is part of the definition of a geometric morphism. The combination of these properties for $\wp^*$, together with basic exactness properties of Grothendieck topoi, has the following rather powerful consequence:

Lemma 3. *Suppose that $\wp : \mathcal{F} \to \mathcal{E}$ is a geometric morphism of Grothendieck topoi. Then the following are equivalent:*

(1) *The inverse image functor $\wp^* : \mathcal{E} \to \mathcal{F}$ is faithful.*
(2) *The functor $\wp^*$ reflects isomorphisms.*
(3) *The functor $\wp^*$ reflects epimorphisms.*
(4) *The functor $\wp^*$ reflects monomorphisms.*

Proof: Suppose that $\wp^*$ is faithful. This means that $\wp^*(f_1) = \wp^*(f_2)$ for $f_1, f_2 : A \to B$ implies that $f_1 = f_2$. Then $\wp^*$ reflects monics. In effect, suppose that $m : B \to C$ is a morphism of $\mathcal{E}$ such that $\wp^*(m)$ is a monomorphism of $\mathcal{F}$. Suppose that $m \circ f_1 = m \circ f_2$. Then $\wp^*(m)\wp^*(f_1) = \wp^*(m)\wp^*(f_2)$ implies $\wp^*(f_1) = \wp^*(f_2)$, so that $f_1 = f_2$ in $\mathcal{E}$. Similarly, $\wp^*$ reflects epimorphisms. A morphism of $\mathcal{E}$ is an isomorphism if and only if it is both a monomorphism and an epimorphism, so it follows that $\wp^*$ reflects isomorphisms.

To see that (3) implies (1), observe that the maps $f_1, f_2 : A \to B$ coincide if and only if their equalizer $m : C \to A$ is an isomorphism. Suppose that $\wp^*(f_1) = \wp^*(f_2)$. Then $\wp^*(m)$ is the equalizer of $\wp^*(f_1)$ and $\wp^*(f_2)$, by exactness of $\wp^*$, so that $\wp^*(m)$ is an isomorphism. Thus, by assumption, m is an epimorphism and hence an isomorphism, so that $f_1 = f_2$. Statement (4) implies statement (1) by a dual argument. □

2. Closed model structures.

In this section, we show that any fixed Boolean localization $\wp : Shv(\mathcal{B}) \to Shv(\mathcal{C})$ determines a class of local weak equivalences of simplicial presheaves on the site $\mathcal{C}$. We further show that this class, along with the cofibrations (or monomorphisms) of simplicial presheaves, creates closed model structures for both simplicial presheaves and simplicial sheaves on $\mathcal{C}$, in such a way that the associated homotopy theories are equivalent (Theorem 18). These closed model structures are seen to be independent of the choice of Boolean localization $\wp$ in the next section.

The definition of local weak equivalence is based on universally defined notions of local fibration and trivial local fibration for simplicial presheaves on arbitrary sites, which specialize to Kan fibrations (respectively trivial Kan fibrations) in all sections in the case of morphisms of simplicial sheaves on a complete Boolean algebra $\mathcal{B}$, via the axiom of choice. With cofibrations and local weak equivalences in hand, one defines global fibrations by a right lifting property with respect to all maps which are both cofibrations and local weak equivalences, thus effectively forcing the factorization axiom CM5 to be the non-trivial part of Theorem 18. To prove it, one shows that a map is a global fibration if and only if it has the right lifting property with respect to some *set* of trivial cofibrations (Lemma 15). These are the α-bounded trivial cofibrations, defined with respect to a cardinal number α which is sufficiently large (and in particular larger than the cardinality of the set of morphisms of $\mathcal{C}$). The most interesting part, technically, is the proof of Lemma 12.

Suppose that K is a finite simplicial set, and that Y is a simplicial presheaf on the Grothendieck site $\mathcal{C}$. Write Y^K for the presheaf defined by simplicial set morphisms in sections via the formula

$$Y^K(U) = hom_{\mathbf{S}}(K, Y(U))$$

Observe that Y^K is a sheaf if Y is a simplicial sheaf, and that any exact functor preserves this definition, so that, for example, the sheaf associated to Y^K is canonically isomorphic to $(L^2Y)^K$. Also, any geometric topos morphism preserves this construction.

One says that a map $p : X \to Y$ of simplicial presheaves is a *local fibration* if the induced maps

$$X^{\Delta^n} \xrightarrow{(i^*, p_*)} X^{\Lambda^n_k} \times_{Y^{\Lambda^n_k}} Y^{\Delta^n} \tag{1}$$

are local epimorphisms of presheaves for $n > 0$. Implicitly, a map $p : Z \to W$ of simplicial sheaves is a local fibration if and only if the maps (1) are sheaf epimorphisms. More than this is true in the Boolean topos setting:

Lemma 4. *A map $p : Z \to W$ of simplicial sheaves on a complete Boolean algebra $\mathcal{B}$ is a local fibration if and only if the induced maps in sections*

$$p : Z(b) \to W(b)$$

are Kan fibrations of simplicial sets for all $b \in \mathcal{B}$.

Proof: The sheaf epimorphisms

$$Z^{\Delta^n} \xrightarrow{(i^*,p_*)} Z^{\Lambda^n_k} \times_{W^{\Lambda^n_k}} W^{\Delta^n}$$

have sections, by Proposition 2, so that the maps

$$Z^{\Delta^n}(b) \xrightarrow{(i^*,p_*)} Z^{\Lambda^n_k}(b) \times_{W^{\Lambda^n_k}(b)} W^{\Delta^n}(b)$$

in sections are surjective, for all $b \in \mathcal{B}$. □

One says that a map $f : X \to Y$ of simplicial presheaves has the *local right lifting property with respect to the simplicial set inclusions* $\partial\Delta^n \subset \Delta^n$ if all of the maps

$$X^{\Delta^n} \xrightarrow{(i^*,f_*)} X^{\partial\Delta^n} \times_{Y^{\partial\Delta^n}} Y^{\Delta^n}$$

are local epimorphisms. One can speak, as well, about local right lifting properties with respect to more general collections of inclusions $K \subset L$ of finite simplicial sets. In particular, a local fibration is a map which has the local right lifting property with respect to all inclusions $\Lambda^n_k \subset \Delta^n$.

Suppose that X is a simplicial presheaf. The simplicial presheaf $Ex^m X$ has n-simplices defined by

$$(Ex^m X)_n = hom(sd^m\Delta^n, X).$$

with simplicial structure maps induced by precomposition with the induced simplicial set maps $sd^m\Delta^k \to sd^m\Delta^n$. The subdivision functor that we use here is the classical one: the subdivision $sd\Delta^n$ is the nerve of the poset of non-degenerate simplices of Δ^n, and the subdivision sdK of a simplicial set K is a colimit of simplicial sets $sd\Delta^m$, indexed over the simplices $\Delta^m \to K$ of K. The collection of "last vertex" maps $sd\Delta^m \to \Delta^m$, $m \geq 0$, induce a natural map $X \to ExX$, and iteration of the construction produces a sequence of simplicial presheaf maps

$$X \to ExX \to Ex^2X \to Ex^3X \to \dots$$

The simplicial presheaf $Ex^\infty X$ is defined to be the filtered colimit of these maps in the simplicial presheaf category. Write $\nu : X \to Ex^\infty X$ for the canonical map.

To put it a different way, Kan's original Ex^∞-construction [8], [2] is natural, so that it certainly applies to simplicial presheaves, and that's all we're doing here. In particular, the map $\nu : X \to Ex^\infty X$ consists of weak equivalences $\nu : X(U) \to Ex^\infty X(U)$, $U \in \mathcal{C}$, in all sections.

Now, fix a Boolean localization $\wp : Shv(\mathcal{B}) \to \mathcal{E}$, and consider the functors

$$\mathbf{S}Pre(\mathcal{C}) \xrightarrow{L^2} \mathbf{S}\mathcal{E} \xrightarrow{\wp^*} \mathbf{S}Shv(\mathcal{B})$$

relating the categories of simplicial objects in the categories $Pre(\mathcal{C})$ of presheaves on $\mathcal{C}$ and the toposes $\mathcal{E}$ and $Shv(\mathcal{B})$, where L^2 is the associated sheaf functor. We shall say that a map $f : X \to Y$ of $\mathbf{S}Shv(\mathcal{B})$ is a *pointwise weak equivalence* if it induces weak equivalences

$$f : X(b) \to Y(b)$$

of simplicial sets for all $b \in \mathcal{B}$. A map $f : X \to Y$ of simplicial presheaves on $\mathcal{C}$ is said to be a *local weak equivalence* if the induced map $\wp^*L^2(f) : \wp^*L^2Ex^\infty X \to \wp^*L^2Ex^\infty Y$ is a pointwise weak equivalence.

Remark 5. All of the decorations that appear in the definition of local weak equivalence are necessary. The categories of simplicial presheaves and sheaves on the site defined by the power set $\mathcal{P}(I)$ of an infinite set I are very good examples to keep in mind. The power set $\mathcal{P}(I)$ is, of course, a complete Boolean algebra, so that the Boolean localization $\wp$ can be taken to be the identity in this case. A simplicial presheaf X on $\mathcal{P}(I)$ is nothing more than a contravariant functor defined on the category of subsets of I, and taking values in simplicial sets. The stalks of the simplicial presheaf X are the simplicial sets $X_i = X(\{i\})$ corresponding to sections in the various singleton subsets of I, and the associated sheaf L^2X is defined in sections at a subset U of I by

$$L^2X(U) = \prod_{i\in U} X_i.$$

One says that a map $f : X \to Y$ of simplicial sheaves on $\mathcal{P}(I)$ is a stalkwise weak equivalence if all induced maps $f_i : X_i \to Y_i$, $i \in I$ are weak equivalences of simplicial sets. Observe that all induced maps in sections for the simplicial sheaf map f have the form

$$\prod f_i : \prod_{i\in U} X_i \to \prod_{i\in U} Y_i$$

for $U \subset I$. It is known that infinite products do not necessarily preserve weak equivalences (see the next paragraph), so that a stalkwise weak equivalence f of simplicial sheaves may not induce weak equivalences of simplicial sets in all sections. Infinite products do, however, preserve weak equivalences when all of the spaces X_i and Y_i are Kan complexes. The assertion that all of the X_i, $i \in I$, are Kan complexes is exactly what it means for the simplicial sheaf (or presheaf) X on $\mathcal{P}(I)$ to be locally fibrant. Thus, local weak equivalences as defined above coincide with stalkwise weak equivalences for simplicial sheaves and presheaves defined on $\mathcal{P}(I)$, and the implicit passage to locally fibrant models is fundamental.

Example 6. Here's an example of a countable collection of contractible simplicial sets X_n, $n \geq 0$, such that the product $\prod_{i\geq 0} X_i$ is not contractible. Let X_n be the subcomplex of Δ^n which is the union of the 1-simplices $\Delta^1 \subset \Delta^n$ defined by pairs of vertices $(i, i+1)$. The sequence of simplicial sets can therefore be identified with the graphs

$$0,\ 0 \to 1,\ 0 \to 1 \to 2, \dots$$

with no compositions allowed. Then the vertices $(0,0,0,0,\dots)$ and $(0,1,2,3,\dots)$ cannot be in the same path component of the product $\prod_{n\geq 0} X_n$. This observation can be expanded to a calculation of the homotopy type of the product: its path components are contractible, and two vertices $x = (x_n)$, $y = (y_n)$ of $\prod_{n\geq 0} X_n$ are in the same path component if and only if the list of combinatorial distances $d(x_n, y_n) = |y_n - x_n|$ (ie. number of 1-simplices between them in X_n) has a finite uniform bound.

Lemma 7. *Suppose, for a map $f : X \to Y$ of $\mathbf{S}Pre(\mathcal{C})$, the preheaf maps*

$$X^{\Delta^n} \xrightarrow{(i^*, f_*)} X^{\partial\Delta^n} \times_{Y^{\partial\Delta^n}} Y^{\Delta^n}$$

are local epimorphisms for $n \geq 0$. Then f is a local weak equivalence and a local fibration.

Proof: If f has the local right lifting property with respect to all $\partial\Delta^n \subset \Delta^n$, then f has the local right lifting property with respect to all inclusions of finite simplicial sets $K \subset L$. In effect, the morphisms

$$\wp^* L^2 X^{\Delta^n} \xrightarrow{(i^*, f_*)} \wp^* L^2 X^{\partial\Delta^n} \times_{\wp^* L^2 Y^{\partial\Delta^n}} \wp^* L^2 Y^{\Delta^n}$$

are sheaf epimorphisms in $Shv(\mathcal{B})$, and hence pointwise epimorphisms, so that all maps $\wp^* L^2 X(b) \to \wp^* L^2 Y(b)$ in sections are trivial Kan fibrations. But this means that the sheaf maps

$$\wp^* L^2 X^L \xrightarrow{(i^*, f_*)} \wp^* L^2 X^K \times_{\wp^* L^2 Y^K} \wp^* L^2 Y^L$$

are pointwise epimorphisms by standard nonsense about trivial Kan fibrations, and are therefore sheaf epimorphisms. It follows that the maps

$$X^L \xrightarrow{(i^*, f_*)} X^K \times_{Y^K} Y^L$$

are local epimorphisms. In particular, the map f is a local fibration.

Also, if f has the local right lifting property with respect to all $\partial\Delta^n \subset \Delta^n$, then f has the local right lifting property with respect to all induced inclusions $sd^m \partial\Delta^n \subset sd^m \Delta^n$, so that $Ex^m f : Ex^m X \to Ex^m Y$ has the local right lifting property with respect to all $\partial\Delta^n \subset \Delta^n$. But then $Ex^\infty f$ has the same local lifting property, and so does $\wp^* L^2 Ex^\infty f$. In particular, $\wp^* L^2 Ex^\infty f$ is a pointwise trivial fibration of simplicial sheaves on $\mathcal{B}$, and is therefore a weak equivalence. □

Corollary 8. *For any simplicial presheaf X, the canonical map $\eta : X \to L^2 X$ has the local right lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$, and is therefore a local fibration and a local weak equivalence.* □

Lemma 9. *Suppose that a map $f : X \to Y$ of simplicial presheaves on $\mathcal{C}$ is a pointwise weak equivalence in the sense that all maps in sections*

$$f : X(U) \to Y(U), \quad U \in \mathcal{C},$$

are weak equivalences of simplicial sets. Then f is a local weak equivalence.

Proof: The canonical map $\nu : X \to Ex^\infty X$ is a pointwise weak equivalence of simplicial presheaves, so it's enough to assume that $f : X \to Y$ is a pointwise weak equivalence of presheaves of Kan complexes, and then deduce that the map $\wp^* L^2 f : \wp^* L^2 X \to \wp^* L^2 Y$ is a pointwise weak equivalence of simplicial sheaves on $\mathcal{B}$.

Since X and Y are presheaves of Kan complexes, the classical method of replacing a map by a fibration gives a factorization

$$\begin{tikzcd} X \arrow[r, "i"] \arrow[dr, "f"'] & X \times_Y \mathbf{hom}(\Delta^1, Y) \arrow[d, "p"] \\ & Y \end{tikzcd}$$

of f in the simplicial presheaf category $\mathbf{S}Pre(\mathcal{C})$, where p is a map which is a pointwise Kan fibration and a pointwise weak equivalence, and the map i is right inverse to a map π which is a pointwise Kan fibration and a pointwise weak equivalence. The maps p and π have the local lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$, so both maps are local fibrations and local weak equivalences by Lemma 7. In particular, the maps $\wp^*L^2i$ and $\wp^*L^2p$ are pointwise weak equivalences, so that $\wp^*L^2f = (\wp^*L^2p)(\wp^*L^2i)$ is a pointwise weak equivalence. □

Corollary 10. *A map $f : X \to Y$ is a local weak equivalence of $\mathbf{S}Pre(\mathcal{C})$ if and only if $\wp^*L^2f : \wp^*L^2X \to \wp^*L^2Y$ is a local weak equivalence of $\mathbf{S}Shv(\mathcal{B})$.*

Proof: Observe that (by definition, and with respect to the Boolean localization $1 : Shv(\mathcal{B}) \to Shv(\mathcal{B})$) a map $g : Z \to W$ of $\mathbf{S}Shv(\mathcal{B})$ is a weak equivalence if and only if the map $L^2Ex^\infty g : L^2Ex^\infty Z \to L^2Ex^\infty W$ is a pointwise weak equivalence of $\mathbf{S}Shv(\mathcal{B})$

Also notice that there are natural isomorphisms

$$L^2Ex^\infty\wp^*L^2X \cong \wp^*L^2Ex^\infty X$$

for $X \in \mathbf{S}Pre(\mathcal{C})$. Thus, $L^2Ex^\infty\wp^*L^2f$ is a pointwise weak equivalence if and only if $\wp^*L^2Ex^\infty f$ is a pointwise weak equivalence. □

Generally, for a fixed property $\mathcal{P}$ of simplicial sets, one says that a map $f : X \to Y$ has the property $\mathcal{P}$ *pointwise* if each of the simplicial set maps $f : X(U) \to Y(U)$, $U \in \mathcal{C}$, in sections has the property $\mathcal{P}$. The class of pointwise weak equivalences appearing in the statement of Lemma 9 is a common example. Pointwise (Kan) fibrations and pointwise trivial fibrations also occur frequently: a map $f : X \to Y$ of simplicial presheaves is a pointwise fibration (respectively pointwise trivial fibration) if all of the maps $f : X(U) \to Y(U)$, $U \in \mathcal{C}$, are fibrations (respectively trivial fibrations) of simplicial sets. We have already met such maps in the context of simplicial presheaves on a complete Boolean algebra $\mathcal{B}$.

We shall also need the following partial converse to Lemma 7:

Lemma 11. *Suppose that X and Y are locally fibrant simplicial presheaves on $\mathcal{C}$, and that the map $q : X \to Y$ is a local fibration and a local weak equivalence. Then q has the local right lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$.*

Proof: It suffices to assume that X and Y are locally fibrant simplicial sheaves, since the associated sheaf functor L^2 preserves local fibrations and local weak equivalences, and reflects the desired local right lifting property.

The induced map

$$\wp^*L^2Ex^\infty q : \wp^*L^2Ex^\infty X \to \wp^*L^2Ex^\infty Y$$

is a pointwise weak equivalence of simplicial sheaves on $\mathcal{B}$, since q is assumed to be a local weak equivalence. There is a natural isomorphism $\wp^*L^2Ex^\infty \cong L^2Ex^\infty\wp^*$, so the map

$$L^2Ex^\infty\wp^*q : L^2Ex^\infty\wp^*X \to L^2Ex^\infty\wp^*Y$$

is a pointwise weak equivalence. The simplicial sheaf $\wp^*X$ is a presheaf of Kan complexes on $\mathcal{B}$, as is the object $Ex^\infty\wp^*X$. Furthermore, the natural map $L^2\nu$:

$\wp^* X \to L^2 Ex^\infty \wp^* X$ can be identified with the effect of applying the associated sheaf functor L^2 to the canonical pointwise weak equivalence $\nu : \wp^* X \to Ex^\infty \wp^* X$. If we can prove that the associated sheaf functor on $\mathcal{B}$ preserves pointwise weak equivalences between presheaves of Kan complexes, then we'd be done, since then $L^2\nu$ would be a pointwise weak equivalence, and so the map $\wp^* q : \wp^* X \to \wp^* Y$ would be a pointwise Kan fibration and a pointwise weak equivalence, and would therefore have the local right lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$. Finally, our faithful functor $\wp^*$ reflects this local right lifting property.

Suppose that $f : Z \to W$ is a pointwise weak equivalence between presheaves of Kan complexes on $\mathcal{B}$, and form a diagram of simplicial presheaf maps

$$\begin{tikzcd} Z \arrow[r, "i"] \arrow[dr, "f"'] & \overline{Z} \arrow[d, "\pi"] \\ & W \end{tikzcd}$$

such that π is a pointwise trivial fibration and i is right inverse to a pointwise trivial fibration $\pi' : \overline{Z} \to Z$. The associated sheaf functor L^2 preserves the local right lifting property with respect to the maps $\partial\Delta^n \subset \Delta^n$, and of course $Shv(\mathcal{B})$ satisfies the axiom of choice, so that the maps $L^2\pi$ and $L^2\pi'$ are pointwise trivial fibrations, and so $L^2 f$ is a pointwise weak equivalence. □

Pick some infinite cardinal α such that α is strictly larger than the cardinality of the set of morphisms of the site $\mathcal{C}$. A simplicial presheaf X on $\mathcal{C}$ is said to be *α-bounded* if

$$|X_n(U)| < \alpha$$

for all $n \geq 0$ and all objects U of $\mathcal{C}$. Standard cardinal arithmetic implies that if X is α-bounded, then so is its associated simplicial sheaf $L^2 X$.

Suppose that K is a simplicial set and U is an object of $\mathcal{C}$. Then the simplicial presheaf $L_U K$ is defined for $V \in \mathcal{C}$ by

$$L_U K(V) = \bigsqcup_{\phi : V \to U} K.$$

Observe that morphisms of simplicial presheaves $L_U K \to X$ are in one to one correspondence with simplicial set maps of the form $K \to X(U)$. If the simplicial set K is α-bounded in the sense that $|K_n| < \alpha$ for $n \geq 0$, then the simplicial presheaf $L_U K$ is α-bounded.

Lemma 12. *Suppose that $f : X \to Y$ is a local weak equivalence of simplicial presheaves on $\mathcal{C}$, and that pullback along f preserves α-bounded subcomplexes in the sense that if T is an α-bounded subcomplex of Y then $T \times_Y X$ is an α-bounded subcomplex of X. Suppose that there is a simplicial presheaf monomorphism $i : Z \hookrightarrow Y$ where Z is α-bounded. Then i has a factorization $Z \subset W \subset Y$ such that W is α-bounded and such that the projection map $f_* : W \times_Y X \to W$ is a local weak equivalence.*

Proof: First of all, one sees that any map of simplicial presheaves $f : X \to Y$ has a factorization

$$\begin{tikzcd} X \arrow[dr, "i_*"] \arrow[dd, "f"'] & \\ & \overline{X} \arrow[dl, "p_*"] \\ Y & \end{tikzcd}$$

such that p_* is a pointwise Kan fibration and i_* is a pointwise weak equivalence, and that this factorization is natural and preserves filtered colimits in f. In effect, take the standard factorization

$$\begin{tikzcd} Ex^\infty X \arrow[dr, "i"] \arrow[dd, "Ex^\infty f"'] & \\ & Ex^\infty X \times_{Ex^\infty Y} \mathbf{hom}(\Delta^1, Ex^\infty Y) \arrow[dl, "p"] \\ Ex^\infty Y & \end{tikzcd}$$

and pull it back to Y using the diagram

$$\begin{tikzcd} X \arrow[r, "\nu"] \arrow[d, "f"'] & Ex^\infty X \arrow[d, "Ex^\infty f"] \\ Y \arrow[r, "\nu"'] & Ex^\infty Y, \end{tikzcd}$$

so that

$$\overline{X} = Y \times_{Ex^\infty Y} Ex^\infty X \times_{Ex^\infty Y} \mathbf{hom}(\Delta^1, Ex^\infty Y).$$

Note finally that if X and Y are α-bounded simplicial presheaves, then so is $\overline{X}$.

The pulled back map p_* has the local right lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$, since Lemma 9 implies that p is a local weak equivalence as well as a pointwise fibration, so that Lemma 11 applies.

This means explicitly that given any diagram of simplicial set maps of the form

$$\begin{tikzcd} \partial\Delta^n \arrow[r, "a"] \arrow[d, hook] & \overline{X}(U) \arrow[d, "p_*"] \\ \Delta^n \arrow[r, "b"'] & Y(U), \end{tikzcd}$$

there is a covering sieve $R \subset hom(\ ,U)$ such that for each $\phi : V \to U$ in R, there is a commutative diagram

$$\begin{tikzcd} \partial\Delta^n \arrow[r, "\phi^* a"] \arrow[d, hook] & \overline{X}(V) \arrow[d, "p_*"] \\ \Delta^n \arrow[r, "\phi^* b"'] \arrow[ur, "x_\phi"] & Y(V). \end{tikzcd}$$

Suppose given a diagram

$$
\begin{tikzcd}
\partial\Delta^n \arrow[r] \arrow[d, hook] & \overline{Z \times_Y X}(U) \arrow[r] \arrow[d, "p_*"] & \overline{X}(U) \arrow[d, "p_*"] \\
\Delta^n \arrow[r] & Z(U) \arrow[r] & Y(U).
\end{tikzcd}
\tag{2}
$$

Then $\overline{X}$ is a filtered colimit of simplicial presheaves of the form $\overline{T \times_Y X}$, where T is an α-bounded subobject of Y containing Z. It follows that there is an α-bounded S containing Z such that all the liftings x_ϕ corresponding to the outer square live in $\overline{S \times_Y X}$. Taking the union of all such subcomplexes S over the α-bounded set of diagrams of the form (2) gives an α-bounded subcomplex Z_1 of Y such that $Z \subset Z_1 \subset Y$, and such that all local lifting problems (2) are solved in $\overline{Z_1 \times_Y X}$. Repeat the construction to obtain a sequence of α-bounded subobjects

$$Z = Z_0 \subset Z_1 \subset Z_2 \subset \dots$$

such that all local lifting problems

$$
\begin{tikzcd}
\partial\Delta^n \arrow[r] \arrow[d, hook] & \overline{Z_i \times_Y X}(U) \arrow[d, "p_*"] \\
\Delta^n \arrow[r] & Z_i(U)
\end{tikzcd}
$$

are solved over Z_{i+1}.

Let $W = \cup_i Z_i$. Then the map $p_* : \overline{W \times_Y X} \to W$ is a local weak equivalence by Lemma 7. Furthermore, $f_* : W \times_Y X \to W$ is a composite

$$
\begin{tikzcd}
W \times_Y X \arrow[r, "i_*"] \arrow[dr, "f_*"'] & \overline{W \times_Y X} \arrow[d, "p_*"] \\
& W,
\end{tikzcd}
$$

where i_* is a pointwise weak equivalence. The map i_* is therefore a local weak equivalence by Lemma 9, so that f_* is also a local weak equivalence. □

Corollary 13. *Suppose that $f : X \to Y$ is a local weak equivalence of simplicial sheaves which satisfies the boundedness condition of Lemma 12, and that there is a simplicial sheaf monomorphism $i : Z \hookrightarrow Y$ where Z is α-bounded. Then i has a factorization $Z \subset W \subset Y$ such that W is α-bounded and such that the projection map $f_* : W \times_Y X \to W$ is a local weak equivalence.*

Proof: Apply the associated sheaf functor to the output of Lemma 12. □

A *cofibration* of simplicial presheaves is a monomorphism $A \hookrightarrow B$ of simplicial presheaves. A map of simplicial presheaves which is both a cofibration and a local weak equivalence is called a *trivial cofibration.* A *global fibration* is a morphism $p : X \to Y$ of simplicial presheaves which has the right lifting property with respect to all trivial cofibrations. Finally a map which is simultaneously a global fibration and a local weak equivalence is said to be a *trivial global fibration.*

Lemma 14.

(1) *Trivial cofibrations of simplicial presheaves are closed under pushout.*

(2) *Suppose that γ is an limit ordinal, thought of as a poset, and that there is a functor $X : \gamma \to \mathbf{S}Pre(\mathcal{C})$ such that for each morphism $i \leq j$ of γ, the induced map $X(i) \to X(j)$ is a trivial cofibration. Then the canonical maps*

$$X(i) \xrightarrow{\tau_i} \varinjlim_{j\in\gamma} X(j)$$

are trivial cofibrations.

(3) *Suppose that the morphisms $f_i : X_i \to Y_i$ are local weak equivalences for $i \in I$. Then the morphism*

$$\bigsqcup_{i\in I} f_i : \bigsqcup_{i\in I} X_i \to \bigsqcup_{i\in I} Y_i$$

is a local weak equivalence.

Proof: It suffices, by Corollary 10 and Corollary 8, to prove all three statements for the category $\mathbf{S}Pre(\mathcal{B})$ of simplicial presheaves on the complete Boolean algebra $\mathcal{B}$.

For statement (1), suppose that the diagram

$$\begin{tikzcd} A \arrow[r] \arrow[d, "i"'] & C \arrow[d, "i_*"] \\ B \arrow[r] & B \cup_A C \end{tikzcd}$$

is a pushout of simplicial presheaves on $\mathcal{B}$, where i is a cofibration and a local weak equivalence. To show that i_* is a local weak equivalence, it suffices to show that the map i' in the pushout diagram of simplicial presheaves

$$\begin{tikzcd} L^2Ex^\infty A \arrow[r] \arrow[d, "L^2Ex^\infty i"'] & L^2Ex^\infty C \arrow[d, "i'"] \\ L^2Ex^\infty B \arrow[r] & L^2Ex^\infty B \cup_{L^2Ex^\infty A} L^2Ex^\infty C \end{tikzcd}$$

is a local weak equivalence. To see this, one invokes the ordinary patching lemma for simplicial sets and Corollary 8. But then the map i' is a pointwise weak equivalence since $L^2Ex^\infty i$ is a pointwise weak equivalence, so we're done.

For (2), let $X : \gamma \to \mathbf{S}Pre(\mathcal{B})$ be a functor as in the statement, and form a new functor $Ex^\infty X$ with $Ex^\infty X(i)$ defined in the obvious way for $i \in \gamma$, and consider the natural transformation $\nu : X \to Ex^\infty X$ arising from the pointwise weak equivalences

$$\nu : X(i) \to Ex^\infty X(i), \quad i \in \gamma.$$

Then each morphism $i \leq j$ of γ induces a trivial cofibration $Ex^\infty X(i) \to Ex^\infty X(j)$ by Lemma 9, and there is a commutative diagram

$$\begin{tikzcd} X(i) \arrow[r, "\tau_i"] \arrow[d, "\nu"'] & \varinjlim X(i) \arrow[d, "\nu_*"] \\ Ex^\infty X(i) \arrow[r, "\tau_i"'] & \varinjlim Ex^\infty X(i) \end{tikzcd}$$

where the filtered colimits are formed in the presheaf category, so that ν_* is a pointwise weak equivalence. It follows from Lemma 9 that one instance of τ_i in the diagram is a local weak equivalence if and only if the other is, so it suffices to assume that each of the simplicial presheaves $X(i)$ is a presheaf of Kan complexes.

Now suppose that X is a diagram of presheaves of Kan complexes, and form the diagram

$$\begin{tikzcd} X(i) \arrow[r, "\tau_i"] \arrow[d, "\eta"'] & \varinjlim X(i) \arrow[r, "\eta"] \arrow[d, "\eta_*"] & L^2(\varinjlim X(i)) \arrow[d, "\cong"] \\ L^2X(i) \arrow[r, "\tau_i"'] & \varinjlim L^2X(i) \arrow[r, "\eta"'] & L^2(\varinjlim L^2X(i)) \end{tikzcd}$$

which is induced the comparison transformation $\eta : X \to L^2X$ induced by the associated sheaf construction. The induced morphisms $L^2X(i) \to L^2X(j)$ are local weak equivalences of locally fibrant simplicial sheaves on the complete Boolean algebra $\mathcal{B}$, and are therefore pointwise weak equivalences, so that the simplicial presheaf maps $\tau_i : L^2X(i) \to \varinjlim L^2X(i)$ are pointwise weak equivalences and therefore local weak equivalences, by Lemma 9. The associated sheaf maps η are local weak equivalences by Corollary 8, so that the original maps $\tau_i : X(i) \to \varinjlim X(i)$ are local weak equivalences as well.

In the case of statement (3), the Ex^∞ construction preserves disjoint unions of simplicial sets, so it suffices to assume that the simplicial presheaves X_i and Y_i are presheaves of Kan complexes. In that case, the induced morphisms $L^2f_i : L^2X_i \to L^2Y_i$ are local weak equivalences of locally fibrant simplicial sheaves on $\mathcal{B}$, so that they are all pointwise weak equivalences. It follows that the induced morphism

$$\bigsqcup_{i\in I} L^2X_i \xrightarrow{\bigsqcup L^2f_i} \bigsqcup_{i\in I} L^2Y_i$$

are pointwise and hence local weak equivalences. One finishes by observing that there is a natural commutative diagram

$$\begin{tikzcd} \bigsqcup_{i\in I} X_i \arrow[r, "\bigsqcup_i \eta"] \arrow[d, "\eta"'] & \bigsqcup_{i\in I} L^2X_i \arrow[d, "\eta"] \\ L^2(\bigsqcup_{i\in I} X_i) \arrow[r, "\cong"'] & L^2(\bigsqcup_{i\in I} L^2X_i) \end{tikzcd}$$

in the category of simplicial presheaves on $\mathcal{B}$, so that the morphism $\bigsqcup_i \eta$ is a natural local weak equivalence by Corollary 8. □

A cofibration $A \hookrightarrow B$ of simplicial presheaves is said to be an *α-bounded cofibration* if the target simplicial presheaf B is α-bounded.

Lemma 15. *A map $p : X \to Y$ of simplicial presheaves is a global fibration if and only if it has the right lifting property with respect to all α-bounded trivial cofibrations.*

Proof: Suppose that p has the right lifting property with respect to all α-bounded trivial cofibrations, and consider the diagram

$$\begin{tikzcd} U \arrow[r] \arrow[d, "i"'] & X \arrow[d, "p"] \\ V \arrow[r] & Y, \end{tikzcd}$$

where i is a trivial cofibration. We shall assume that $U \neq V$. Consider the set of all partial lifts

$$\begin{tikzcd} U \arrow[r] \arrow[d, hook, "j"'] & X \arrow[dd, "p"] \\ U' \arrow[ur] \arrow[d, hook, "i'"'] & \\ V \arrow[r] & Y \end{tikzcd} \tag{3}$$

where $i'j = i$, $U' \neq U$, and j is a trivial cofibration. This set is non-trivial: take $x \in V(W) - U(W)$ for some $W \in \mathcal{C}$, and observe that x sits inside some α-bounded subcomplex C of V, namely the image of the map $L_W\Delta^m \to V$ which classifies x. By Lemma 12, there is an α-bounded subcomplex $B \subset V$ such that $C \subset B$ and such that the induced cofibration $j : B \cap U \hookrightarrow B$ is a local weak equivalence. Form the diagram

$$\begin{tikzcd} B \cap U \arrow[r] \arrow[d, "j"'] & U \arrow[r] \arrow[d, hook, "j_*"'] & X \arrow[dd, "p"] \\ B \arrow[r] & B \cup U \arrow[ur, dashed] \arrow[d, hook, "i'"'] & \\ & V \arrow[r] & Y \end{tikzcd}$$

and observe that the indicated lift exists because j_* is a pushout of the α-bounded trivial cofibration j. Then x is a section of $B \cup U$, so that $B \cup U \neq U$. Furthermore, j_* is a trivial cofibration: this is a consequence of Lemma 14.

The set of all partial lifts has maximal elements, by Zorn's lemma and part (2) of Lemma 14. Any maximal element must be a lift

$$\begin{tikzcd} U \arrow[r] \arrow[d, "i"'] & X \arrow[d, "p"] \\ V \arrow[r] \arrow[ur] & Y \end{tikzcd}$$

by the argument (applied to maps of the form i' in (3)) that is used to demonstrate the existence of partial lifts. □

Lemma 16. *Every simplicial presheaf map $f : X \to Y$ has a factorization*

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "i"'] & & Y \\ & Z \arrow[ur, "p"'] & \end{tikzcd}$$

where p is a global fibration and i is a trivial cofibration.

Proof: The proof is a transfinite small object argument.

Take a cardinal $\beta > 2^{\alpha}$, where α is the cardinality of the set of morphisms of the site $\mathcal{C}$. We define a functor $F : \beta \to \mathbf{S}Pre(\mathcal{C}) \downarrow Y$ by first setting $F(0) = f : X \to Y$. We let

$$X(\zeta) = \varinjlim_{\gamma<\zeta} X(\gamma)$$

for limit ordinals ζ. Finally, the map $X(\gamma) \to X(\gamma+1)$ is defined by taking the set of all diagrams

$$D : \quad \begin{tikzcd} U_D \arrow[r] \arrow[d, "i_D"'] & X(\gamma) \arrow[d, "F(\gamma)"] \\ V_D \arrow[r] & Y \end{tikzcd}$$

such that i_D is an α-bounded trivial cofibration, and then forming the pushout

$$\begin{tikzcd} \bigsqcup_D U_D \arrow[r] \arrow[d, "\bigsqcup_D i_D"'] & X(\gamma) \arrow[d, "i_*"] \\ \bigsqcup_D V_D \arrow[r] & X(\gamma+1) \end{tikzcd}$$

Then i_* is a trivial cofibration, by Lemma 14, as is the map i_β in the resulting diagram

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "i_\beta"'] & & Y \\ & X(\beta), \arrow[ur, "F(\beta)"'] & \end{tikzcd}$$

where $X(\beta) = \varinjlim_{\gamma<\beta} X(\gamma)$, and $F(\beta)$ is induced by all maps $F(\gamma)$. In any diagram

$$\begin{tikzcd} U \arrow[r, "g"] \arrow[d, "i"'] & X(\beta) \arrow[d, "F(\beta)"] \\ V \arrow[r] \arrow[ur, dashed] & Y \end{tikzcd}$$

where i is a trivial α-bounded cofibration, the simplicial presheaf U is α-bounded, so that g must factor through some subcomplex $X(\gamma) \subset X(\beta)$ with $\gamma < \beta$. It follows that the dotted arrow exists, making the diagram commute. □

Lemma 17. *Any simplicial presheaf map $f : X \to Y$ has a factorization*

$$\begin{tikzcd} X \arrow[rr, "f"] \arrow[dr, "j"'] & & Y \\ & W \arrow[ur, "q"'] & \end{tikzcd}$$

where q is a trivial global fibration and j is a cofibration.

Proof: First of all, if a map $f : X \to Y$ has the right lifting property with respect to all morphisms of the form $A \subset L_U\Delta^n$, then f is a global fibration and a local weak equivalence. In effect, f has the right lifting property with respect to all cofibrations by an argument similar to that of Lemma 15, so that f is a global fibration, and f has the right lifting property with respect to all cofibrations of the form $L_U\partial\Delta^n \subset L_U\Delta^n$, so that f is a pointwise weak equivalence and hence a local weak equivalence by Lemma 9.

The existence of the required factorization is now a consequence of a transfinite small object argument similar to that given for Lemma 16. □

Theorem 18. *With respect to the definitions of cofibration, weak equivalence and global fibration given above,*

(1) *the category $\mathbf{S}Pre(\mathcal{C})$ of simplicial presheaves is a closed model category,*
(2) *the category $\mathbf{S}Shv(\mathcal{C})$ is a closed model category,*
(3) *the inclusion $\mathbf{S}Shv(\mathcal{C}) \subset \mathbf{S}Pre(\mathcal{C})$ induces an equivalence*

$$Ho(\mathbf{S}Shv(\mathcal{C})) \simeq Ho(\mathbf{S}Pre(\mathcal{C}))$$

of the associated homotopy categories.

Proof: The only non-trivial parts of the respective demonstrations are the factorization axiom and CM4, for both simplicial presheaves and simplicial sheaves. But the factorization axioms follow from Lemma 16 and Lemma 17, and their simplicial sheaf counterparts (which have the same arguments), and CM4 is a consequence of the assertion that every trivial global fibration has the right lifting property with respect to all cofibrations, for both categories.

For the latter, observe that if $p : X \to Y$ is a global fibration and a local weak equivalence, then the proof of Lemma 17 shows that p has a factorization

$$\begin{tikzcd} X \arrow[rr, "p"] \arrow[dr, "j"'] & & Y \\ & W, \arrow[ur, "q"'] & \end{tikzcd}$$

where j is a cofibration and q has the right lifting property with respect to all cofibrations and is a local weak equivalence. But then j is a trivial cofibration, so that there is a commutative diagram

$$\begin{tikzcd} X \arrow[r, "1_X"] \arrow[d, "j"'] & X \arrow[d, "p"] \\ W \arrow[r, "q"'] \arrow[ur] & Y, \end{tikzcd}$$

and so p is a retract of q.

The equivalence of categories

$$Ho(\mathbf{S}Shv(\mathcal{C})) \simeq Ho(\mathbf{S}Pre(\mathcal{C}))$$

is induced by the inclusion $\mathbf{S}Shv(\mathcal{C}) \subset \mathbf{S}Pre(\mathcal{C})$ and its left adjoint, the associated sheaf functor $L^2 : \mathbf{S}Pre(\mathcal{C}) \to \mathbf{S}Shv(\mathcal{C})$. Both of these functors preserve local weak equivalences, and the canonical simplicial presheaf map $X \to L^2X$ is a weak equivalence, by Corollary 8. □

Suppose that X is a simplicial presheaf and that K is a simplicial set. There is a simplicial presheaf $\mathbf{hom}(K, X)$, which is defined in sections by

$$\mathbf{hom}(K, X)(U) = \mathbf{hom}(K, X(U)), \quad U \in \mathcal{C},$$

where $\mathbf{hom}(K, X(U))$, denotes the ordinary function space object in the category of simplicial sets. The simplicial presheaf $\mathbf{hom}(\Delta^1, X)$ is the path object that was used in the proof of Lemma 9.

The ordinary exponential law for simplicial sets induces a natural isomorphism of the form

$$hom(X, \mathbf{hom}(K, Y)) \cong hom(X \times K, Y),$$

where the indicated morphisms are in the category of simplicial presheaves, and $X \times K$ is the simplicial presheaf defined in sections by

$$(X \times K)(U) = X(U) \times K, \quad U \in \mathcal{C}.$$

The main homotopical result about function spaces of this type is the following:

Lemma 19. *Suppose that $q : X \to Y$ is a local fibration of simplicial presheaves on $\mathcal{C}$, and that $i : K \hookrightarrow L$ is a cofibration of simplicial sets where L is finite in the sense that it has only finitely many non-degenerate simplices. Then the induced simplicial presheaf map*

$$\mathbf{hom}(L, X) \xrightarrow{(i^*, q_*)} \mathbf{hom}(K, X) \times_{\mathbf{hom}(K,Y)} \mathbf{hom}(L, Y)$$

is a local fibration, and this map is a local weak equivalence if q is a local weak equivalence or if i is a trivial cofibration of simplicial sets.

Proof: There is a natural isomorphism

$$\wp^* L^2 \mathbf{hom}(K, X) \cong \mathbf{hom}(K, \wp^* L^2 X)$$

for all finite simplicial sets K and simplicial presheaves X, since the associated sheaf functor L^2 and the Boolean localization functor $\wp^*$ both preserve finite limits. The map $\wp^* L^2 q : \wp^* L^2 X \to \wp^* L^2 Y$ is a pointwise Kan fibration, so that the map

$$\mathbf{hom}(L, \wp^* L^2 X) \xrightarrow{(i^*, \wp^* L^2 q_*)} \mathbf{hom}(K, \wp^* L^2 X) \times_{\mathbf{hom}(K, \wp^* L^2 Y)} \mathbf{hom}(L, \wp^* L^2 Y)$$

is a pointwise Kan fibration, which is a pointwise weak equivalence if i is a trivial cofibration or if $\wp^* L^2 q$ is pointwise trivial. □

Corollary 20. *Suppose that X is a locally fibrant simplicial presheaf, and that $i : K \hookrightarrow L$ is a cofibration of finite simplicial sets. Then the induced map*

$$i^* : \mathbf{hom}(L, X) \to \mathbf{hom}(K, X)$$

is a local fibration. The map i^* is a local weak equivalence if i is a trivial cofibration. □

Remark 21. Corollary 20 is the central device behind the path object and associated local fibration constructions that appear in the proof of Lemma 9.

Suppose that X and Y are simplicial presheaves. The *function space* $\mathbf{hom}(X, Y)$ is the simplicial set defined by having n-simplices

$$\mathbf{hom}(X, Y)_n = hom(X \times \Delta^n, Y),$$

where the morphism set on the right is in the category of simplicial presheaves. The standard exponential law for the simplicial set category also induces a natural isomorphism

$$hom(X \times K, Y) \cong hom_{\mathbf{S}}(K, \mathbf{hom}(X, Y)),$$

so that the category of simplicial presheaves acquires the structure of a simplicial category in the sense of Quillen.

Similar observations obtain for the category of simplicial sheaves on $\mathcal{C}$. In that case, one writes $X \otimes K = L^2(X \times K)$ for $X \in \mathbf{S}Shv(\mathcal{C})$. Then, if Y is a simplicial sheaf, there is an isomorphism

$$hom(X \otimes K, Y) \cong hom_{\mathbf{S}}(K, \mathbf{hom}(X, Y)),$$

so that the category of simplicial sheaves on $\mathcal{C}$ also has the structure of a simplicial category.

Lemma 22. *Suppose that $i : A \to B$ is a cofibration and $q : X \to Y$ is a global fibration of simplicial presheaves. Then the induced simplicial set map*

$$\mathbf{hom}(B, X) \xrightarrow{(i^*, q_*)} \mathbf{hom}(A, X) \times_{\mathbf{hom}(A,Y)} \mathbf{hom}(B, Y)$$

is a Kan fibration which is trivial if either i or q is a local weak equivalence.

Proof: The map

$$(B \times \Lambda^n_k) \cup_{(A \times \Lambda^n_k)} (A \times \Delta^n) \subset B \times \Delta^n$$

is a cofibration and a pointwise weak equivalence; it is therefore a local weak equivalence by Lemma 9. Finish the argument that (i^*, q_*) is a Kan fibration with the standard adjointness trick [11], [2].

It remains to show that the cofibration

$$(B \times \partial\Delta^n) \cup_{(A \times \partial\Delta^n)} (A \times \Delta^n) \subset B \times \Delta^n$$

is a local weak equivalence in the case where the cofibration $i : A \to B$ is a local weak equivalence.

One should know first that the functor $X \mapsto X \times K$ preserves local weak equivalences in simplicial presheaves X, for all simplicial sets K. For this, there are natural local equivalences

$$\wp^* L^2(X \times K) \xrightarrow{\wp^* L^2(\nu \times K)} \wp^* L^2(Ex^\infty X \times K) \cong L^2(\wp^* L^2 Ex^\infty X \times K).$$

The functor $X \mapsto \wp^* L^2 Ex^\infty X \times K$ takes local weak equivalences to pointwise weak equivalences, and so the desired result follows from Corollary 10.

It follows that, in the diagram

$$\begin{tikzcd} A \times \Delta^n \arrow[d, "i_*"'] \arrow[dr, "i \times \Delta^n"] & \\ (B \times \partial\Delta^n) \cup_{(A \times \partial\Delta^n)} (A \times \Delta^n) \arrow[r, hook] & B \times \Delta^n, \end{tikzcd}$$

the maps $i \times \Delta^n$ and i_* are trivial cofibrations. ◻

There is a corresponding statement about simplicial sheaves, which is an immediate corollary of Lemma 22.

Corollary 23. *The simplicial presheaf category* $\mathbf{S}Pre(\mathcal{C})$ *and the simplicial sheaf category* $\mathbf{S}Shv(\mathcal{C})$ *are both closed simplicial model categories.*

One says that a closed model category is *proper* if weak equivalences are preserved by pullback along fibrations and by pushout along cofibrations.

Theorem 24. *The simplicial presheaf category* $\mathbf{S}Pre(\mathcal{C})$ *and the simplicial sheaf category* $\mathbf{S}Shv(\mathcal{C})$ *are proper closed simplicial model categories.*

Properness is very commonly used: it is fundamental to all patching lemmas [2, 2.8], and is essential for constructing stable homotopy theories for simplicial presheaves [1], [5], [6].

Proof of Theorem 24: Suppose that the diagram of simplicial presheaf morphisms

$$\begin{tikzcd} Z \times_Y X \arrow[r, "g_*"] \arrow[d] & X \arrow[d, "q"] \\ Z \arrow[r, "g"'] & Y \end{tikzcd}$$

is a pullback with q a global fibration and g a local weak equivalence. To show that g_* is a local weak equivalence, it suffices, by Corollary 8 and Corollary 10, to assume that X, Y, and Z are simplicial sheaves on the complete Boolean algebra $\mathcal{B}$. By Corollary 8 and exactness, applying the composite functor $L^2 Ex^\infty$ doesn't change the problem, so it suffices to assume that X, Y and Z are locally fibrant simplicial sheaves on $\mathcal{B}$. But then g is a pointwise weak equivalence, so that g_* is a pointwise, hence local, weak equivalence.

Suppose given a pushout diagram

$$\begin{tikzcd} A \arrow[r, "f"] \arrow[d, hook, "i"'] & X \arrow[d] \\ B \arrow[r, "f_*"'] & B \cup_A X \end{tikzcd}$$

with i a cofibration and f a local weak equivalence. By the patching lemma for simplicial sets, Corollary 8 and Corollary 10, it suffices to assume that A, B and X are locally fibrant simplicial sheaves on $\mathcal{B}$ and that the pushout is formed in the category of simplicial presheaves on $\mathcal{B}$. In that case, f is a pointwise weak equivalence, so that f_* is a pointwise hence local weak equivalence, by Lemma 9. □

3. Homotopy groups.

Traditionally, weak equivalences of simplicial sheaves and presheaves have been defined via sheaves of homotopy groups, which we haven't even mentioned yet. We have so far used a definition of weak equivalence that appears to depend on a fixed Boolean localization $\wp : Shv(\mathcal{B}) \to \mathcal{E} = Shv(\mathcal{C})$. In this section we will show that this apparent dependence on $\wp$ can be removed by introducing a notion of fibred homotopy group objects which is preserved by the inverse image functor $\wp^*$ and specializes to the standard homotopy groups for ordinary simplicial sets over all vertices (but see also Remark 28 below). These homotopy group objects are made up of sheaves of homotopy groups in the usual sense, and our definition of weak equivalence is seen to coincide with the familiar one.

Suppose that X is a Kan complex, with base point x. The set underlying the homotopy group $\pi_n(X, x)$ can be identified with the set of path components $\pi_0 F_{n,x}X$, where the $F_{n,x}X$ is defined by the pullback diagram

$$\begin{tikzcd} F_{n,x}X \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^n, X) \arrow[d, "i^*"] \\ * \arrow[r, "x"'] & \mathbf{hom}(\partial\Delta^n, X) \end{tikzcd}$$

and i^* is the fibration induced by the inclusion $i : \partial\Delta^n \hookrightarrow \Delta^n$. Note, in particular, that $F_{n,x}X$ is a Kan complex, so that $\pi_0 F_{n,x}X$ can be identified with a set of homotopy classes of vertices.

To collect all such definitions together, use the notation X_0 for the discrete simplicial set $\bigsqcup_{x \in X_0} *$ on the set of vertices of X as well as for the set of vertices itself, and form the pullback

$$\begin{tikzcd} F_nX \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^n, X) \arrow[d, "i^*"] \\ X_0 \arrow[r] & \mathbf{hom}(\partial\Delta^n, X), \end{tikzcd}$$

where the map $X_0 \to \mathbf{hom}(\partial\Delta^n, X)$ takes the vertex x to the map $x : \partial\Delta^n \to X$ which factors through x. The simplicial set X_0 is a Kan complex, so that

$$F_nX \cong \bigsqcup_{x \in X_0} F_{n,x}X$$

is a Kan complex fibred over X_0, and we write

$$\pi_nX = \pi_0F_nX = \bigsqcup_{x \in X_0} \pi_0F_{n,x}X = \bigsqcup_{x \in X_0} \pi_n(X, x).$$

There is a canonical function

$$\pi_n X = \pi_0 F_n X \to X_0$$

which gives $\pi_n X$ a fibred structure over the set of vertices X_0.

To see the group multiplication, let $\Lambda^{[0,n-2]} \subset \Delta^{n+1}$ be the subcomplex which is generated by the simplices $d^i : \mathbf{n} \to \mathbf{n+1}$, $0 \le i \le n-2$, and write $K_n = \Lambda^{[0,n-2]} \cup sk_{n-1}\Delta^{n+1}$ let j denote the inclusion $K_n \subset \Delta^{n+1}$. Form the pullback diagram

$$\begin{tikzcd} G_n X \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^{n+1}, X) \arrow[d, "j^*"] \\ X_0 \arrow[r] & \mathbf{hom}(K_n, X) \end{tikzcd}$$

in the category of simplicial sets. The maps $d^i : \Delta^n \to \Delta^{n+1}$ induce morphisms $d_i : G_n X \to F_n X$ of spaces fibred over X_0 for $n-1 \le i \le n+1$. Furthermore, the induced map $(d_{n-1}, d_{n+1}) : G_n X \to F_n X \times_{X_0} F_n X$ is surjective, since it is induced by pulling back a trivial fibration $\mathbf{hom}(\Delta^{n+1}, X) \to \mathbf{hom}(\Delta^n \times_{\Delta^{n-1}} \Delta^n, X)$. By looking at vertices and taking path components one sees, via the standard constructions, that there is a unique map $m : \pi_0 F_n X \times_{X_0} \pi_0 F_n X \to \pi_0 F_n X$ of objects fibred over X_0 making the following diagram commute:

$$\begin{tikzcd} G_n X_0 \arrow[r, "d_{n*}"] \arrow[d, two heads, "(d_{n-1}, d_{n+1})_*"'] & F_n X_0 \arrow[r, two heads] & \pi_0 F_n X \\ (F_n X \times_{X_0} F_n X)_0 \arrow[d, two heads] & & \\ \pi_0 F_n X \times_{X_0} \pi_0 F_n X \arrow[uurr, "m"'] & & \end{tikzcd}$$

Observe that the map m can be identified with the map

$$\bigsqcup_{x \in X_0} \pi_n(X, x) \times \pi_n(X, x) \to \bigsqcup_{x \in X_0} \pi_n(X, x)$$

that one obtains by collecting all of the ordinary homotopy group multiplication maps together.

The group inverse $\sigma : \pi_n X \to \pi_n X$ is defined as a fibred map over X_0 by letting $\Lambda^{n+1}_{n-1,n+1}$ be the subcomplex of Δ^{n+1} generated by the simplices d^i for $i \neq n-1, n+1$, and forming the pullback

$$\begin{tikzcd} H_n X \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^{n+1}, X) \arrow[d] \\ X_0 \arrow[r] & \mathbf{hom}(\Lambda^{n+1}_{n-1,n+1}, X) \end{tikzcd}$$

The maps d_{n-1} and d_{n+1} induce functions $d_{n-1*}, d_{n+1*} : H_n X \to F_n X$ of spaces fibred over X_0, and both of these maps are surjective because they are induced by trivial fibrations of the form $(d^j)^* : \mathbf{hom}(\Delta^{n+1}, X) \to \mathbf{hom}(\Delta^n, X)$. Then σ is the unique map of sets fibred over X_0 which makes the following diagram commute:

$$\begin{tikzcd} H_nX_0 \arrow[r, two heads, "d_{n+1*}"] \arrow[d, two heads, "d_{n-1*}"'] & F_nX_0 \arrow[r, two heads] & \pi_0F_nX \\ F_nX_0 \arrow[d, two heads] & & \\ \pi_0F_nX \arrow[uurr, "\sigma"'] & & \end{tikzcd}$$

Again, the map σ can be identified with the map

$$\bigsqcup_{x\in X_0} \pi_n(X,x) \to \bigsqcup_{x\in X_0} \pi_n(X,x)$$

which consists of the group inverses for the regular homotopy groups.

The identity $e : X_0 \to \pi_nX$ is the section of the structure map $\pi_nX \to X_0$ which is induced by the canonical section of the simplicial set map $F_nX \to X_0$. Of course e specializes to the map $* \to \pi_n(X,x)$ which picks out the identity map of the group $\pi_n(X,x)$ over each summand of X_0.

The defining axioms for the group structures on the various $\pi_n(X,x)$ can now be used to show that the fibred objects $\pi_nX \to X_0$, together with the multiplication map m, the inverse map σ and the identity section e, give π_nX the structure of a group object in the category of sets fibred over X_0. This group object is abelian if $n \geq 2$. The existence of the group object isn't news by itself, but the descriptions of the maps m, σ and e are combinatorial and functorial, and are therefore more broadly applicable.

Observe that a map $f : X \to Y$ of Kan complexes is a weak equivalence if and only if

(1) the induced map $f_* : \pi_0X \to \pi_0Y$ of path components is a bijection, and
(2) the induced diagrams

$$\begin{tikzcd} \pi_nX \arrow[r, "f_*"] \arrow[d] & \pi_nY \arrow[d] \\ X_0 \arrow[r, "f"'] & Y_0 \end{tikzcd}$$

are pullbacks for $n \geq 1$.

This is easily verified, given that the displayed group objects consist of ordinary homotopy groups.

Suppose that Y is a simplicial presheaf, and define a presheaf π_0^pY by forming the coequalizer diagram

$$\begin{tikzcd} Y_1 \arrow[r, shift left, "d_0"] \arrow[r, shift right, "d_1"'] & Y_0 \arrow[r, "c"] & \pi_0^pY \end{tikzcd}$$

in the presheaf category. Let π_0Y denote the associated sheaf for π_0^p; one oftens says that π_0Y is the *sheaf of path components* of Y. Observe that the canonical map $Y \to L^2Y$ from Y to its associated sheaf induces an isomorphism $\pi_0Y \cong \pi_0L^2Y$.

Lemma 25. *Suppose that X is a locally fibrant simplicial sheaf on a complete Boolean algebra $\mathcal{B}$. Then the associated sheaf map*

$$\eta : \pi_0^p X \to \pi_0 X$$

is an isomorphism of presheaves.

Proof: The locally fibrant simplicial sheaf X is a presheaf of Kan complexes, by Lemma 4. It follows that the canonical presheaf map

$$X_1 \xrightarrow{(d_1,d_0)} X_0 \times_{\pi_0^p X} X_0$$

is a pointwise epimorphism.

Form the comparison diagram

$$\begin{tikzcd} X_1 \arrow[r, two heads, "{(d_1, d_0)}"] \arrow[dr, dashed] & X_0 \times_{\pi_0^p X} X_0 \arrow[r, shift left] \arrow[r, shift right] \arrow[d, "\eta"'] & X_0 \arrow[r, "c"] \arrow[d, "1_{X_0}"'] & \pi_0^p X \arrow[d, "\eta"'] \\ & X_0 \times_{\pi_0 X} X_0 \arrow[r, shift left] \arrow[r, shift right] & X_0 \arrow[r, "L^2 c"'] & \pi_0 X \end{tikzcd}$$

The bottom sequence is a coequalizer in the sheaf category, while the top sequence is a coequalizer in the presheaf category.

The sheaf epimorphism L^2c is a pointwise epimorphism, by the axiom of choice (Proposition 2), so that the canonical presheaf map $\eta : \pi_0^p X \to \pi_0 X$ is also a pointwise epi. The composite map displayed by the dotted arrow can be identified with the sheaf map associated to the presheaf epimorphism (d_0, d_1), so it's a sheaf epi and hence a pointwise epi, again by the axiom of choice. If $L^2c(x) = L^2c(y)$ in $\pi_0 X$, then (x, y) defines an element of $X_0 \times_{\pi_0 X} X_0$, and so there is a section z of X_1 which maps to (x, y) under the dotted composite. But then $x = d_1 z$ and $y = d_0 z$, so that x and y represent the same element of $\pi_0^p X$. The associated sheaf map $\eta : \pi_0^p X \to \pi_0 X$ is therefore pointwise monic as well as pointwise epi. □

Suppose that X is a locally fibrant simplicial sheaf on the site $\mathcal{C}$. The homotopy group sheaves $\pi_n X \to X_0$ are defined as sheaves fibred over the sheaf of vertices X_0 by letting $F_n X$ be the locally fibrant simplicial sheaf defined by the pullback diagram

$$\begin{tikzcd} F_n X \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^n, X) \arrow[d, "i^*"] \\ X_0 \arrow[r] & \mathbf{hom}(\partial\Delta^n, X), \end{tikzcd}$$

and then by defining $\pi_n X = \pi_0 F_n X$, where the latter denotes the *sheaf* of path components of the simplicial sheaf $F_n X$, as above. The group object multiplication $m : \pi_n X \times_{X_0} \pi_n X \to \pi_n X$ is defined by analogy with the group object multiplication for Kan complexes: define a locally fibrant simplicial sheaf $G_n X$ by requiring that the diagram

$$\begin{tikzcd} G_n X \arrow[r] \arrow[d] & \mathbf{hom}(\Delta^{n+1}, X) \arrow[d, "j^*"] \\ X_0 \arrow[r] & \mathbf{hom}(K_n, X) \end{tikzcd}$$

is a pullback, and then consider the resulting diagram

$$
\begin{tikzcd}
R \arrow[d, shift left] \arrow[d, shift right] & & \\
G_nX_0 \arrow[rr, "d_{n*}"] \arrow[d, two heads, "(d_{n-1}, d_{n+1})_*"'] & & F_nX_0 \arrow[r, two heads] & \pi_0F_nX \\
(F_nX \times_{X_0} F_nX)_0 \arrow[d, two heads] & & & \\
\pi_0F_nX \times_{X_0} \pi_0F_nX \arrow[rrruu, dashed, "m"'] & & &
\end{tikzcd}
$$

We haven't exactly shown that the morphism m exists yet, but the indicated morphisms $R \rightrightarrows G_nX_0$ are supposed to denote the kernel pair of the composite sheaf epimorphism

$$G_nX_0 \to \pi_0F_nX \times_{X_0} \pi_0F_nX.$$

A unique morphism $m : \pi_0F_nX \times_{X_0} \pi_0F_nX \to \pi_0F_nX$ exists and makes the diagram commute if it can be shown that the horizontal composite $G_nX_0 \to \pi_0F_nX$ equalizes the arrows $R \rightrightarrows G_nX_0$ in the sense that it gives the same result when composed with each of them. This is shown by applying the Boolean localization functor $\wp^*$. This functor commutes with the constructions π_0, F_n and G_n, and $\wp^*X$ is a presheaf of Kan complexes. There is an isomorphism $\pi_0F_n\wp^*X \cong \pi_0^pF_n\wp^*X$ by Lemma 25, so that the ordinary group object structure on the presheaves of homotopy group objects for the presheaf of Kan complexes $\wp^*X$ determines a map

$$m : \pi_0F_n\wp^*X \times_{\wp^*X_0} \pi_0F_n\wp^*X \to \pi_0F_n\wp^*X.$$

In other words, applying the functor $\wp^*$ to the map $G_nX_0 \to \pi_0F_nX$ gives a morphism which equalizes the induced maps $\wp^*R \rightrightarrows \wp^*G_nX_0$. The functor $\wp^*$ is faithful, so that $G_nX_0 \to \pi_0F_nX$ equalizes the morphisms $R \rightrightarrows G_nX_0$, and the multiplication map m is defined uniquely.

The inverse map $\sigma : \pi_0F_nX \to \pi_0F_nX$ of sheaves over X_0 exists by a completely analogous argument, and the identity $e : X_0 \to \pi_0F_nX$ is a canonical section. Finally, the maps m, σ and e define a group object structure on $\pi_0F_nX = \pi_nX \to X_0$: just use the fact that $\wp^*$ is faithful (and exact) again, together with the observation that the corresponding group object structure for $\pi_n\wp^*X$ already exists, since $\wp^*X$ is a presheaf of Kan complexes, and the sheaves of homotopy groups for $\wp^*X$ coincide with their underlying presheaves.

Lemma 26. *A map $f : X \to Y$ of simplicial presheaves on a site $\mathcal{C}$ is a weak equivalence if and only if*

(1) *the induced map*

$$f_* : \pi_0L^2Ex^\infty X \to \pi_0L^2Ex^\infty Y$$

is an isomorphism of sheaves, and

(2) *the diagrams*

$$\begin{tikzcd} \pi_n L^2 Ex^\infty X \arrow[r, "f_*"] \arrow[d] & \pi_n L^2 Ex^\infty Y \arrow[d] \\ (L^2 Ex^\infty X)_0 \arrow[r, "f_*"'] & (L^2 Ex^\infty Y)_0 \end{tikzcd}$$

are pullbacks for $n \geq 1$.

Proof: The map f is a local weak equivalence if and only if the induced map $f_* : L^2 Ex^\infty X \to L^2 Ex^\infty Y$ is a local weak equivalence, so it's enough to show that a map $f : X \to Y$ of locally fibrant simplicial sheaves on $\mathcal{C}$ is a weak equivalence if and only if the map

$$f_* : \pi_0 X \to \pi_0 Y$$

is a sheaf isomorphism, and all of the diagrams

$$\begin{tikzcd} \pi_n X \arrow[r, "f_*"] \arrow[d] & \pi_n Y \arrow[d] \\ X_0 \arrow[r, "f_*"'] & Y_0 \end{tikzcd}$$

are pullbacks. The Boolean localization functor $\wp^*$ reflects isomorphisms and pullbacks, so that these conditions are equivalent to the assertions that

$$\wp^* f_* : \pi_0 \wp^* X \to \pi_0 \wp^* Y$$

is a sheaf isomorphism, and all diagrams

$$\begin{tikzcd} \pi_n \wp^* X \arrow[r, "\wp^* f_*"] \arrow[d] & \pi_n \wp^* Y \arrow[d] \\ \wp^* X_0 \arrow[r, "\wp^* f_*"'] & \wp^* Y_0 \end{tikzcd}$$

are pullbacks. The simplicial sheaves $\wp^* X$ and $\wp^* Y$ are presheaves of Kan complexes, and their associated presheaves of homotopy group objects coincide with the respective sheaves of homotopy group objects, so these last conditions are jointly equivalent to the assertion that $\wp^* f : \wp^* X \to \wp^* Y$ is a pointwise weak equivalence. □

We can now give our independence result:

Theorem 27. *Suppose that $\mathcal{C}$ is an arbitrary Grothendieck site. Say that a cofibration of simplicial presheaves on $\mathcal{C}$ is a pointwise monomorphism, a local weak equivalence is a map satisfying the conditions of Lemma 26, and a global fibration is a map which has the right lifting property with respect to all maps which are simultaneously cofibrations and local weak equivalences. Then, with these definitions, the categories $\mathbf{S}Pre(\mathcal{C})$ and $\mathbf{S}Shv(\mathcal{C})$, respectively, of simplicial presheaves and simplicial sheaves on the site $\mathcal{C}$ satisfy the axioms for a proper closed simplicial model category. Furthermore, the associated sheaf functor induces an equivalence*

$$Ho(\mathbf{S}Pre(\mathcal{C})) \simeq Ho(\mathbf{S}Shv(\mathcal{C}))$$

between the associated homotopy categories. □

Suppose that U is an object of the site $\mathcal{C}$ and that x is a vertex of the simplicial set $X(U)$, where X is a locally fibrant simplicial sheaf. Then there is a pullback diagram

$$\begin{tikzcd} \pi_n(X|_U, x)(U) \arrow[r] \arrow[d] & \pi_n X(U) \arrow[d] \\ * \arrow[r, "x"'] & X_0(U) \end{tikzcd}$$

where $\pi_n(X|_U, x)$ is the n^{th} ordinary sheaf of homotopy groups for the restricted simplicial sheaf $X|_U$ on the site $\mathcal{C} \downarrow U$ of objects over U, based at the global section x. Also if $f : X \to Y$ is a map of locally fibrant simplicial sheaves, then the diagram

$$\begin{tikzcd} \pi_n X \arrow[r, "f_*"] \arrow[d] & \pi_n Y \arrow[d] \\ X_0 \arrow[r, "f_*"'] & Y_0 \end{tikzcd}$$

is a pullback if and only if all of the induced maps

$$\pi_n(X|_U, x) \xrightarrow{f_*} \pi_n(Y|_U, f(x))$$

are isomorphisms for $U \in \mathcal{C}$, $x \in X_0(U)$, so the definition of local weak equivalence given here coincides with the standard form.

Remark 28. There is another, much easier, way to see the independence result for the closed model structure on $\mathbf{S}Pre(\mathcal{C})$. The key point is a combination of Lemma 7, Lemma 9, and Lemma 11: if $f : X \to Y$ is a map of $\mathbf{S}Pre(\mathcal{C})$, then there is a commutative diagram of the form

$$\begin{tikzcd} X \arrow[r, "\nu"] \arrow[d, "f"'] & Ex^\infty X \arrow[d, "Ex^\infty f"'] \arrow[dr, "i"] & \\ Y \arrow[r, "\nu"'] & Ex^\infty Y & Z \arrow[l, "q"] \end{tikzcd}$$

where i is a pointwise weak equivalence and a cofibration, and q is a pointwise Kan fibration. The maps ν are pointwise weak equivalences, so Lemma 9 says that f is a local weak equivalence if and only if the pointwise Kan fibration q is a local weak equivalence. The objects Z and $Ex^\infty Y$ are presheaves of Kan complexes, so that one infers from Lemma 7 and Lemma 11 that q is a local weak equivalence if and only if it has the local right lifting property with respect to all inclusions $\partial\Delta^n \subset \Delta^n$. This local right lifting property is an internal criterion for simplicial presheaves on the site $\mathcal{C}$, and is independent of any Boolean localization.

Bibliography

[1] A.K. Bousfield and E.M. Friedlander, *Homotopy theory of* Γ*-spaces, spectra, and bisimplicial sets*, Springer Lecture Notes in Math. **658** (1978), 80–150.

[2] P.G. Goerss and J.F. Jardine, *Simplicial Homotopy Theory*, Preprint (1996), http://www.math.uwo.ca/∼jardine/papers/simp-sets/.

[3] L. Illusie, *Complexe cotangent et déformations I*, Springer Lecture Notes in Math., Vol. 239, Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[4] J.F. Jardine, *Simplicial presheaves*, J. Pure Applied Algebra **47** (1987), 35–87.

[5] J.F. Jardine, *Stable homotopy theory of simplicial presheaves*, Can. J. Math. **39(3)** (1987), 733–747.

[6] J.F. Jardine, *Generalized Etale Cohomology Theories*, to appear in Birkhäuser Progress in Math. series, http://www.math.uwo.ca/∼jardine/papers/gen-et/.

[7] A. Joyal, *letter to A. Grothendieck*, (1984).

[8] D.M. Kan, *On c.s.s. complexes*, Amer. J. Math. **79** (1957), 449–476.

[9] S. Mac Lane and I. Moerdijk, *Sheaves in Geometry and Logic: A First Introduction to Topos Theory*, Springer-Verlag, Berlin Heidelberg New York, 1992.

[10] J. Peter May, *Simplicial Objects in Algebraic Topology*, Van Nostrand, Princeton, 1967.

[11] D. Quillen, *Homotopical Algebra*, Lecture Notes in Math., Vol. 43, Springer, Berlin-Heidelberg- New York, 1967.

[12] D. Quillen, *Rational homotopy theory*, Ann. of Math. **90** (1969), 205–295.

[13] H. Schubert, *Categories*, Springer-Verlag, New York Heidelberg Berlin, 1972.

[14] D.H. Van Osdol, *Simplicial homotopy in an exact category*, Amer. J. Math. **99** (1977), 1193–1204.

Mathematics Department
University of Western Ontario
London, Ont., N6A 5B7
CANADA
jardine@jardine.math.uwo.ca

Multirelative K-Theory and Axioms for the K-Theory of Rings

Frans Keune

Received: July 17, 1996

Communicated by Ulf Rehmann

Abstract. K-groups are defined for a special type of m-tuples of ideals in a ring. It is shown that some of the properties of this multirelative K-theory characterize the K-theory of rings.

1991 Mathematics Subject Classification: Primary 19D99.

Introduction

Multirelative K-groups $K_n(R, \mathfrak{a}_1, \dots, \mathfrak{a}_m)$ of an m-tuple $(\mathfrak{a}_1, \dots, \mathfrak{a}_m)$ of ideals of a ring R are recently used to derive properties of the absolute K-groups, e.g. by Levine [4] and by Bloch and Lichtenbaum [1]. Here it is shown how K-theory as defined in [3] can easily be extended to the multirelative case and that some of its properties can be taken as axioms for the K-theory of rings. Special types of m-tuples of ideals—the 'normal' m-tuples—play a crucial role. In fact we will only define multirelative K-groups for such m-tuples. The notion of normal m-tuple of ideals is introduced in Section 2. It already appeared in 1981 in a paper by Dayton and Weibel [2] on the K-theory of affine glued schemes under the name of 'condition (CRT)' (= Chinese Remainder Theorem).

In Section 4 we review briefly higher K-theory as defined in [3]. In Section 6 multirelative K-groups are defined, and in Section 7 it is shown that from some of their properties one can reconstruct the K-theory of rings.

1 Notations

In this paper 'ring' stands for a non-unital ring. Non-unital rings form a category which is denoted by $\mathcal{R}$.

Since the functors GL, E and K_1 are product preserving functors from unital rings to groups, they can be extended to functors defined on $\mathcal{R}$ in the usual way: if T is one of these functors, then put

$$T(R) := \operatorname{Ker}(T(R^+) \to T(\mathbb{Z})),$$

where $R^+ = R \times \mathbb{Z}$ with multiplication given by

$$(r,k)(s,l) = (rs + ks + lr, kl)$$

is a ring with $(0,1)$ as unity element.

Here 'ideal' will always stand for 'twosided ideal'.

By $\mathcal{A}$ we will denote the category of Abelian groups, by $\mathcal{G}$ the category of all groups, and by $\mathcal{S}$ the category of sets. The category of simplicial objects in a category $\mathcal{C}$ is denoted by $s\mathcal{C}$.

2 m-CUBES AND NORMAL m-TUPLES

In this section the notion of normality of an m-tuple of ideals is considered. Only the group structure is involved in its definition, and since we can use later a similar notion for groups instead of rings we give a more general definition. By $\underline{m}$ we will denote the set $\{1,\ldots,m\}$.

DEFINITION 1. An m-tuple $(B_1,\ldots,B_m)$ of normal subgroups of a group A—also denoted as $(A,B_1,\ldots,B_m)$—is called *normal* if for all subsets I and J of $\underline{m}$

$$\bigcap_{i\in I} B_i \cdot \prod_{j\in J} B_j = \bigcap_{i\in I}\Big(B_i \cdot \prod_{j\in J} B_j\Big).$$

The condition is trivially fulfilled when $I \cap J \neq \emptyset$. In the case of Abelian groups it reads in the additive notation as

$$\bigcap_{i\in I} B_i + \sum_{j\in J} B_j = \bigcap_{i\in I}\Big(B_i + \sum_{j\in J} B_j\Big).$$

Note that in the special case of an m-tuple of ideals in a commutative ring the condition is a local one since it involves only intersections and sums of ideals.

The subsets of $\underline{m}$ are ordered by inclusion. This ordered set determines in the usual way a category $\mathcal{C}_m$. For every pair (I,J) of subsets with $I \subseteq J$ there is the unique morphism ρ^I_J from I to J in $\mathcal{C}_m$.

DEFINITION 2. Let $\mathcal{D}$ be a category. An *m-cube* in $\mathcal{D}$ is a functor

$$D\colon \mathcal{C}_m \to \mathcal{D}, \quad I \mapsto D_I, \quad \rho^I_J \mapsto r^I_J.$$

The morphisms in $\mathcal{C}_m$ are generated by the ρ^I_J, where $\#J = \#I + 1$. An m-cube in a category $\mathcal{D}$ is a commutative diagram in $\mathcal{D}$ having the shape of an m-dimensional cube. The edges of the cube correspond to the images of these generating morphisms.

DEFINITION 3. Let $D\colon \mathcal{C}_m \to \mathcal{D}$ be an m-cube in $\mathcal{D}$. It is said to be a *split* m-cube if for every pair of subsets (I,J) of $\underline{m}$ satisfying $I \subseteq J$ there is a morphism $s^J_I\colon D_J \to D_I$ in $\mathcal{D}$ such that

(S1) $s^J_I s^K_J = s^K_I$ for all $I \subseteq J \subseteq K$,

(S2) $r^I_J s^J_I = 1_{D_J}$ for all $I \subseteq J$, (S3) $r_J^{I\cap J} s_{I\cap J}^I = s_J^{I\cup J} r_{I\cup J}^I$ for all I and J.

(Of course such a split m-cube can also be seen as a functor defined on a category which is obtained from $\mathcal{C}_m$ by adjoining extra morphisms $\sigma_I^J\colon J \to I$.)

In condition (S3) one only needs the case where $\#(I\setminus J) = \#(J\setminus I) = 1$. It then reads

(S3') $r_{I\cup\{k\}}^I s_I^{I\cup\{j\}} = s_{I\cup\{k\}}^{I\cup\{j,k\}} r_{I\cup\{j,k\}}^{I\cup\{j\}}$ for all $j,k \notin I$ with $j \neq k$.

This can easily be seen as follows. Put $K = I\cap J$, $I\setminus K = \{i_1,\dots,i_p\}$ and $J\setminus K = \{j_1,\dots,j_q\}$. Then the result follows from the diagram

$$\begin{tikzcd}
D_K \arrow[r] & D_{K\cup\{i_1\}} \arrow[r] & \cdots \arrow[r] & D_I \\
D_{K\cup\{j_1\}} \arrow[u] \arrow[r] & D_{K\cup\{i_1,j_1\}} \arrow[u] \arrow[r] & \cdots \arrow[r] & D_{I\cup\{j_1\}} \arrow[u] \\
\vdots \arrow[u] & \vdots \arrow[u] & & \vdots \arrow[u] \\
D_J \arrow[u] \arrow[r] & D_{J\cup\{i_1\}} \arrow[u] \arrow[r] & \cdots \arrow[r] & D_{I\cup J} \arrow[u]
\end{tikzcd}$$

where the horizontal maps are r-maps and the vertical maps are s-maps.

Definition 4. An m-tuple $T = (A, B_1, \dots, B_m)$ of normal subgroups determines an m-cube in $\mathcal{G}$:

$$I \mapsto T_I = A \Big/ \prod_{i\in I} B_i.$$

When $I \subseteq J$, then $\prod_{i\in I} B_i \subseteq J$ and 1_A induces a grouphomomorphism $r_J^I\colon T_I \to T_J$. This m-cube is said to be *induced* by the m-tuple T. Similarly for an m-tuple of ideals in a ring.

Proposition 2.1. *Let $D\colon \mathcal{C}_m \to \mathcal{D}$ be an m-cube in $\mathcal{G}$, which is split as an m-cube in $\mathcal{S}$. Then D is induced by a normal m-tuple of normal subgroups of $D_\emptyset$.*

Proof. For $i \in \underline{m}$ put

$$B_i = \mathrm{Ker}\Big(r_{\{i\}}^\emptyset\colon D_\emptyset \to D_{\{i\}}\Big).$$

We will first show that the cube is induced by the m-tuple $(D_\emptyset, B_1, \dots, B_m)$. Since the cube splits in $\mathcal{S}$, the homomorphisms $D_\emptyset \to D_I$ are surjective. To show that for each $I \subseteq \underline{m}$

$$\mathrm{Ker}(D_\emptyset \to D_I) = \prod_{i\in I} B_i.$$

This can be done by induction on $\#(I)$. For $\#(I) = 0$ it is trivial. Let $\#(I) > 0$. Choose $k \in I$. By induction hypothesis

$$\mathrm{Ker}\Big(D_\emptyset \to D_{I\setminus\{k\}}\Big) = \prod_{i\in I\setminus\{k\}} B_i.$$

Since the cube splits in $\mathcal{S}$ we have a commutative diagram with exact rows and columns:

$$\begin{array}{ccccccccc}
 & & 1 & & 1 & & 1 & & \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & B_k \cap \prod_{i\in I\setminus\{k\}} B_i & \longrightarrow & B_k & \longrightarrow & \mathrm{Ker}\left(r_I^{I\setminus\{k\}}\right) & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & \prod_{i\in I\setminus\{k\}} B_i & \longrightarrow & D_\emptyset & \longrightarrow & D_{I\setminus\{k\}} & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
1 & \longrightarrow & \mathrm{Ker}\left(r_I^{\{k\}}\right) & \longrightarrow & D_{\{k\}} & \longrightarrow & D_I & \longrightarrow & 1 \\
 & & \downarrow & & \downarrow & & \downarrow & & \\
 & & 1 & & 1 & & 1 & &
\end{array}$$

Hence

$$\mathrm{Ker}(r_I^\emptyset)/B_k \cong \prod_{i\in I\setminus\{k\}} B_i \Big/ \left(B_k \cap \prod_{i\in I\setminus\{k\}} B_i\right) \cong \prod_{i\in I}(B_i/B_k),$$

and therefore,

$$\mathrm{Ker}(r_I^\emptyset) = \prod_{i\in I} B_i.$$

For the normality of the m-tuple let $I, J \subseteq \underline{m}$ and consider the commutative square

$$\begin{array}{ccc}
D_\emptyset & \xrightarrow{(r_{\{i\}}^\emptyset)} & \bigtimes_{i\in I} D_\emptyset/B_i \\
\Big\downarrow r_J^\emptyset & & \Big\downarrow (r_{J\cup\{i\}}^{\{i\}}) \\
D_\emptyset \Big/ \prod_{j\in J} B_j & \xrightarrow{(r_{J\cup\{i\}}^J)} & \bigtimes_{i\in I} D_\emptyset \Big/ \prod_{j\in J\cup\{i\}} B_j.
\end{array}$$

Since the m-cube is split in $\mathcal{S}$ the vertical homomorphisms have compatible sections in $\mathcal{S}$. So $r_J^\emptyset$ induces a surjective homomorphism on the kernels of the horizontal homomorphisms. This holds for all $I, J \subseteq \underline{m}$. Therefore, the m-tuple $(D_\emptyset, B_1, \ldots, B_m)$ is normal. □

For the Abelian case we also prove the converse.

Proposition 2.2. *Let $T = (A, B_1, \ldots, B_m)$ be a normal m-tuple of subgroups of an Abelian group A. Then the induced m-cube is split in the category $\mathcal{S}$.*

Proof. By taking kernels of the surjective homomorphisms in the induced m-cube it can be extended to a diagram of 3^m Abelian groups. We will give a detailed description of this diagram and show how a splitting of the cube can be obtained from it.

For each pair (I, J) of disjoint subsets of $\underline{m}$ define

$$C_J^I = \bigcap_{i\in I} B_i + \sum_{j\in J} B_j \Big/ \sum_{j\in J} B_j.$$

Then for each such pair (I, J) and each $k \notin I \cup J$ we have a surjective homomorphism $C_J^I \to C_{J\cup\{k\}}^I$, induced by $r_{J\cup\{k\}}^J \colon A_J \to A_{J\cup\{k\}}$, where we use the notation

$$A_J = A \Big/ \sum_{j\in J} B_j.$$

Thus $A_J = C_J^{\emptyset}$. The kernel of the surjective homomorphism $C_J^I \to C_{J\cup\{k\}}^I$ is

$$\Big(\bigcap_{i\in I} B_i + \sum_{j\in J} B_j\Big) \cap \Big(B_k + \sum_{j\in J} B_j\Big) \Big/ B_k + \sum_{j\in J} B_j.$$

We have the inclusions

$$\bigcap_{i\in I\cup\{k\}} B_i + \sum_{j\in J} B_j \subseteq \Big(\bigcap_{i\in I} B_i + \sum_{j\in J} B_j\Big) \cap \Big(B_k + \sum_{j\in J} B_j\Big) \subseteq \bigcap_{i\in I\cup\{k\}} \Big(B_i + \sum_{j\in J} B_j\Big).$$

By normality these groups are equal, so we have a short exact sequence

$$0 \to C_J^{I\cup\{k\}} \to C_J^I \to C_{J\cup\{k\}}^I \to 0.$$

For each pair (I, J) of disjoint subsets of $\underline{m}$ satisfying $I \cup J = \underline{m}$ choose a section

$$t_J^I \colon C_J^I \to C_\emptyset^I (\subseteq C_\emptyset^\emptyset = A)$$

of the map $C_\emptyset^I \to C_J^I$ induced by $r_J^\emptyset \colon A \to A_J$ and satisfying $t_J^I(0) = 0$. Next define maps $t_J^I \colon C_J^I \to C_\emptyset^I$ for every disjoint pair (I, J) using induction to the number of elements of the complement of $I \cup J$. So, let (I, J) be a disjoint pair of subsets of $\underline{m}$ with $\#(I \cup J) = n < m$ and assume that sections $t_L^K \colon C_L^K \to C_\emptyset^K$ have already been defined for pairs (K, L) with $K \cup L$ having more than n elements.

Choose $k \in \underline{m} \setminus (I \cup J)$. Let $x \in C_J^I$, then for $y = r_J^\emptyset t_{J\cup\{k\}}^I r_{J\cup\{k\}}^J(x)$ we have

$$r_{J\cup\{k\}}^J(y) = r_{J\cup\{k\}}^\emptyset t_{J\cup\{k\}}^I r_{J\cup\{k\}}^J(x) = r_{J\cup\{k\}}^J(x),$$

so, $x - y \in C_J^{I\cup\{k\}}$. Now define t_J^I by

$$t_J^I(x) = t_J^{I\cup\{k\}}(x - y) + t_{J\cup\{k\}}^I r_{J\cup\{k\}}^J(x).$$

It easily verified that this map is a section of $r \colon C_\emptyset^I \to C_J^I$. Furthermore it is independent of the choice of k: if also $l \notin I \cup J$, then in both cases the image of an $x \in C_J^I$ under t_J^I is determined in the same way by the images of the same elements in the following four groups

$$C_J^{I\cup\{l,k\}},\ C_{J\cup\{k\}}^{I\cup\{l\}},\ C_{J\cup\{l\}}^{I\cup\{k\}},\ \text{and}\ C_{J\cup\{k,l\}}^I :$$

$$\begin{array}{ccccccccc}
& & 0 & & 0 & & 0 & & \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C_J^{I\cup\{l,k\}} & \longrightarrow & C_J^{I\cup\{l\}} & \longrightarrow & C_{J\cup\{k\}}^{I\cup\{l\}} & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C_J^{I\cup\{k\}} & \longrightarrow & C_J^{I} & \longrightarrow & C_{J\cup\{k\}}^{I} & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & C_{J\cup\{l\}}^{I\cup\{k\}} & \longrightarrow & C_{J\cup\{l\}}^{I} & \longrightarrow & C_{J\cup\{k,l\}}^{I} & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
& & 0 & & 0 & & 0 & &
\end{array}$$

Thus we obtain a splitting of the cube, where the sections s_I^J of the homomorphisms r_J^I, where $I \subseteq J$, are the maps $r_I^{\emptyset} t_J^{\emptyset}$. In particular, condition (S3') follows from the above diagram for $I = \emptyset$. □

3 Operations on normal m-tuples of ideals

By $\mathcal{R}_m$ we will denote the category of all normal m-tuples of ideals. Such an m-tuple is denoted as $(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m)$, where R is a ring and $\mathfrak{a}_1, \ldots, \mathfrak{a}_m$ are ideals of R. A morphism $\phi\colon (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m) \to (S, \mathfrak{b}_1, \ldots, \mathfrak{b}_m)$ is just a ringhomomorphism $\phi\colon R \to S$ satisfying $\phi(\mathfrak{a}_i) \subseteq \mathfrak{b}_i$ for all $i \in \underline{m}$.

The following notations will simplify notations for long exact sequences of multirelative K-theory. Another advantage will be that they are useful to indicate funtoriality properties.

For each $m \geq 1$ the functor $D\colon \mathcal{R}_m \to \mathcal{R}_{m-1}$ is the functor that deletes the last ideal:

$$D(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m) = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_{m-1})$$

and which has no effect on morphisms.

For each $m \geq 1$ the functor $M\colon \mathcal{R}_m \to \mathcal{R}_{m-1}$ is the functor that deletes the last ideal and that takes the ring and the other ideals modulo this ideal:

$$M(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m) = (R/\mathfrak{a}_m, \overline{\mathfrak{a}}_1, \ldots, \overline{\mathfrak{a}}_{m-1}),$$

where $\overline{\mathfrak{a}}_j = \mathfrak{a}_j + \mathfrak{a}_i/\mathfrak{a}_i$, and which maps a morphism to the induced morphism.

A functor morphism $\phi\colon D \to M$ of the functors $D, M\colon \mathcal{R}_m \to \mathcal{R}_{m-1}$ is defined as follows: let $A = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m)$, then $\phi_A\colon D(A) \to M(A)$ is the canonical ringhomomorphism $R \to R/\mathfrak{a}_m$.

Every $A \in \mathcal{R}_m$ has an *underlying ideal* $I(A)$, which is defined as the intersection of the m ideals in A: when $A = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m)$, then

$$I(A) = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_m.$$

Thus defined, $I(A)$ is functorial in A.

4 Higher K-theory of rings

In [3] the definition of higher K-groups is as follows. Let $R \in \mathcal{R}$. Choose a simplicial ring $\mathbf{R}$ with an augmentation $\varepsilon\colon \mathbf{R} \to R$ such that

- $\mathbf{R}$ is aspherical, i.e. $\pi_n(\mathbf{R}) = 0$ for all $n \geq 1$,
- $\mathbf{R}_m$ is free for all $m \geq 0$, say $\mathbf{R}_m$ is free on a set X_m of generators,
- the sets X_m of free generators are stable under degeneracies: $s_j(X_m) \subseteq X_{m+1}$ for all $m \geq 0$,
- the augmentation ε induces an isomorphism $\pi_0(\mathbf{R}) \xrightarrow{\sim} R$.

Then for $n \geq 3$ the group $K_n(R)$ is defined as the $(n-2)$nd homotopy group of the simplicial group $GL(\mathbf{R})$, and the groups $K_1(R)$ and $K_2(R)$ are given by the exactness of

$$0 \to K_2(R) \to \pi_0(GL\mathbf{R}) \to GL(R) \to K_1(R) \to 0.$$

The groups $K_n(R)$ for $n \geq 3$ are Abelian because $GL(\mathbf{R})$ is a simplicial group. The group $K_1(R)$ is Abelian since it is the cokernel of $GL(\mathbf{R}_0) \to GL(R)$, and $K_2(R)$ is Abelian because it is the cokernel of $GL(\mathbf{R}_1) \to GL(Z_0)$, where $Z_0 = \{\, (x_0, x_1) \mid \epsilon(x_0) = \epsilon(x_1) \,\}$. In [3] it is shown using a comparison theorem that the higher K-groups are thus well-defined and that they are actually functors. For the purpose of this paper we will confine to a functorial resolution $\mathbf{Fr}(R)$ of a ring R, which we now describe. Let $F\colon \mathcal{S} \to \mathcal{R}$ the free ring functor and let $U\colon \mathcal{R} \to \mathcal{S}$ be the underlying set functor, then the functor $FU\colon \mathcal{R} \to \mathcal{R}$ together with the obvious functor morphisms $\nu\colon FU \to (FU)^2$ and $\eta\colon FU \to I$ is a cotriple. Put

$$\mathbf{Fr}_n = (FU)^{n+1}.$$

Face and degeneracy morphisms are given by

$$d_i = (FU)^i \eta (FU)^{n-1-i} \quad \text{and} \quad s_j = (FU)^i \nu (FU)^{n-1-i}.$$

The augmentation is then given by η.

A property of this functorial resolution is that, when applied to a surjective ringhomomorphism $R \to S$, it gives a dimensionwise surjective homomorphism $\mathbf{Fr}R \to \mathbf{Fr}S$ of simplicial rings, and since the ringhomomorphisms are dimensionwise split it also gives a surjective simplicial grouphomomorphism $GL(\mathbf{Fr}R) \to GL(\mathbf{Fr}S)$. This is often convenient when considering homotopy fibres, because surjective simplicial grouphomomorphisms are fibrations themselves. So instead of taking a homotopy fibre one just takes a fibre, i.e. the kernel of the simplicial group homomorphism.

5 Cubes in a simplicial group

Let $\mathbf{A}$ be a simplicial group with augmentation $d_0\colon \mathbf{A}_0 \to A$. It is a contravariant functor $\mathbf{A}\colon \Omega_+^{\mathrm{op}} \to \mathcal{G}$ from the category Ω_+ of finite ordered sets

$$[n] = \{0, \ldots, n\} \qquad (n \geq -1)$$

(where $[-1] = \emptyset$) and monotone (= order preserving) maps to the category of groups. (Here we use the notation $\mathbf{A}_{-1} = A$.) We will show that $\mathbf{A}$ determines an m-cube of groups for every nonnegative integer m. In stead of the ordered set of subsets of $\underline{m}$ for the description of an m-cube the ordered set of subsets of $[m-1]$ will be used for this purpose.

Let $\Omega(m)$ be the category of injective monotone maps

$$\alpha\colon [k] \to [m-1].$$

A morphism from $\alpha\colon [k] \to [m-1]$ to $\beta\colon [l] \to [m-1]$ is a monotone map $\gamma\colon [k] \to [l]$ such that $\beta\gamma = \alpha$. It exists if and only if $\mathrm{Im}(\alpha) \subseteq \mathrm{Im}(\beta)$, and it is unique if it exists.

For each $I \subseteq [m-1]$ there is a unique injective monotone map

$$\alpha_I\colon [k] \to [m-1],$$

where $k = m-1-\#(I)$ and $\mathrm{Im}(\alpha_I) = [m-1] \setminus I$. If $I \subseteq J \subseteq [m-1]$, then $\mathrm{Im}(\alpha_I) \supseteq \mathrm{Im}(\alpha_J)$, so then there is a unique

$$\gamma_I^J : \alpha_J \to \alpha_I,$$

i.e. a monotone $\gamma_I^J : [m-1-\#(J)] \to [m-1-\#(I)]$ such that $\alpha_I \gamma_I^J = \alpha_J$.

DEFINITION 5. Let $\mathbf{A}$ be an augmented simplicial group and let m be a nonnegative integer. Then the *m-cube of* $\mathbf{A}$ is the m-cube $\mathbf{A}(m)\colon \mathcal{C}_m \to \mathcal{G}$ with

$$\begin{cases} \mathbf{A}(m)_I = \mathbf{A}_{[m-1-\#(I)]} & \text{for all } I \subseteq [m-1], \\ r_J^I = \mathbf{A}(\gamma_I^J)\colon \mathbf{A}(m)_I \to \mathbf{A}(m)_J & \text{for all } I \subseteq J \subseteq [m-1]. \end{cases}$$

LEMMA 5.1. *Let the augmentation $d_0\colon \mathbf{A}_0 \to \mathbf{A}_{-1}$ induce a surjective homomorphism $\pi_0(\mathbf{A}) \to \mathbf{A}_{-1}$. Then for all integers i, j, m such that $0 \le j < i \le m$*

$$d_i^{(m)}\Big(\mathrm{Ker}\Big(d_j^{(m)}\Big)\Big) = \mathrm{Ker}\Big(d_j^{(m-1)}\Big).$$

Proof. Let $x \in \mathrm{Ker}\Big(d_j^{(m)}\Big)$. Then, since $i > j$, $d_j d_i(x) = d_{i-1} d_j(x) = 1$. So $d_i(\mathrm{Ker}(d_j)) \subseteq \mathrm{Ker}(d_j)$. Now, let $y \in \mathrm{Ker}\Big(d_j^{(m-1)}\Big)$. There is an $x \in \mathbf{A}_m$ such that $d_j(x) = 1$ and $d_i(x) = y$. For $m > 1$ this is the case because a simplicial group is a Kan-complex, while for $m = 1$ it follows from the condition on the augmentation. ☐

PROPOSITION 5.1. *Let $\mathbf{A}$ be a simplicial group with an augmentation $d_0\colon \mathbf{A} \to A$ that induces an isomorphism $\pi_0(\mathbf{A}) \to A$. Then for all $m \ge 1$ the m-cube $\mathbf{A}(m)$ is induced by the m-tuple*

$$(\mathbf{A}_{m-1}, \mathrm{Ker}(d_0), \ldots, \mathrm{Ker}(d_{m-1})).$$

Proof. All face maps are surjective, so it remains to show that for all $J \subseteq [m-1]$

$$\mathrm{Ker}(r_J^{\emptyset}) = \prod_{j \in J} \mathrm{Ker}\Big(d_j^{(m-1)}\Big).$$

For $J = \emptyset$ this is trivially true. Let J be nonempty and proceed by induction. Let $x \in \operatorname{Ker}(r_J^\emptyset)$. Let $k \in J$ be maximal. Then $r_{\{k\}}^\emptyset(x) = d_k(x) \in \operatorname{Ker}\Big(r_J^{\{k\}}\Big)$. By induction this group is equal to $\prod_{j\in J'} \operatorname{Ker}\Big(d_j^{(m-2)}\Big)$, where $J' = J \setminus \{k\}$. (Here we used the maximality of k in J and the same result for the $(m-1)$-cube $\mathbf{A}(m-1)$.) By the lemma we have

$$d_k\Bigg(\prod_{j\in J'} \operatorname{Ker}\Big(d_j^{(m-1)}\Big)\Bigg) = \prod_{j\in J'} \operatorname{Ker}\Big(d_j^{(m-2)}\Big).$$

Choose $y \in \prod_{j\in J'} \operatorname{Ker}(d_j^{(m-1)}$ such that $d_k(y) = d_k(x)$. Then $xy^{-1} \in \operatorname{Ker}(d_k)$. It follows that

$$\operatorname{Ker}(r_J^\emptyset) \subseteq \prod_{j\in J} \operatorname{Ker}\Big(d_j^{(m-1)}\Big).$$

For the other inclusion note that $d_j = r_{\{j\}}^\emptyset$ and

$$r_J^{\{j\}} r_{\{j\}}^\emptyset = r_J^\emptyset.$$

□

PROPOSITION 5.2. *Let* $\mathbf{A}$ *be as in Proposition 5.1 and assume moreover that* $\mathbf{A}$ *is aspherical. Then the* m*-tuple*

$$(\mathbf{A}_{m-1}, \operatorname{Ker}(d_0), \ldots, \operatorname{Ker}(d_{m-1}))$$

is normal.

Proof. The edges of the m-cube are face maps of the simplicial group (A). Normality means that these maps preserve intersections of (the images of) the normal subgroups $\operatorname{Ker}(d_0), \ldots, \operatorname{Ker}(d_{m-1})$. By induction it suffices to show this for the face maps $d_i^{(m-1)}$. Let $J \subseteq [m-1]$. Then to show that

$$d_i\Bigg(\bigcap_{j\in J} \operatorname{Ker}(d_j)\Bigg) = \bigcap_{j\in J} d_i(\operatorname{Ker}(d_j)).$$

for $i \notin J$. The inclusion of the left hand side in the right hand side is trivial. So let $x \in \bigcap_{j\in J} d_i(\operatorname{Ker}(d_j))$. Then for $j \in J$ there is an $y_j \in \operatorname{Ker}(d_j)$ such that $x = d_i(y_j)$. For $j < i$ it follows that $d_j(x) = d_j d_i(y_j) = d_{i-1} d_j(x_j) = 1$. Similarly for $j > i$ we have $d_{j-1}(x) = 1$. So, since a simplicial group is a Kan-complex and for $J = [m-1]$ since $\mathbf{A}$ is aspherical, there is a $y \in \mathbf{A}_{m-1}$ such that $d_j(y) = 1$ for all $j \in J$ and $d_i(y) = x$. This shows that $x \in d_i\Big(\bigcap_{j\in J} \operatorname{Ker}(d_j)\Big)$. □

6 MULTIRELATIVE K-THEORY

A normal m-tuple of ideals $A = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m)$ induces an m-cube in $\mathcal{R}$

$$A \colon I \mapsto R \Big/ \sum_{i\in I} \mathfrak{a}_i,$$

which by Proposition 2.2 is split in $\mathcal{S}$. Application of **Fr** to this m-cube gives an m-cube of simplicial rings which is dimensionwise split in $\mathcal{R}$. Put

$$\mathbf{Fr}(R,\mathfrak{a}_i) := \mathrm{Ker}(\mathbf{Fr}(R) \to \mathbf{Fr}(R/\mathfrak{a}_i)).$$

This is a simplicial ideal. The m-cube is then induced by the m-tuple

$$(\mathbf{Fr}(R), \mathbf{Fr}(R,\mathfrak{a}_1), \dots, \mathbf{Fr}(R,\mathfrak{a}_m)),$$

of simplicial ideals, an object of the category $s\mathcal{R}_m$ of normal m-tuples of simplicial ideals. We also define the simplicial ideal

$$\mathbf{Fr}(R,\mathfrak{a}_1,\dots,\mathfrak{a}_m) := \bigcap_{i=1}^{m} \mathbf{Fr}(R,\mathfrak{a}_i).$$

Application of GL gives an m-cube of simplicial groups, which is dimensionwise split in $\mathcal{G}$. This m-cube is induced by the m-tuple

$$(GL\mathbf{Fr}(R), GL\mathbf{Fr}(R,\mathfrak{a}_1), \dots, GL\mathbf{Fr}(R,\mathfrak{a}_m))$$

of simplicial normal subgroups. For $n \geq 3$ we define multirelative K_n by

$$K_n(R,\mathfrak{a}_1,\dots\mathfrak{a}_m) := \pi_{n-2}(GL\mathbf{Fr}(R,\mathfrak{a}_1,\dots,\mathfrak{a}_m)).$$

Multirelative K_2 and K_1 are then given by the exactness of

$$\begin{aligned}0 \to K_2(R,\mathfrak{a}_1,\dots\mathfrak{a}_m) \to \pi_0(GL\mathbf{Fr}(R,\mathfrak{a}_1,\dots,\mathfrak{a}_m)) \to \\ GL(\mathfrak{a}_1\cap\dots\cap\mathfrak{a}_m) \to K_1(R,\mathfrak{a}_1,\dots\mathfrak{a}_m) \to 0.\end{aligned}$$

These multirelative K_1 and K_2 are Abelian groups for the same reason as in the absolute case.

Now let $A \in \mathcal{R}_m$ with $m \geq 1$. Then $\phi_*\colon GL\mathbf{Fr}(DA) \to GL\mathbf{Fr}(MA)$ is a fibration with fibre $GL\mathbf{Fr}(A)$. The long exact sequence of homotopy groups is a long exact sequence of multirelative K-groups which can easily be extended to include multirelative K_2 and K_1.

Proposition 6.1. *Let $A \in \mathcal{R}_m$ with $m \geq 1$. Then we have a functorial exact sequence*

$$\cdots \to K_n(A) \to K_n(DA) \to K_n(MA) \to K_{n-1}(A) \to \cdots \to K_1(MA).$$

□

The connecting map $K_n(MA) \to K_{n-1}(A)$ will be denoted by δ and the map $K_n(A) \to K_n(DA)$ by ι. To put it in an even more functorial way, we have an exact sequence of functors and functor morphisms

$$\cdots \to K_n \xrightarrow{\iota} K_nD \xrightarrow{K_n(\phi)} K_nM \xrightarrow{\delta} K_{n-1} \to \cdots \to K_1M.$$

In the remaining part of this section multirelative K_0 is defined and the long exact sequence for multirelative K-theory is extended with multirelative K_0-groups.

DEFINITION 6. For a normal m-tuple A of ideals we define

$$K_0(A) = K_0(IA).$$

Thus defined, K_0 is a functor from $\mathcal{R}_m$ to $\mathcal{A}$.

For $m = 1$ we take the long exact sequence to be the long exact sequence of an ideal in a ring. Now assume that $m \geq 1$ and that we have an extended long exact sequence

$$\cdots \to K_1 D \to K_1 M \to K_0 \to K_0 D \to K_0 M$$

of functors $\mathcal{R}_m \to \mathcal{A}$. We will show that there is also such a sequence of functors $\mathcal{R}_{m+1} \to \mathcal{A}$.

Let $A = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_{m+1}) \in \mathcal{R}_{m+1}$. Put $\mathfrak{b} = IA = \bigcap_{i=1}^{m+1} \mathfrak{a}_i$. We have exact sequences for the following m-tuples of ideals

$$\begin{aligned} B &= DA = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m), \\ \overline{B} &= (R/\mathfrak{b}, \mathfrak{a}_1/\mathfrak{b}, \ldots, \mathfrak{a}_m/\mathfrak{b}) \end{aligned}$$

and

$$(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_{m-1}, \mathfrak{b}).$$

These m-tuples are normal and their K-groups fit into a commutative diagram

$$\begin{tikzcd}
 & & K_1(DB) \arrow[dr] \arrow[rr, bend left] & & K_1(MB) \arrow[dr] & & \\
 & K_1(B) \arrow[ur] \arrow[dr] & & K_1(D\overline{B}) \arrow[ur] \arrow[dr] & & K_0(B) \arrow[dr] & \\
K_2(MB) \arrow[ur] \arrow[rr, bend right] & & K_1(\overline{B}) \arrow[ur] \arrow[rr, dashed, bend right] & & K_0(\mathfrak{b}) \arrow[ur] \arrow[rr, bend right] & & K_0(DA).
\end{tikzcd}$$

Let the dashed arrow be the composition $K_1(\overline{B}) \to K_1(D\overline{B}) \to K_0(\mathfrak{b})$. By an easy diagram chase we see that the sequence with the dashed arrow is exact as well. The identity on R is a morphism

$$(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m, \mathfrak{b}) \to A$$

in $\mathcal{R}_{m+1}$. So we have a commutative diagram with exact rows:

$$\begin{array}{ccccccc}
K_1(R,\mathfrak{a}_1,\dots,\mathfrak{a}_m,\mathfrak{b}) & \longrightarrow & K_1(B) & \longrightarrow & K_1(\overline{B}) & \longrightarrow & K_0(\mathfrak{b}) \\
\downarrow & & \downarrow{\scriptstyle 1} & & \downarrow{\scriptstyle \alpha} & & \\
K_1(A) & \longrightarrow & K_1(DA) & \longrightarrow & K_1(MA). & &
\end{array}$$

It now suffices to show that the morphism α in this diagram is an isomorphism. The $(m+1)$-tuple $(R/\mathfrak{b},\mathfrak{a}_1/\mathfrak{b},\dots,\mathfrak{a}_{m+1}/\mathfrak{b})$ induces an exact sequence

$$K_1(R/\mathfrak{b},\mathfrak{a}_1/\mathfrak{b},\dots,\mathfrak{a}_{m+1}/\mathfrak{b}) \to K_1(\overline{B}) \to K_1(MA).$$

The group $K_1(R/\mathfrak{b},\mathfrak{a}_1/\mathfrak{b},\dots,\mathfrak{a}_{m+1}/\mathfrak{b})$ is a quotient of $GL((\mathfrak{a}_1/\mathfrak{b})\cap\dots\cap(\mathfrak{a}_{m+1}/\mathfrak{b})) = \{1\}$, so α is injective. On the other hand, since the $(m+1)$-tuple A of ideals is normal, the identity on R induces an isomorphism $I(\overline{B}) \to I(MA)$ and hence also an isomorphism

$$GL(I(\overline{B})) \xrightarrow{\sim} GL(I(MA)).$$

Since the multirelative K_1 is a quotient of the general linear group of the underlying ideal, the map α is surjective. This proves:

THEOREM 1. *Let $A \in \mathcal{R}_m$ for $m \geq 1$. Then we have a functorial exact sequence*

$$\dots \to K_n(A) \to K_n(DA) \to K_n(MA) \to K_{n-1}(A) \to \dots \to K_0(MA).$$

□

7 AXIOMS FOR MULTIRELATIVE K-THEORY

It will be shown in this section that an axiomatic approach to multirelative K-theory is possible. We take some of the properties of multirelative K-groups as axioms and show that they determine all of multirelative K-theory.

AXIOMS

MULTIRELATIVE K-THEORY consists of functors

$$K_n\colon \mathcal{R}_m \to \mathcal{A} \qquad \text{for } m \text{ and } n \text{ integers} \geq 0,$$

morphisms

$$\delta\colon K_{n+1}M \to K_n \qquad (\text{for } m \text{ and } n \text{ integers} \geq 0)$$

of functors $\mathcal{R}_{m+1} \to \mathcal{A}$ and morphisms

$$\iota\colon K_n \to K_nD \qquad (\text{for } m \text{ and } n \text{ integers} \geq 0)$$

of functors $\mathcal{R}_{m+1} \to \mathcal{A}$, such that

(MK1) the following sequence is an exact sequence of functors $\mathcal{R}_{m+1} \to \mathcal{A}$ for all non-negative integers m and n

$$K_{n+1}D \xrightarrow{K_{n+1}\phi} K_{n+1}M \xrightarrow{\delta} K_n \xrightarrow{\iota} K_nD \xrightarrow{K_n\phi} K_nM.$$

(MK2) $K_n(R) = 0$ for all $n \geq 0$ and all free associative non-unital rings R,

(MK3) $K_0(A) = K_0(IA)$ for all $A \in \mathcal{R}_m$ for all m.

Loosely speaking, the multirelative K-groups are only defined for normal m-tuples of ideals and they fit into exact sequences the way one can expect, the (absolute) K-groups of free non-unital rings are trivial and the multirelative K_0 is just the Grothendieck group of the intersection of the ideals.

Let $(R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m)$ be a normal m-tuple of ideals. It induces an m-cube

$$I \mapsto R_I = R \Big/ \sum_{i \in I} \mathfrak{a}_i,$$

which is split in $\mathcal{S}$. Application of $\mathbf{Fr}$ gives an m-cube

$$I \mapsto \mathbf{Fr}(R_I)$$

of aspherical simplicial rings, which is dimensionwise split in $\mathcal{R}$.

PROPOSITION 7.1. *Let m and n be positive integers. Then the $(m+n)$-tuple*

$$\left(\mathbf{Fr}(R)_{n-1}, \mathbf{Fr}(R, \mathfrak{a}_1)_{n-1}, \ldots, \mathbf{Fr}(R, \mathfrak{a}_m)_{n-1}, \operatorname{Ker}\left(d_0^{(n-1)}\right), \ldots, \operatorname{Ker}\left(d_{n-1}^{(n-1)}\right)\right)$$

is normal.

Proof. First we show that the induced $(m+n)$-cube is

$$(I_1, I_2) \mapsto \mathbf{Fr}(R_{I_1})_{n-1-\#(I_2)},$$

where the cube is indexed by pairs of subsets of $\underline{m}$ and $[n-1]$. This set of pairs is ordered by componentwise inclusion:

$$(I_1, I_2) \leq (J_1, J_2) \iff I_1 \subseteq J_1 \quad \text{and} \quad I_2 \subseteq J_2.$$

The homomorphism

$$\mathbf{Fr}(R)_{n-1} \to \mathbf{Fr}(R_{I_1})_{n-1-\#(I_2)}$$

is the composition

$$\mathbf{Fr}(R)_{n-1} \to \mathbf{Fr}(R_{I_1})_{n-1} \to \mathbf{Fr}(R_{I_1})_{n-1-\#(I_2)},$$

the first map being induced by $\emptyset \subseteq I_1$ and the second by $[n-1] \setminus I_2 \subseteq [n-1]$. Both homomorphisms are surjective. The first one has kernel $\bigcap_{i \in I_1} \mathbf{Fr}(R, \mathfrak{a}_i)_{(n-1)}$ and the second one $\bigcap_{i \notin I_2} \operatorname{Ker}(d_i)$, where the d_i are face maps of $\mathbf{Fr}(R_{I_1})$. Since $\mathbf{Fr}(R)$ and $\mathbf{Fr}(R_{I_1})$ are both aspherical, elements of the second kernel can be lifted to elements of $\bigcap_{i \notin I_2} \operatorname{Ker}(d_i)$, where the d_i are face maps of $\mathbf{Fr}(R)$.

For the $(m+n)$-tuple to be normal it suffices that the intersections of the images of the $m+n$ ideals are preserved under the maps on the edges of the induced $(m+n)$-cube. These are the homomorphisms

$$\mathbf{Fr}(R_J)_l \to \mathbf{Fr}(R_{J \cup \{k\}})_l,$$

where $J \subseteq \underline{m}$, $k \in \underline{m} \setminus J$ and $l \in [n-1]$, and also the face maps

$$d_i \colon \mathbf{Fr}(R_J)_p \to \mathbf{Fr}(R_J)_{p-1},$$

where $p \in [n-1]$ and $0 \leq i \leq p$. Without loss of generality we may assume that $J = \underline{m}$, $l = n-1$ and $p = n-1$.

Because the m-cube $J \mapsto \mathbf{Fr}(R_J)$ is dimensionwise split we have short exact sequences

$$0 \to \bigcap_{i \in I \cup \{k\}} \mathbf{Fr}(R, \mathfrak{a}_i) \to \bigcap_{i \in I} \mathbf{Fr}(R, \mathfrak{a}_i) \to \bigcap_{i \in I} \mathbf{Fr}(R/\mathfrak{a}_k, \overline{\mathfrak{a}}_i) \to 0$$

of aspherical simplicial rings. It follows that for all $J \subseteq [n-1]$ we have

$$\bigcap_{i \in I} \mathbf{Fr}(R, \mathfrak{a}_i)_{n-1} \cap \bigcap_{j \in J} \mathrm{Ker}(d_j) = \bigcap_{j \in J} \mathrm{Ker}(d'_j),$$

where the d'_j are the face maps of $\bigcap_{i \in I} \mathbf{Fr}(R, \mathfrak{a}_i)$. Under $\mathbf{Fr}(R) \to \mathbf{Fr}(R/\mathfrak{a}_k)$ this maps onto

$$\bigcap_{j \in J} \mathrm{Ker}(d''_j) = \bigcap_{i \in I} \mathbf{Fr}(R/\mathfrak{a}_k, \overline{\mathfrak{a}}_i)_{n-1} \cap \bigcap_{j \in J} \mathrm{Ker}(d'''_j),$$

where the d''_j are the face maps of $\bigcap_{i \in I} \mathbf{Fr}(R/\mathfrak{a}_k, \overline{\mathfrak{a}}_i)$ and d'''_j those of $\bigcap_{i \in I} \mathbf{Fr}(R/\mathfrak{a}_k)$.

Because the simplicial rings $\bigcap_{i \in I} \mathbf{Fr}(R, \mathfrak{a}_i)$ are aspherical also the face maps $d_i \colon \mathbf{Fr}(R)_{n-1} \to \mathbf{Fr}(R)_{n-2}$ preserve intersections

$$\bigcap_{i \in I} \mathbf{Fr}(R, \mathfrak{a}_i)_{n-1} \cap \bigcap_{j \in J} \mathrm{Ker}(d_j).$$

□

Theorem 2. *Let $A = (R, \mathfrak{a}_1, \ldots, \mathfrak{a}_m) \in \mathcal{R}$. Then for all $n \geq 0$ it follows from the axioms (MK1) and (MK2) that $K_n(A)$ is naturally isomorphic to K_0 of the following object of $\mathcal{R}_{m+n}$:*

$$(\mathbf{Fr}(R)_{n-1}, \mathbf{Fr}(R, \mathfrak{a}_1)_{n-1}, \ldots, \mathbf{Fr}(R, \mathfrak{a}_m)_{n-1}, \mathrm{Ker}(d_0), \ldots, \mathrm{Ker}(d_{n-1})).$$

From axiom (MK3) it then follows that $K_n(A)$ is determined. So (MK1), (MK2) and (MK3) can be taken as axioms for the (multirelative) K-theory of rings.

Proof. The proof follows from the following three lemmas. □

Lemma 7.1. *Let $m \geq -1$ and $q, n \geq 0$. Then*

$$K_q(\mathbf{Fr}(R)_n, \mathbf{Fr}(R, \mathfrak{a}_1)_n, \ldots, \mathbf{Fr}(R, \mathfrak{a}_m)_n) = 0.$$

Proof. Since for $m \geq 0$ the $(m-1)$-tuples $D(A)$ and $M(A)$ are of the same type, the proof reduces by (MK1) to the case $m = -1$. For $m = -1$ the lemma follows from (MK2). □

Put

$$A[n,p] = (\mathbf{Fr}(R)_n, \mathbf{Fr}(R,\mathfrak{a}_1)_n, \ldots, \mathbf{Fr}(R,\mathfrak{a}_m)_n, \mathrm{Ker}(d_0), \ldots, \mathrm{Ker}(d_p)),$$

where $-1 \leq p \leq n$. It is an object of $\mathcal{R}_{m+p+1}$.

Lemma 7.2. *For all $p < n$ and all $q > 0$ we have*

$$K_q(A[n,p]) = 0.$$

Proof. For $p \geq 0$ we have

$$D(A[n,p]) = A[n,p-1] \quad \text{and} \quad M(A[n,p]) = A[n-1,p-1].$$

By (MK1) the problem reduces to the case $p = -1$, which is covered by the previous lemma. □

Lemma 7.3. *For all $q, n \geq 0$ we have*

$$K_q(A[n,n]) \cong K_{q+1}(A[n-1,n-1]).$$

Proof. This follows from (MK1) and the previous lemma. □

From this lemma the theorem follows:

$$K_n(A) = K_n(A[-1,-1]) \cong K_{n-1}(A[0,0]) \cong \cdots \cong K_0(A[n-1,n-1]).$$

References

[1] Bloch, S. and S. Lichtenbaum, *A Spectral Sequence for Motivic Cohomology*, preprint (K-theory Preprint Archives 62, 1995 Mar 3).

[2] Dayton, B.H. and C.A. Weibel, *A spectral sequence for the K-theory of affine glued schemes*, in "Algebraic K-theory", Proceedings Evanston 1980, Lecture Notes in Mathematics 854, Springer-Verlag, Berlin, 1981.

[3] Keune, Frans, *Nonabelian derived functors and algebraic K-theory*, in "Algebraic K-theory," Vol. I, Proceedings Seattle 1972, Lecture Notes in Mathematics 341, Springer-Verlag, Berlin, 1973.

[4] Levine, Marc, *The Weight Two K-Theory of Fields*, K-Theory 9, 443–501 (1995).

Frans Keune
Department of Mathematics
University of Nijmegen
Toernooiveld 1
6525 ED Nijmegen
The Netherlands
keune@sci.kun.nl

On the Average Values of the Irreducible Characters of Finite Groups of Lie Type on Geometric Unipotent Classes

Meinolf Geck

Received: August 16, 1996

Communicated by Wolfgang Soergel

Abstract. In 1980, Lusztig posed the problem of showing the existence of a unipotent support for the irreducible characters of a finite reductive group $G(\mathbb{F}_q)$. This is defined in terms of certain average values of the irreducible characters on unipotent classes. The problem was solved by Lusztig [16] for the case where q is a power of a sufficiently large prime. In this paper we show that, in general, these average values can be expressed in terms of the Green functions of G. In good characteristic, these Green functions are given by polynomials in q. Combining this with Lusztig's results, we can then establish the existence of unipotent supports whenever q is a power of a good prime.

1991 Mathematics Subject Classification: Primary 20C33, secondary 20G40.

1 Introduction

Let G be a connected reductive group defined over the finite field with q elements, and let $F : G \to G$ be the corresponding Frobenius map. We are interested in the average values of the irreducible characters of the finite group of Lie type G^F on the F-fixed points of F-stable unipotent classes of G. In 1980, Lusztig [9] has stated the following problem.

Problem 1.1 *Let ρ be an irreducible character of G^F. Show that there exists a unique F-stable unipotent class C of maximal possible dimension such that the average value of ρ on C^F is non-zero, that is,*

$$\sum_{j=1}^{r}[G^F : C_G(u_j)^F]\,\rho(u_j) \neq 0,$$

where $u_1, \ldots, u_r \in G^F$ are representatives for the G^F-conjugacy classes contained in C^F and $C_G(u_j)$ denotes the centralizer of u_j. If this is the case, we call C the unipotent support *of ρ.*

In 1992, Lusztig [16] addressed this problem in the framework of his theory of character sheaves and its application to Kawanaka's theory [8] of generalized Gelfand-Graev representations. In this context, one is lead to consider the following related question.

PROBLEM 1.2 *Let ρ be an irreducible character of G^F. Show that there exists a unique F-stable unipotent class C of maximal possible dimension such that*

$$\sum_{j=1}^{r}[A(u_j) : A(u_j)^F]\,\rho(u_j) \neq 0,$$

where $A(u_j)$ denotes the group of components of $C_G(u_j)$.

Assuming that q is a sufficiently large power of a sufficiently large prime p, Lusztig proves in [16], (9.11), a formula which expresses a 'modified' average value as above in terms of the scalar products of the Alvis–Curtis–Kawanaka dual of ρ with the characters of the various generalized Gelfand-Graev representations corresponding to C. (The bound on p comes from the condition that, roughly speaking, one wants to operate with the Lie algebra of G as if it were in characteristic 0.) It is then an easy consequence of [16], Theorem 11.2, that Problem 1.2 has a positive solution.

Using the results in [16] and [6], we shall prove in Proposition 2.5 below a formula which expresses an average value as in Problem 1.1 in similar terms as above. Then the solution of Problem 1.1[1] also is an easy and formally completely analogous consequence of [16], Theorem 11.2. For this argument we have to assume, as in [*loc. cit.*], that q and p are large enough. It is one purpose of this paper to show that this condition on p can be relaxed so that Problem 1.1 and Problem 1.2 have a positive solution (and yield the same unipotent class) whenever p is a good prime for G. It may be true that, eventually, no condition on p will be needed but this seems to require some new arguments. (I have checked, using [19], that things go through for exceptional groups in characteristic $p \neq 2$. A more detailed discussion of the bad characteristic case appears in [7], where it is shown that Problem 1.2 always has a positive solution – Problem 1.1 in bad characteristic remains open.)

The idea of our argument is as follows. It is clear that an average value as in Problem 1.1 is given by the scalar product of ρ with the class function f_C on G^F such that

$$f_C(g) = \begin{cases} |G^F| & \text{if } g \in C^F, \\ 0 & \text{if } g \in G^F \setminus C^F. \end{cases}$$

A similar interpretation can also be given for the modified average value in Problem 1.2, using the class function f'_C on G^F with support on C^F and such that

$$f'_C(u_j) = [A(u_j) : A(u_j)^F]|C_G(u_j)^F| \quad \text{for } 1 \leq j \leq r,$$

where (as above) $u_1, \ldots, u_r \in C^F$ are representatives for the G^F-classes contained in C^F.

The statement concerning f_C in the following result was already conjectured by Lusztig in [9], (2.16). For large p, it follows easily from the results on Green functions in [17] (see also Kawanaka [8], (1.3.8)).

[1] Lusztig has informed me that this solution was known to him, but it was not included in [16]

Proposition 1.3 *The functions f_C and f'_C are uniform, that is, they can be written as linear combinations of various Deligne-Lusztig generalized characters $R^G_{T,\theta}$.*

The proof of this result in (3.6) below will be based on Proposition 3.5, where we show that the known algorithm for computing the ordinary Green functions in [17] works without any restriction on p and q. This may also be of independent interest. It uses heavily the description of this algorithm in terms of Lusztig's character sheaves in [13], Section 24. The (mild) restrictions on p in [*loc. cit.*] can be removed by using Shoji's results [18] on cuspidal character sheaves in bad characteristic and the fact, also proved in [18], that the ordinary Green functions of G^F always coincide with those defined in terms of character sheaves.

It then follows that in order to compute our average values we only need to consider the uniform projection of ρ. We can also reduce to the case where G has a connected center and is simple modulo its center, see Lemmas 5.1 and 5.2. Then our average values can be expressed as linear combinations of Green functions of G^F where the coefficients are 'independent of q', by [11], Main Theorem 4.23. Up to this point we don't need any assumption on p or q.

Let now q be a power of a prime p which is good for G. Recall that this is the case if p is good for each simple factor involved in G, and that the conditions for the various simple types are as follows.

$$\begin{array}{rl} A_n: & \text{no condition,} \\ B_n, C_n, D_n: & p \neq 2, \\ G_2, F_4, E_6, E_7: & p \neq 2, 3, \\ E_8: & p \neq 2, 3, 5. \end{array}$$

Then the Green functions of G^F are given by evaluating certain well-defined polynomials at q (see [17]), and we obtain a similar statement for our average values. We can then replace a given q by a power of a larger prime p for which the results in [16] are applicable and thus deduce results about these average value polynomials being zero or not. Finally, we deduce from the formulae in Proposition 2.5 that our polynomials have the property that if one of them is non-zero then its evaluation at every prime power is non-zero. The details and the precise formulation of this argument can be found in Section 4, especially Proposition 4.4. Then the main result of this paper will be established in Section 5.

Theorem 1.4 *Assume that q is a power of a good prime p for G. Let ρ be an irreducible character of G^F.*

(a) Both Problem 1.1 and Problem 1.2 have a positive solution for ρ, and they yield the same unipotent class, C say.

(b) The p-part in the degree of ρ is given by q^d where d is the dimension of the variety of Borel subgroups of G containing a fixed element in C.

The characterization of the p-part of ρ in terms of C was also conjectured in [9]. Lusztig [16] proves the following refinement (again assuming that q is a power of a large enough prime): Let $g \in G^F$ be any element such that $\rho(g) \neq 0$. Then the unipotent part of g lies in the unipotent support C of ρ or in a unipotent class of strictly smaller dimension than C. Note that it is not clear how to pass from results about the vanishing or non-vanishing of individual character values to results about the non-vanishing of average values.

We remark that, as far as this refinement is concerned, the situation definitely is different in the bad characteristic case. Consider, for example, the simple group G of type G_2 defined over a finite field of characteristic 3. Let C be the class of regular unipotent elements. Then there exist unipotent characters of G^F which are non-zero on some element in C^F but whose average value on C^F is zero (see the character table in [5]).

Completing earlier results of Lusztig's (see [14]), A.-M. Aubert [1] has shown that such a refinement holds for classical groups in good characteristic and with g unipotent. For that purpose, one has to use the full power of the theory of character sheaves and Shoji's proof of Lusztig's conjecture about almost characters and characteristic functions of character sheaves (see [18]). I have checked that this also works for exceptional groups in good characteristic. This will be discussed elsewhere.

I thank A.-M. Aubert for carefully reading earlier versions of this paper.

2 Generalized Gelfand-Graev representations and average values

Let G be a connected reductive group defined over $\mathbb{F}_q$, with corresponding Frobenius map F. All of our characters and class functions will have values in an algebraic closure of $\mathbb{Q}_l$, where l is prime not dividing q. If f, f' are two class functions on G^F we denote by

$$(f, f') := \frac{1}{|G^F|} \sum_{g \in G^F} f(g)\overline{f'(g)}$$

their usual hermitian product, where $x \mapsto \bar{x}$ is a field automorphism which maps roots of unity to their inverses. We denote by G_{uni} the set of unipotent elements in G. For each element $g \in G$ we let C_g denote the G-conjugacy class of g. There is a canonical partial order on the set of unipotent classes of G: if C, C' are two such classes we write $C \leq C'$ if C is contained in the Zariski closure of C'. We write $C < C'$ if $C \leq C'$ but $C \neq C'$.

2.1 Unipotently supported class functions on G^F

Let C be an F-stable unipotent class in G. Let $u \in C^F$ and $A(u)$ be the group of components of $C_G(u)$. If we twist u with any element $y \in A(u)$ we obtain an element $u_y \in C^F$, well-defined up to G^F-conjugacy. If we choose representatives for the F-conjugacy classes of $A(u)$ we obtain in this way a full set of representatives of the G^F-classes contained in C^F; denote such a set of representatives by $u_1, \ldots, u_r \in C^F$, where we let $u_1 = u$.

Let $I(C)^F$ be the set of pairs $i = (C, E)$ where E is an irreducible representation of $A(u)$ over $\bar{\mathbb{Q}}_l$ (given up to isomorphism) for which there exists an automorphism $\alpha_E : E \to E$ of finite order such that $\alpha_E \circ y = F(y) \circ \alpha_E$ for all $y \in A(u)$. We define a class function $Y_i : G^F \to \bar{\mathbb{Q}}_l$ by

$$Y_i(g) = \begin{cases} \text{Trace}(\alpha_E \circ y, E) & \text{if } g \text{ is } G^F\text{-conjugate to } u_y \text{ for some } y \in A(u), \\ 0 & \text{otherwise.} \end{cases}$$

These functions form a basis of the space of class functions of G^F with support on C^F. (Note that they are only well-defined up to non-zero scalar multiples.) For each j let $a_j := |A(u_j)^F|$. The order of $A(u_j)$ is independent of j; we denote it by a. With this notation we have the following orthogonality relations:

$$\sum_{j=1}^{r} \frac{a}{a_j} \overline{Y_i(u_j)} Y_{i'}(u_j) = a\delta_{ii'} \quad \text{and} \quad \sum_{i \in I_0(C)^F} \overline{Y_i(u_j)} Y_i(u_{j'}) = a_j \delta_{jj'},$$

for all $i, i' \in I(C)^F$, or all $1 \leq j, j' \leq r$, respectively.

The trivial module for $A(u)$ always satisfies the above condition. The corresponding pair will be denoted $i_0 = (C, \bar{\mathbb{Q}}_l)$, and the isomorphism α_E can be chosen so that the function Y_{i_0} is identically 1 on C^F. Thus, we have

$$f_C = |G^F| Y_{i_0} \quad \text{with} \quad i_0 = (C, \bar{\mathbb{Q}}_l).$$

On the other hand, using the definition of f_C' and the above orthogonality relations we compute that

$$(f_C', Y_i) = a\delta_{i,i_0} \quad \text{for all } i \in I(C)^F.$$

Note that these relations determine f_C' uniquely.

2.2 GGGR's

Recall that if q is a power of a good prime for G then Kawanaka [8] has defined generalized Gelfand-Graev representations (GGGR's for short) for every unipotent class in G^F. (Usually, we will identify a GGGR with its character.) Very roughly, this is done as follows. Let C be an F-stable unipotent class in G. Using the corresponding weighted Dynkin diagram we can associate with C a pair of unipotent subgroups $U_2 \subseteq U_1$ where U_1 is the unipotent radical of an F-stable parabolic subgroup P of G and U_2 is an F-stable closed normal subgroup in P. Furthermore, $C \cap U_2$ is dense in U_2 and the centralizer in G of any element $u \in U_2 \cap C$ is already contained in P. (Note that Kawanaka [8] has checked that these statements indeed are true whenever the characteristic is good.) Hence the subgroup U_2^F contains representatives for all G^F-classes in C^F. Using a Killing type form on U_2 we can associate with each such representative $u \in C \cap U_2^F$ a certain linear character λ_u of U_2^F such that

$$\mathrm{Ind}_{U_2^F}^{G^F}(\lambda_u) = [U_1^F : U_2^F]^{1/2} \Gamma_u,$$

where Γ_u is the GGGR associated with u. With the notation in (2.1), we can assume that $u_j \in U_2$ for $1 \leq j \leq r$. As in [16], (7.5), we define the following 'twisted' version of GGGR's:

$$\Gamma_i = \sum_{j=1}^{r} \frac{a}{a_j} Y_i(u_j) \Gamma_{u_j} \quad \text{for } i \in I(C)^F.$$

2.3 Basic properties of GGGR's

We shall need two basic properties of GGGR which we now explain. Denote by D_G the Alvis–Curtis–Kawanaka duality operation on the character ring of G^F.

Assume that p and q are large enough so that the results in [16] are applicable.

(a) *For all* $g \in G^F_{\text{uni}}$ *we have*

$$D_G(\Gamma_i)(g) \neq 0 \quad \Rightarrow \quad C \leq C_g.$$

(b) *For all* $i, i' \in I(C)^F$ *we have*

$$(D_G(\Gamma_i), Y_{i'}) = a\zeta_i' q^{d_i} \delta_{i,i'}$$

where ζ_i' *is a certain 4-th root of unity and* d_i *is half a certain integer.*

Proofs of (a) and (b) can be obtained by combining [16], (8.6), with [16], (6.13)(i), and (6.13)(iii), respectively. Properties (a) and (b) are also contained in [6], Corollary 3.6(b) and Lemma 3.5. (Actually, the formula in the latter reference involves a certain function $X_{i'}$ instead of $Y_{i'}$, but $X_{i'}$ is zero on all elements $g \in G^F_{\text{uni}}$ unless $C_g \leq C$ and coincides with $Y_{i'}$ on C^F; using (a) we can therefore take $Y_{i'}$.) Note also that in [16] it is generally assumed that G is a split group, and the results in [6] referred to above are also proved under this assumption. However, by [16], (8.7), everything goes through for non-split groups as well, with only minor changes. Especially, properties (a) and (b) remain valid. Finally, we have the following special property of the numbers ζ', d_i appearing in (b).

(c) *If* $i_0 = (C, \bar{\mathbb{Q}}_l)$ *then* $\zeta_{i_0}' = 1$ *and* $d_{i_0} = -d$ *where* d *is the dimension of the variety of Borel subgroups of* G *containing* u.

For the proof see [6], Lemma 3.5, and the remarks concerning equation (a) in the proof of [16], Theorem 11.2. We also use the formula $\dim G - \dim C = \text{rank}(G) + 2d$ (see [3], Theorem 5.10.1).

LEMMA 2.4 *Assume that* p *and* q *are large enough so that the results in [16] are applicable. Let* f_C *and* f_C' *be the functions introduced in Section 1. Then the following hold.*

$$\begin{aligned} f_C(g) &= q^d \sum_{j=1}^{r} [G^F : C_G(u_j)^F]\, D_G(\Gamma_{u_j})(g) \qquad \textit{for all } g \in C^F, \\ f_C'(g) &= q^d D_G(\Gamma_{i_0})(g) = q^d \sum_{j=1}^{r} \frac{a}{a_j} D_G(\Gamma_{u_j})(g) \quad \textit{for all } g \in C^F. \end{aligned}$$

Proof. Let $i \in I(C)^F$ and Y_i the corresponding class function as in (2.1). Since the various functions Y_i form a basis of the space of class functions on G^F with support on C^F it will be sufficient to show that the scalar product of Y_i with the left and right hand sides of the above expressions are equal.

Consider at first f_C. The scalar product with the left hand side is just (f_C, Y_i). On the other hand, using the orthogonality relations in (2.1), we conclude that

$$\Gamma_{u_j} = \frac{1}{a} \sum_{i' \in I(C)^F} \overline{Y_{i'}(u_j)} \Gamma_{i'} \qquad \text{for all } 1 \leq j \leq r.$$

Inserting this into the expression on the right hand side we obtain that

$$\begin{aligned}(\text{r.h.s.}, Y_i) &= q^d \sum_{j=1}^{r} \frac{1}{a}[G^F : C_G(u_j)^F] \sum_{i' \in I(C)^F} \overline{Y_{i'}(u_j)}(D_G(\Gamma_{i'}), Y_i) \\ &= q^{d+d_i}\zeta_i' \sum_{j=1}^{r}[G^F : C_G(u_j)^F]\, \overline{Y_i(u_j)} \quad \text{by (2.3)(b)} \\ &= q^{d+d_i}\zeta_i'(f_C, Y_i) \quad \text{by definition of the scalar product.}\end{aligned}$$

Hence it remains to prove that if $(f_C, Y_i) \neq 0$ then $\zeta_i' = 1$ and $d_i = -d$. This follows from the fact that the set $I(C)^F$ can be partitioned into 'blocks' according to the generalized Springer correspondence (see [16], (4.4)) and that the scalar product between $(Y_i, Y_{i'})$ is zero unless i, i' lie in the same block (see [16], (6.5)). Now remember that $f_C = |G^F|Y_{i_0}$. Hence, if $(f_C, Y_i) \neq 0$ then i lies in the same block as i_0. In this case, $\zeta_i' = \zeta_{i_0}'$ and $d_i = d_{i_0}$ by [6], Lemma 3.5. So we are done by (2.3)(c).

Now consider f_C'. By (2.1) we have $(f_C', Y_i) = a\delta_{i,i_0}$. The scalar product with the right hand side evaluates to the same expression using (2.3)(b) and (c). □

Proposition 2.5 (Cf. [16], (9.11)) *Assume that p and q are large enough so that the results in [16] are applicable. Let ρ be an irreducible character of G^F such that*

(*) $\rho(g) = 0$ for all $g \in G^F_{\text{uni}}$ with $C < C_g$.

Then we have

$$\begin{aligned}(\rho, f_C) = \sum_{j=1}^{r}[G^F : C_G(u_j)^F]\, \rho(u_j) &= q^d \sum_{j=1}^{r}[G^F : C_G(u_j)^F]\, (\Gamma_{u_j}, D_G(\bar{\rho})), \\ (\rho, f_C') = \sum_{j=1}^{r}[A(u) : A(u_j)^F]\, \rho(u_j) &= q^d \sum_{j=1}^{r}[A(u) : A(u_j)^F]\, (\Gamma_{u_j}, D_G(\bar{\rho})).\end{aligned}$$

Since these expressions are rational integers the above formulae are also valid with ρ instead of $\bar{\rho}$ on the right hand side.

Proof. It is clear that in order to evaluate the left hand sides of the above expressions we only need to know the values of ρ on C^F. Let us check that the same also holds for the expressions on the right hand side. We start by looking at the scalar product of $\bar{\rho}$ with $D_G(\Gamma_i)$, for $i \in I(C)^F$, that is, the expression

$$(D_G(\Gamma_i), \bar{\rho}) = \frac{1}{|G^F|} \sum_{g \in G^F} D_G(\Gamma_i)(g)\rho(g).$$

First, the sum need only be extended over $g \in G^F_{\text{uni}}$ since Γ_i, and hence also its dual, is zero outside G^F_{uni}. Now assume that $g \in G^F_{\text{uni}}$ gives a non-zero contribution to the above sum. On one hand, by (2.3)(a), we must have $C \leq C_g$. On the other hand, our assumption (*) then forces $C = C_g$. Hence, in order to evaluate the above scalar product we only need to look at the values of ρ and $D_G(\Gamma_i)$ on C^F. A similar remarks holds, of course, if we consider Γ_{u_j} instead of Γ_i. Using the self-adjointness of D_G

we can therefore conclude that the right hand sides of our desired equalities are also determined by the restriction of ρ to C^F.

To complete the proof, it remains to use the expressions for f_C and f'_C which are given in Lemma 2.4. $\square$

COROLLARY 2.6 (LUSZTIG) *Assume that p and q are large enough so that the results in [16] are applicable. Let ρ be an irreducible character of G^F. Then both Problem 1.1 and Problem 1.2 have a positive solution for ρ, and the corresponding unipotent classes are equal.*

Proof. (Compare with the argument in the last part of the proof of [16], Theorem 11.2.) Let ρ' be the irreducible character such that $\rho' = \pm D_G(\rho)$. By [16], Theorem 11.2, there exists an F-stable unipotent class C such that the following two conditions hold (among others).

(1) There exists some $u \in C^F$ such that $(\Gamma_u, \rho') \neq 0$.

(2) If C' is an F-stable unipotent class such that $\rho(g) \neq 0$ for some $g \in {C'}^F$ then $\dim C' \leq \dim C$ with equality only if $C = C'$.

We show that C satisfies the requirements for both Problem 1.1 and Problem 1.2.

If C' is some F-stable unipotent class such that an average value on ${C'}^F$ as in Problem 1.1 or Problem 1.2 is non-zero then ρ has a non-zero value on some element in ${C'}^F$ and (2) implies that $\dim C' \leq \dim C$.

Recall that our average values are given by (ρ, f_C) and (ρ, f'_C), respectively. It remains to prove that these two scalar products are non-zero. By (2), assumption (*) in Proposition 2.5 is satisfied. So we have

$$\begin{aligned} (\rho, f_C) &= \pm q^d \sum_j [G^F : C_G(u_j)^F]\,(\Gamma_{u_j}, \rho'), \\ (\rho, f'_C) &= \pm q^d \sum_j [A(u_j) : A(u_j)^F]\,(\Gamma_{u_j}, \rho'). \end{aligned}$$

In both cases all terms in the sums on the right hand sides are non-negative and at least one of them is non-zero by (1). Hence there are no cancellations and the left hand sides must be non-zero, too. This completes the proof. $\square$

EXAMPLE 2.7 Assume that p and q are large enough so that the above results are applicable. Let ρ be an irreducible character of G^F and C its unipotent support. The assumption (*) in Proposition 2.5 is satisfied (see Property (2) in the proof of Corollary 2.6). Assume that the centralizer of an element in C is connected. In this case we have $r = 1$ in the formulae in Proposition 2.5 and the left and the right hand sides contain just one summand. So we find that

$$\rho(u) = q^d(\Gamma_u, D_G(\rho)).$$

In particular, the character value $\rho(u)$ is an integer divisible by q^d. In fact, using a similar argument as in [6], Proposition 5.4, one can show that the scalar product between $D_G(\rho)$ and Γ_u must be ± 1. Hence we have

$$\rho(u) = \pm q^d \qquad \text{where the sign is such that } \pm D_G(\rho)(1) > 0.$$

Theorem 1.4 and the results in Section 4 will imply that this last formula holds whenever q is a power of a good prime p. We omit further details.

3 Average values and uniform functions

The first aim of this section is to prove Proposition 1.3. We then derive in Corollary 3.8 a formula for the scalar products of an irreducible character of G^F with the functions f_C and f'_C for an F-stable unipotent class C. This will be in terms of Lusztig's parametrization of irreducible characters in [11].

We shall now introduce some notation and recall some facts from [13] which will be needed for the proof of Proposition 1.3. With each F-stable maximal torus T in G, we can associate two types of Green functions: one is the ordinary Green function Q_T^G defined as the restriction of a corresponding Deligne-Lusztig generalized character to G_{uni}^F; the other is a special case of a more general construction of generalized Green functions which are defined in terms of characteristic functions of F-stable character sheaves on G (see [13], (8.3.1)). Shoji has shown in [18], Theorem 5.5 (part II), that these two types of Green functions coincide (without any restriction on p or q).

We shall need some more detailed properties about the values of Green functions. For this purpose we take a closer look at Lusztig's algorithm in [13], Theorem 24.4, for the computation of all generalized Green functions. The properties that we need can be obtained from this algorithmic description. However, there is a mild restriction on p in [*loc. cit.*] which comes from the fact that certain properties of character sheaves on G are not yet established in complete generality. We will now go through [13], Section 24, and check that everything works without any restriction on p, if we only consider those generalized Green functions which correspond to the ordinary Green functions. This will use in an essential way Shoji's results in [18] on cuspidal character sheaves in bad characteristic.

3.1 The generalized Springer correspondence

Let I be the set of all pairs $(C, \mathcal{E})$ where C is a unipotent class in G and $\mathcal{E}$ is an irreducible G-equivariant $\bar{\mathbb{Q}}_l$-local system, given up to isomorphism. If $i = (C, \mathcal{E})$ and $i' = (C', \mathcal{E}')$ are elements in I we write $i \leq i'$ if $C \leq C'$, and $i \sim i'$ if $C = C'$. With each pair $i \in I$ there is associated a triple $(L, C_1, \mathcal{E}_1)$ consisting of a Levi subgroup L in some parabolic subgroup of G and $(C_1, \mathcal{E}_1)$ is a pair like i for L, but where $\mathcal{E}_1$ is 'cuspidal' (in the sense of [12]). The pairs in I associated with a fixed triple as above are parameterized by the irreducible characters of a group $W_G(L, C_1, \mathcal{E}_1)$ which is the inertia group of the pair $(C_1, \mathcal{E}_1)$ in the normalizer of L. This correspondence is the generalized Springer correspondence defined and studied in [12].

A pair $i \in I$ which corresponds to a triple where the Levi subgroup L is a maximal torus, the class C_1 is the trivial class and the local system $\mathcal{E}_1$ is trivial, will be called *uniform* (see the remark following [13], Theorem 24.4). In this case, the inertia group $W_G(T, \{1\}, \bar{\mathbb{Q}}_l)$ is nothing but the Weyl group of G with respect to T, and the above correspondence reduces to Springer's original correspondence. We will denote by I_0 the subset of I consisting of uniform pairs.

Remark 3.2 Let $i_0 = (C, \bar{\mathbb{Q}}_l) \in I$ where $\bar{\mathbb{Q}}_l$ denotes the trivial local system. Then i_0 is uniform.

Proof. This is a general property of the generalized Springer correspondence. Let $i = (C, \mathcal{E}) \in I$ and $\mathcal{B}_u^G$ be the variety of Borel subgroups containing a fixed element $u \in C$.

Recall from [12] that $\mathcal{E}$ corresponds to an irreducible representation of $A(u)$, and that i is uniform if and only if that representation appears with non-zero multiplicity in the permutation representation of $A(u)$ on the irreducible components of $\mathcal{B}_u^G$.

Now the trivial local system on C corresponds to the trivial representation of $A(u)$, and this certainly appears with non-zero multiplicity in any permutation representation of $A(u)$. Hence $i_0 = (C, \bar{\mathbb{Q}}_l)$ is uniform. $\square$

3.3 Basic relations

The Frobenius map F acts naturally on I. An F-stable pair $i = (C, \mathcal{E}) \in I^F$ gives rise to a pair in $I(C)^F$ as in (2.1) and hence to a function Y_i (cf. the proof of [13], (24.2.7).) This function can be extended to a function X_i on the Zariski closure of C by the construction in [10], (24.2.8), so that we have equations of the form

$$X_i = \sum_{i' \in I^F} P_{i',i} Y_{i'} \quad \text{ with } P_{i',i} \in \bar{\mathbb{Q}}_l \text{ for all } i, i' \in I^F,$$

and where $P_{i,i} = 1$ and $P_{i',i} = 0$ if $i' \not\leq i$ or if $i' \sim i$, $i' \neq i$. Now we also have 'contragredient' versions of these functions which will be denoted by $\tilde{X}_i$ and $\tilde{Y}_i$ (see [13], (24.2.12) and (24.2.13)). We have $\tilde{Y}_i = \overline{Y}_i$, see [13], (25.6,4). Correspondingly, we have similar equations as above with coefficients $\tilde{P}_{i',i}$. The various class functions introduced so far are only well-defined up to some scalar multiple, but [13], (24.2.1) and (24.2.2), singles out a certain 'good' normalization which we also assume chosen here. Finally, we define

$$\lambda_{i,i'} := (Y_i, Y_{i'}) \quad \text{ and } \quad \omega_{i,i'} := \frac{1}{|G^F|} \sum_{g \in G^F_{\mathrm{uni}}} X_i(g) \tilde{X}_{i'}(g) \quad \text{ for all } i, i' \in I^F.$$

As in [13], (24.3), we see that $\lambda_{i,i'} = 0$ unless $i \sim i'$, and that the matrix $(\lambda_{i,i'})_{i,i' \in I^F}$ is invertible. (The functions Y_i form a basis of the space of class functions on G^F_{uni}.) We obtain the following basic relations:

$$\sum_{i'_1, i'_2 \in I^F} P_{i'_1,i_1} \tilde{P}_{i'_2,i_2} \lambda_{i'_1,i'_2} = \omega_{i_1,i_2} \quad \text{ for all } i_1, i_2 \in I^F.$$

Theorem 24.4 in [13] states that the coefficients $P_{i',i}$, $\tilde{P}_{i',i}$ and $\lambda_{i',i}$ are determined by this system of equations once the right hand side coefficients $\omega_{i',i}$ are known. Now, under some mild restriction on p, the coefficient $\omega_{i',i}$ is given by the equation [13], (24.3.4) (arising from a scalar product formula for characteristic functions of character sheaves). In general, we can at least obtain the following information.

Lemma 3.4 *Assume that the center of G is connected. Let $i \in I_0^F$ and $i' \in I^F$. Then the following hold.*

(i) If $i' \notin I_0^F$ then $\omega_{i',i} = \omega_{i,i'} = 0$.

(ii) If $i' \in I_0^F$ then $\omega_{i',i} = \omega_{i,i'}$ is a rational number.

Proof. Given any $i, i' \in I^F$ the relevant scalar product formula for the evaluation of $\sum_g X_i(g)\tilde{X}_{i'}(g)$ (sum over all $g \in G^F_{\mathrm{uni}}$) can be found in [13], Theorem 10.9. Let $(L, C_1, \mathcal{E}_1)$ and $(L', C'_1, \mathcal{E}'_1)$ be the triples associated with i and i', and K_1, K'_1 the corresponding cuspidal perverse sheaves on L, L', respectively. One of the assumptions for the validity of [13], Theorem 10.9, is that K_1 and K'_1 must be 'strongly cuspidal' (see the description of these assumptions in [13], (10.7)).

We claim that a cuspidal perverse sheaf on any group G with a connected center is always strongly cuspidal (hence in particular K_1 and K'_1; note that L, L' also have a connected center). This can be seen as follows. By [13], (7.1.6), it is sufficient to show that a cuspidal perverse sheaf on G is a character sheaf. By the reduction arguments in [13], (17.10) and (17.11), we can reduce to the case where G is simple of adjoint type. If p is an almost good prime the result is already covered by [13], Theorem 23.1(b). For G of type E_6 or E_7, see [13], Proposition 20.3. It remains to consider G of type G_2, F_4, E_8. The result in this case is contained in [18], Theorem 7.3(a) in part I and Proposition 5.3 in part II. So our claim is established.

Another assumption for the validity of [13], Theorem 10.9, is that if L, L' are conjugate in G then K_1, K'_1 must be 'clean' (see again [13], (7.7)). Now if $i, i' \in I_0$ then both L and L' are maximal tori and the 'cleanness' is clear (we have to consider the trivial local system on the trivial class). If one of i, i' is uniform and the other is not, then one of the Levi subgroups L, L' is a maximal torus and the other is not, hence the above condition is vacuous. In combination with the 'good' normalization of $X_i, X_{i'}$ mentioned in (3.3), this proves both (i) and (ii) (cf. [13], (24.3.5)). $\square$

Now we can state the analogue of [13], Theorem 24.4, for uniform pairs $i \in I_0$.

Proposition 3.5 *Assume that the center of G is connected. Let $i \in I_0^F$ and $i' \in I^F$. Then the following hold.*

(i) $P_{i',i} = \tilde{P}_{i',i}$ and $\lambda_{i',i} = \lambda_{i,i'}$ are rational numbers.

(ii) $P_{i',i}$ and $\lambda_{i',i}$ are zero if $i' \notin I_0^F$.

Moreover, the coefficients $P_{i',i}$ and $\lambda_{i',i}$ (for $i, i' \in I_0^F$) are determined from the basic relations in (3.3) by an algorithm as described in [13], Theorem 24.4 or [17], Remark 5.4.

Proof. This is almost completely analogous to the proof of [13], Theorem 24.4, with some minor changes concerning the ordering of the arguments. We will go through that proof and check that things go through as desired for uniform pairs in I_0^F. For any integer δ consider the following two statements.

(A_δ) If $i' = (C', \mathcal{E}') \in I^F$ with $\dim C' \leq \delta$ and if $i \in I^F$, then $P_{i',i} = \tilde{P}_{i',i}$ is a rational number if i or i' is uniform, and it is zero if one of i, i' is uniform and the other is not.

(B_δ) If $i' = (C', \mathcal{E}') \in I^F$ with $\dim C' \leq \delta$ and if $i \in I^F$, then $\lambda_{i',i} = \lambda_{i,i'}$ is a rational number if i or i' is uniform, and it is zero if one of i, i' is uniform and the other is not.

It is clear that these statements are true if $\delta < 0$. As in [*loc. cit.*] we first show that *if $\delta \geq 0$ and $(A_{\delta-1})$, (B_δ) are true then (A_δ) is true.*

Let us just describe this in more detail. Let $i \in I^F$ and $i' \in I^F$ such that $\dim C' = \delta$. If $i' \not\leq i$ or if $i \sim i'$, $i \neq i'$ then $P_{i',i} = \tilde{P}_{i',i} = 0$. So we may assume that $i' < i$. From the basic relations in (3.3) we derive, as in [*loc. cit.*], the following equations for any $a \in I^F$ with $a \sim i'$.

$$\begin{aligned}
\sum_{i'_2 \sim i'} \tilde{P}_{i'_2,i}\lambda_{a,i'_2} &= \omega_{a,i} - \sum_{i'_1<i',i'_2\sim i'_1} P_{i'_1,a}\tilde{P}_{i'_2,i}\lambda_{i'_1,i'_2},\\
\sum_{i'_2 \sim i'} P_{i'_2,i}\lambda_{i'_2,a} &= \omega_{i,a} - \sum_{i'_1<i',i'_2\sim i'_1} P_{i'_2,i}\tilde{P}_{i'_1,a}\lambda_{i'_2,i'_1}.
\end{aligned}$$

We denote the right hand sides of these two equations by $\tilde{r}(a)$ and $r(a)$, respectively. We claim that

(1) if a and i are uniform then $r(a) = \tilde{r}(a)$ is a rational number, and

(2) if one of a, i is uniform and the other is not then $\tilde{r}(a) = r(a) = 0$.

This is proved as follows. Lemma 3.4 shows that it is sufficient to consider the sum over i'_1, i'_2 in each of the defining equations for $r(a)$ and $\tilde{r}(a)$. At first let us consider $r(a)$, and assume that there exists some i'_1, i'_2 such that the corresponding term is non-zero. Then $P_{i'_2,i} \neq 0$, $\tilde{P}_{i'_1,a} \neq 0$, and $\lambda_{i'_2,i'_1} \neq 0$. For each of these terms we can apply $(A_{\delta-1})$ or $(B_{\delta-1})$. If one of a, i is uniform and the other is not we obtain a contradiction; while if both of a, i are uniform we obtain a summand which is a rational number. We can argue similarly for $\tilde{r}(a)$. Moreover, if both a and i are uniform this analysis shows that $r(a) = \tilde{r}(a)$ is a rational number. Our claim is proved.

We have already mentioned above that the matrix of coefficients $(\lambda_{a,a'})$ (where $a, a' \in I^F$, $a \sim a' \sim i'$) is invertible. Let $(\lambda'_{a,a'})$ be the coefficients in the inverse of this matrix. Then we obtain that

$$\begin{aligned}
\tilde{P}_{i',i} &= \sum_{i'_2\sim i'} \tilde{P}_{i'_2,i}\left(\sum_{a\sim i'} \lambda'_{i',a}\lambda_{a,i'_2}\right) = \sum_{a\sim i'} \tilde{r}(a)\lambda'_{i',a},\\
P_{i',i} &= \sum_{i'_2\sim i'} P_{i'_2,i}\left(\sum_{a\sim i'} \lambda_{i'_2,a}\lambda'_{a,i'}\right) = \sum_{a\sim i'} r(a)\lambda'_{a,i'}.
\end{aligned}$$

By (B_δ) we know that $\lambda_{a,a'}$ is zero if one of a, a' is uniform and the other is not; moreover, $\lambda_{a,a'} = \lambda_{a',a}$ is a rational number if both a, a' are uniform. It follows that the matrix of coefficients $(\lambda'_{a,a'})$ has the analogous properties. Hence, if i' is uniform (respectively, not uniform) we can restrict the above sums to those a which are also uniform (respectively, not uniform).

Now assume that both i, i' are uniform. As we have just seen, we can assume that a in the above sums is uniform, and then $r(a) = \tilde{r}(a)$ by (1). Moreover, $\lambda'_{i',a} = \lambda'_{a,i'}$ is a rational number. Hence also $P_{i',i} = \tilde{P}_{i',i}$ is a rational number.

Next assume that i is uniform and i' is not uniform. We can now assume that a in the above sums is not uniform. By (2), we know that then both $r(a)$ and $\tilde{r}(a)$ are zero. Hence $P_{i',i} = \tilde{P}_{i',i} = 0$. A similar argument shows that this is also the case if i' is uniform and i is not uniform. This completes the proof of (A_δ).

In a completely similar way, we can also prove that

if $\delta \geq 0$ and $(A_{\delta-1})$, $(B_{\delta-1})$ are true then (B_δ) is true.

We can then proceed as in [*loc. cit.*] to complete the proof. □

3.6 Uniform pairs and uniform functions

We claim that (without any assumptions on the center of G or on p, q)

(a) *the pair $i \in I^F$ is uniform (cf. (3.1)) if and only if Y_i is a uniform function, and*

(b) *we have $\lambda_{i,i'} = (Y_i, Y_{i'}) = 0$ if one of $i, i' \in I^F$ is uniform and the other is not.*

Before we prove this let us check that this implies Proposition 1.3.

Let C be an F-stable unipotent class and $i_0 = (C, \bar{\mathbb{Q}}_l) \in I^F$. By Remark 3.2 we know that i_0 is uniform. By (2.1) we have $f_C = |G^F|Y_{i_0}$ (for a suitable normalization) hence (a) implies that this is a uniform function and we are done. Now consider f'_C. We can write $f'_C = \sum_i b_i Y_i$ where the sum is over all $i \in I(C)^F$ and $b_i \in \bar{\mathbb{Q}}_l$. By the orthogonality relations in (2.1) we have

$$a\delta_{i',i_0} = (f'_C, Y_{i'}) = \sum_{i \in I(C)^F} b_i (Y_i, Y_{i'}) = \sum_{i \in I(C)^F} b_i \lambda_{i,i'} \quad \text{for all } i' \in I(C)^F.$$

The matrix $(\lambda_{i,i'})$ (where $i', i \in I(C)^F$) is invertible. Let $(\lambda'_{i,i'})$ denote its inverse. Then the above equations imply that $b_i = a\lambda'_{i_0,i}$. Now (b) shows that $\lambda_{i,i'} = 0$ if one of i', i is uniform and the other is not. The coefficients $\lambda'_{i,i'}$ in the inverse matrix then have the analogous property. Since i_0 is uniform we conclude that $b_i = 0$ unless i is uniform. Hence f'_C is uniform. This completes the proof of Proposition 1.3.

We now prove (a). Recall that a class function on G^F_{uni} is uniform if and only it is a linear combination of the Green functions of G^F. Since the functions $\{Y_i \mid i \in I^F\}$ form a basis of the space of class functions on G^F_{uni} it will therefore be sufficient to show that the Green functions can be expressed as linear combinations of the functions $\{Y_i \mid i \in I_0^F\}$ and vice versa.

Assume at first that G has a connected center. By Proposition 3.5, we can write

$$X_i = \sum_{i' \in I_0^F} P_{i',i} Y_{i'} \quad \text{for all } i \in I_0^F.$$

If we choose a total order on I_0 which refines the order relation $i' \leq i$, we see that the matrix of coefficients $P_{i',i}$ has a triangular shape with 1's along the diagonal. Hence these equations can be inverted, and every Y_i (for $i \in I_0^F$) can be expressed as a linear combination of the functions $X_{i'}$, for various $i' \in I_0^F$.

By [13], (10.4.5), and the character formula in [13], Theorem 8.5, a function X_i for which $i \in I^F$ is uniform can be expressed as a linear combination of generalized Green functions corresponding to various F-stable maximal tori in G. (This is because i is uniform; otherwise, one would have to use generalized Green functions corresponding to Levi subgroups in G which are not maximal tori.) But now [18], Theorem 5.5 (part II), states that these generalized Green functions (corresponding to maximal tori) coincide with the ordinary Green functions of G^F. Moreover, this can be reversed and hence every Green function is a linear combination of the functions $\{X_i \mid i \in I_0^F\}$. Combining this with the above relations among the X_i and Y_i we see that, indeed, the Green functions can be expressed in terms of the functions $\{Y_i \mid i \in I_0^F\}$ and vive versa.

If the center of G is not connected let $\iota\colon G \to G'$ be a regular embedding. Recall from [15] that this means that ι is a homomorphism of connected reductive groups over $\mathbb{F}_q$ such that G' has a connected center, ι is an isomorphism onto a closed subgroup of G', and $\iota(G), G'$ have the same derived subgroup. To simplify notation, we identify G and its image $\iota(G)$.

The embedding $G \subseteq G'$ defines a bijection between the F-stable unipotent classes in G and in G'. Let $u \in C^F$ and consider the canonical quotient $C_{G'}(u) \to A_{G'}(u)$. Since $G' = GZ(G')$ the restriction of this map to $C_G(u)$ defines a surjective map $A_G(u) \to A_{G'}(u)$ whose kernel is given by the image of $Z(G)$ in $A_G(u)$. Via this surjection (which is compatible with the action of F) we also obtain a canonical injective map $I_{G'}(C)^F \to I_G(C)^F$. Since this holds for all F-stable unipotent classes C we obtain an injective map $I_{G'}^F \to I_G^F$. The characterization of uniform pairs in terms of multiplicities in permutation representations as in the proof of Remark 3.2 immediately shows that $i \in I_G^F$ certainly is uniform if i is the image of a uniform pair in $I_{G'}^F$ under this map. On the other hand the number of uniform pairs in I_G^F is always given (via the Springer correspondence) by the number of irreducible characters of the Weyl group W which are invariant under the action of F. Since the latter number is the same for G and G' we conclude that the uniform pairs in I_G^F are precisely the images of the uniform pairs in $I_{G'}^F$.

It follows from the definitions that for all $i \in I_{G'}^F$ we have

$$\mathrm{Res}_G^{G'}(Y_i^{G'}) = Y_i^G \text{ where we also regard } i \text{ as an element in } I_G^F.$$

Now it is also known (see [17]) that the Green functions for G^F are the restrictions of the Green functions for G'^F. Hence we can use the results from the connected center case to conclude that the Green functions of G^F are linear combinations of the functions $\{Y_i^G \mid i \in I_G^F \text{ uniform}\}$ and vice versa. This completes the proof of (a).

Finally, let us consider (b). If the center of G is connected then this is already contained in Proposition 3.5(ii). If the center of G is not connected we use a regular embedding $G \subseteq G'$ as above. Recall that we then have a surjective map $A_G(u) \to A_{G'}(u)$ with kernel given by the image of $Z(G)$ in $A_G(u)$. Using the definitions this easily implies that $(Y_i, Y_{i'}) = 0$ if one of $i, i' \in I_G^F$ lies in the image of the map $I_{G'}^F \to I_G^F$ and the other does not. This implies (b), and the proof is complete.

3.7 Series of irreducible characters

We assume for the rest of this section that the center of G is connected. (We will see in Section 5 that this is no loss of generality as far as Problem 1.1 and Problem 1.2 are concerned.) Let $T \subseteq G$ be an F-stable maximal torus contained in some F-stable Borel subgroup of G, and W be the Weyl group of G with respect to T. Let G^* be a group dual to G (see [11], (8.4)). Then G^* is also defined over $\mathbb{F}_q$ and we denote again by F the corresponding Frobenius map. We can identify W with the Weyl group of an F-stable maximal torus $T' \subseteq G^*$ dual to T; note that the actions of the Frobenius maps of G and G^* on W are inverse to each other.

(a) Let $s \in T'$ be a semisimple element such that the G^*-conjugacy class of s is F-stable. Let W_s be the stabilizer of s in W. Then W_s is a reflection subgroup of W. Let $w_1 \in W$ be the unique element of minimal length in the coset $Z_s = \{w \in W \mid F(s) = w(s)\}$. Then we have an induced automorphism $\gamma\colon W_s \to W_s$ defined by $\gamma^{-1}(w) = F(w_1 w w_1^{-1})$ for all $w \in W_s$ (see [11], (2.15) and the remarks in [11], p.258). Let $\bar{X}(W_s, \gamma)$ be the parameter set defined in [11], (4.21.12); this set only depends on W_s and γ.

(b) If $s \in T'$ is as in (a), we let $\tilde{W}_s = W_s\langle\sigma\rangle$ be the semidirect product of W and the cyclic group $\langle\sigma\rangle$ with generator σ such that $\sigma w \sigma^{-1} = \gamma(w)$ for all $w \in W_s$. Let ψ be an irreducible character of W_s which can be extended to $\tilde{W}_s$; we fix one possible extension of ψ and denote it by $\tilde{\psi}$. As in [11], (3.7), we define

$$R^s[\tilde{\psi}] := \frac{1}{|W_s|} \sum_{w \in W_s} \tilde{\psi}(\sigma w) R^G_{T_{w_1 w}, \theta_s},$$

where $T_{w_1 w} \subseteq G$ is an F-stable maximal torus obtained from T by twisting with $w_1 w$ and θ_s is an irreducible character of $T^F_{w_1 w}$ in 'duality' with s. (This 'duality' is described in [11], proof of Lemma 6.2 and the remarks on p.257.)

(c) The irreducible characters of G^F are divided into series corresponding to conjugacy classes of F-stable semisimple elements in G^*. If $s \in T'$ is as in (a), we denote by $\mathcal{E}_s$ the corresponding series. By [11], Main Theorem 4.23, there exists a bijection

$$\mathcal{E}_s \leftrightarrow \bar{X}(W_s, \gamma), \quad \rho \leftrightarrow \bar{x}_\rho,$$

such that the scalar product

$$(\rho, R^s[\tilde{\psi}])$$

is a rational number depending only on $w_1 W_s$, ψ, and $\bar{x}_\rho \in \bar{X}(W_s, \gamma)$. Let us denote this number by $a(w_1 W_s, \psi, \bar{x}_\rho)$.

(d) Consider the special case where $s = 1$. Then $W_s = W$, $w_1 = 1$ and σ is given by the action of F. We denote by $\mathrm{Irr}(W)^F$ the set of irreducible characters of W which can be extended to $\tilde{W}$, and we assume chosen once and for all a fixed extension for such a character. The corresponding functions $R^s[\tilde{\phi}]$ will be denoted by Q_ϕ, where $\phi \in \mathrm{Irr}(W)^F$. (These are the same as the functions in [17], Remark 5.5(i).)

With this notation we can now state the following result, which expresses our average values as linear combinations of Green functions with coefficients 'independent of q'.

Corollary 3.8 *Assume that the center of G is connected. Let $s \in T'$ be as in (3.7a) and $\rho \in \mathcal{E}_s$. Let C be an F-stable unipotent class in G and $u_1, \ldots, u_r$ be representatives for the G^F-classes contained in C^F. Then there exists constants*

$b(w_1W_s, \phi, \bar{x}_\rho) \in \bar{\mathbb{Q}}_l$ (depending only on $w_1W_s, \phi, \bar{x}_\rho$) such that

$$\begin{aligned}
(\rho, f_C) &= \sum_{j=1}^{r}\sum_{\phi}[G^F : C_G(u_j)^F]\, b(w_1W_s, \phi, \bar{x}_\rho)\, Q_\phi(u_j), \\
(\rho, f'_C) &= \sum_{j=1}^{r}\sum_{\phi}[A(u_j) : A(u_j)^F]\, b(w_1W_s, \phi, \bar{x}_\rho)\, Q_\phi(u_j),
\end{aligned}$$

where in both formulae the second sum is over all $\phi \in \mathrm{Irr}(W)^F$.

Proof. Let ρ_{unif} denote the uniform projection of ρ. By Proposition 1.3 we know that f_C and f'_C are uniform. Hence we can replace ρ by ρ_{unif} in order to evaluate the scalar products with f_C and f'_C.

The various functions $R^s[\tilde{\psi}]$ have norm 1 and are mutually orthogonal. The uniform projection of ρ is given by projecting ρ on the space generated by the various $R^s[\tilde{\psi}]$. Hence we have

$$\rho_{\mathrm{unif}} = \sum_{\psi} a(w_1W_s, \psi, \bar{x}_\rho) R^s[\tilde{\psi}]$$

where the sum is over all irreducible characters ψ of W_s which can be extended to $\tilde{W}_s$. We insert the defining equation for $R^s[\tilde{\psi}]$ and note that the value of a Deligne-Lusztig generalized character at a unipotent element is the value of the corresponding Green function. Now the Green functions for G^F can be re-written in terms of the functions Q_ϕ, where $\phi \in \mathrm{Irr}(W)^F$ and where the coefficients are given by the entries in the inverse of the matrix of values $(\tilde{\phi}(Fw))$. This yields the above expressions for the average values.

Finally note that the coefficients in these linear combinations involve the constants $a(w_1W_s, \psi, \bar{x}_\rho)$, the character values $\tilde{\psi}(\sigma w)$, and the entries in the inverse of the matrix of values $(\tilde{\phi}(Fw))$. Having chosen fixed extensions of the various characters involved we see that these coefficients only depend on w_1W_s, ϕ and $\bar{x}_\rho$. This completes the proof. □

4 CONSIDERING q AS A VARIABLE

We continue to assume that G has a connected center. We have seen in Corollary 3.8 that average values of irreducible characters of G^F as in Problem 1.1 and Problem 1.2 can be expressed in terms of certain combinatorial objects associated with various reflection subgroups of the Weyl group of G and the values of the Green functions of G^F. There is a sense in which the latter are given by 'polynomials in q', and hence the same holds for our average value. In this section we will give a precise formulation for this statement, and this will eventually allow us to remove the assumption on p and q in Corollary 2.6. It will be technically simpler if our group G is simple modulo its center. (In Section 5 below we will see that this is no loss of generality as far as Problem 1.1 and Problem 1.2. are concerned.)

For the remainder of this section, our group G has a connected center and is simple modulo its center. As remarked above we will want to say that certain quantities or objects associated with G^F are given by 'polynomials in q' or are classified 'independently of q'. In order to make this precise, we let Ψ be the root datum of G with respect to a fixed F-stable maximal torus T contained in some F-stable Borel subgroup of G. We denote by W the Weyl group of G with respect to T; this only depends on Ψ. Let $X = X(T)$ be the character group of T. Then F acts as q times an automorphism F_0 of finite order on X, and the pair (G, T) together with the Frobenius map F is determined by (Ψ, F_0) and the choice of the prime power q. We now assume given, once and for all, the root datum Ψ, the corresponding Weyl group W, and the automorphism F_0. Then each choice of a prime power q_1 determines a pair (G_1, T_1) and a Frobenius map F_1 such that G_1 has root datum Ψ and F_1 acts as q_1 times F_0 on the character group of T_1.

4.1 Classification of unipotent classes

We summarize the known results on the classification of unipotent classes in good characteristic, as follows. There exists a finite index set A and a map $A \to \mathbb{N}_0$, $\alpha \mapsto d_\alpha$, depending only on (Ψ, F_0) and having the following properties. If q is a power of a good prime and G is the corresponding group over $\mathbb{F}_q$, there is a map

$$A \to G^F_{\text{uni}}, \quad \alpha \mapsto u_\alpha,$$

such that $\{u_\alpha \mid \alpha \in A\}$ is a set of representatives for the F-stable unipotent classes in G and $d_\alpha = \dim C_\alpha$ where C_α is the class of G containing u_α. (This is contained, for example, in [3], Chapter 5).

Moreover, there is a collection of finite groups $(A_\alpha)_{\alpha\in A}$ such that the map $A \to G^F_{\text{uni}}$ can be chosen to have the following additional properties.

(i) For each α, the group of components of the centralizer of u_α is isomorphic to A_α, and the action of F on this group is trivial.

(ii) For each α, the element u_α is split in the sense of [17], Remark 5.1, except possibly when G is of type E_8, $q \equiv -1 \bmod 3$, and u_α lies in the class $D_8(a_3)$ (notation of the table in [3], pp.405).

Each u_α is uniquely determined up to G^F-conjugacy by (i) and (ii). This follows in all cases where split elements exist, see Shoji [17] and the references there. For type E_8, see Kawanaka [8], (1.2.1); the uniqueness of u_α in this case is mentioned in [2], p.590.

For each $\alpha \in A$ we let $\mathrm{Cl}(A_\alpha)$ be a set of representatives of the conjugacy classes of A_α. By property (i), the set $\mathrm{Cl}(A_\alpha)$ parametrizes the various G^F-classes contained in C^F_α (for q and G as above). If $j \in \mathrm{Cl}(A_\alpha)$ we denote by $u_{\alpha,j}$ an element in C^F_α which is obtained from the representative u_α by twisting with j.

4.2 Values of Green functions

We summarize the known results about the values of Green functions in good characteristic as follows. For $\delta = 0, \pm 1$ there exist maps

$$Q^\delta \colon \mathrm{Irr}(W)^F \times \coprod_{\alpha\in A} A_\alpha \to \mathbb{Z}[t] \quad \text{and} \quad h^\delta \colon \coprod_{\alpha\in A} \mathrm{Cl}(A_\alpha) \to \mathbb{Q}[t]$$

depending only on (Ψ, F_0) and having the following properties. If q is a power of a good prime such that $q \equiv \delta \bmod 3$ and G is the corresponding group over $\mathbb{F}_q$ then

$$Q_\phi(u_{\alpha,j}) = Q^\delta(\phi,\alpha,j)(q) \quad \text{for all } w \in W,\ \alpha \in A \text{ and } j \in \mathrm{Cl}(A_\alpha),$$

where $\phi \in \mathrm{Irr}(W)^F$ and $u_{\alpha,j}$ is an element in C_α^F obtained by twisting the representative u_α with j. Moreover, $h^\delta(\alpha,j)(q)$ is the size of the G^F-conjugacy class of $u_{\alpha,j}$.

The results concerning the Green functions are contained in [17]. The existence of the polynomials $h^\delta(\alpha,j)$ follows, for example, from the algorithm for the computation of generalized Green functions in [13], Theorem 24.4. These polynomials (for fixed α) all have the same degree which is the integer $d_\alpha = \dim C_\alpha$. Note that the parameter δ makes a difference only for G of type E_8.

4.3 The average value polynomials

Fix $\delta = 0, \pm 1$. Let q be any power of a good prime with $q \equiv \delta \bmod 3$ and G the corresponding group over $\mathbb{F}_q$ with dual group G^*. Let $s \in T'$, $W_s \subseteq W$ and $w_1 \in W$ as in (3.7a). Then w_1 has minimal length in the coset w_1W_s and we have $F(w_1W_sw_1^{-1}) = W_s$. The cosets w_1W_s arising in this way (for various choices of q and elements $s \in T'$) will be called the *δ-admissible cosets* of W.

Let w_1W' be a δ-admissible coset. We define the automorphism $\gamma : W' \to W'$ and the corresponding semidirect product $\tilde{W}'$ analogously as in (3.7a). The constructions in [11], Chapter 4, yield a parameter set $\bar{X}(W',\gamma)$ and rational numbers $a(w_1W',\phi,\bar{x})$ (as in (3.7b)) for all irreducible characters ϕ of W' which can be extended to $\tilde{W}'$. Moreover, we obtain constants $b(w_1W_s,\phi,\bar{x})$ (for $\phi \in \mathrm{Irr}(W)^F$) by the rewriting process as in the proof of Corollary 3.8. We now define two polynomial functions $A \times \bar{X}(W',\gamma) \to \mathbb{Q}[t]$ by

$$\begin{aligned}
\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x}) &:= \sum_{j\in \mathrm{Cl}(A_\alpha)} \sum_{\phi \in \mathrm{Irr}(W)^F} h^\delta(\alpha,j)\, b(w_1W',\phi,\bar{x})\, Q^\delta(\phi,\alpha,j),\\
\mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x}) &:= \sum_{j\in \mathrm{Cl}(A_\alpha)} \sum_{\phi \in \mathrm{Irr}(W)^F} [A_\alpha : C_{A_\alpha}(j)]\, b(w_1W',\phi,\bar{x})\, Q^\delta(\phi,\alpha,j).
\end{aligned}$$

Given α and $\bar{x}$ we call the corresponding polynomials the *average value polynomials* of type (1.1) and (1.2), respectively.

The relevance of this definition is as follows. Let q be a power of a good prime with $q \equiv \delta \bmod 3$, and G the corresponding group over $\mathbb{F}_q$. Let $s \in T'$ and W_s, w_1, γ be as in (3.7a). Then w_1W_s is a δ-admissible coset, hence $\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x})$ and $\mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x})$ are defined for all $\alpha \in A$ and $\bar{x} \in \bar{X}(W_s,\gamma)$. Corollary 3.8 can now be rephrased by saying that if $\rho \in \mathcal{E}_s$ we have

$$(\rho, f_C) = \mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x}_\rho)(q) \quad \text{and} \quad (\rho, f'_C) = \mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x}_\rho)(q).$$

Proposition 4.4 *Let w_1W' be a δ-admissible coset and $\bar{x}$ a fixed element in the corresponding parameter set $\bar{X}(W',\gamma)$.*

(i) There exists a unique $\alpha \in A$ with maximal possible value d_α such that the polynomial $\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x})$ is non-zero.

(ii) *There exists a unique $\tilde{\alpha} \in A$ with maximal possible value $d_{\tilde{\alpha}}$ such that the polynomial* $\mathrm{AV}^{\delta}_{(1.2)}(\tilde{\alpha}, \bar{x})$ *is non-zero.*

(iii) *We have $\alpha = \tilde{\alpha}$.*

(iv) $\mathrm{AV}^{\delta}_{(1.1)}(\alpha, \bar{x})(q) \neq 0$ *and* $\mathrm{AV}^{\delta}_{(1.2)}(\alpha, \bar{x})(q) \neq 0$ *for all good prime powers q such that $q \equiv \delta$ mod 3.*

(Recall from (4.1) that $d_\alpha = \dim C_\alpha$.)

Proof. Let M be the set of integers q which are powers of various good primes and such that the following conditions are satisfied.

(a) All elements in M are congruent to δ modulo 3.

(b) If an average value polynomial is non-zero then it is non-zero when evaluated at every $q \in M$.

(c) If $q \in M$ and G is the corresponding group over $\mathbb{F}_q$ with Frobenius map F the results in Section 2 are applicable.

(d) If $q \in M$ and G is the corresponding group over $\mathbb{F}_q$ with Frobenius map F then the coset $w_1 W'$ arises from an F-stable semisimple class in G as in (3.7a).

The set M contains infinitely many elements. Indeed, condition (b) holds for all but finitely many good prime powers since we only have a finite number of average value polynomials; condition (c) holds for all large enough powers of large enough primes. Using Dirichlet's Theorem on primes in an arithmetic progression, the set M_1 of good prime powers satisfying (a), (b), (c) is infinite. Finally, Deriziotis has shown in [4], Theorem 3.3, that condition (d) either holds for none of for all but finitely many good prime powers q in a fixed congruence class modulo a certain integer depending only on (Ψ, F_0). The definition of δ-admissibility therefore implies that the set of elements in M_1 which also satisfy (d) is still infinite.

Let us prove (i). Let $q \in M$ and G the corresponding group over $\mathbb{F}_q$. By condition (d), the coset $w_1 W'$ arises from some F-stable semisimple class (s) in G^* as in (3.7a). Let ρ be an irreducible character of G^F in the corresponding series $\mathcal{E}_s$ such that $\bar{x}_\rho$ is the given element $\bar{x} \in \bar{X}(W', \gamma)$. By (c) we can apply Corollary 2.6 and conclude that there exists a unique $\alpha \in A$ with maximal possible value d_α such that

$$(\rho, f_C) = \mathrm{AV}^{\delta}_{(1.1)}(\alpha, \bar{x})(q) \neq 0.$$

But then property (b) implies that (i) holds. The proof of (ii) is completely analogous, and yields the same class C_α by Corollary 2.6. This also proves (iii).

Now we prove (iv). For this purpose note that the class C_α has the properties (1) and (2) in the proof of Corollary 2.6. By [16], Theorem 11.2, we also have the following additional property.

(1') For all $u \in C_\alpha^F$, the absolute value of $(\Gamma_u, D_G(\rho))$ is $\leq |A_\alpha||W|$.

Recall that property (2) implies that assumption (*) in Proposition 2.5 is satisfied for the character ρ and the class C_α, and this holds for each choice of $q \in M$. The formulae in Proposition 2.5 yield that

$$\begin{aligned}\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x})(q) &= q^d \sum_{j\in \mathrm{Cl}(A_\alpha)} h^\delta(\alpha,j)(q)N_j(q) \quad \text{and}\\ \mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x})(q) &= q^d \sum_{j\in \mathrm{Cl}(A_\alpha)} [A_\alpha : C_{A_\alpha}(j)]N_j(q),\end{aligned}$$

where $N_j(q)$ denotes the multiplicity of $D_G(\rho)$ in the GGGR associated with the representative in C_α corresponding to j. Property (1') gives a bound on the absolute value of $N_j(q)$ 'independently of q'. So there exists an infinite subset $M' \subseteq M$ such that $(N_j(q))_{j\in \mathrm{Cl}(A_\alpha)}$ is constant for all $q \in M'$. Let N_j denote this constant for $j \in \mathrm{Cl}(A_\alpha)$. We conclude that

$$\begin{aligned}\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x})(q) &= q^d \sum_{j\in \mathrm{Cl}(A_\alpha)} h^\delta(\alpha,j)(q)N_j \quad \text{for all } q \in M',\\ \mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x})(q) &= q^d \sum_{j\in \mathrm{Cl}(A_\alpha)} [A_\alpha : C_{A_\alpha}(j)]N_j \quad \text{for all } q \in M'.\end{aligned}$$

So we actually obtain identities of polynomials in $\mathbb{Q}[t]$:

$$\begin{aligned}\mathrm{AV}^\delta_{(1.1)}(\alpha,\bar{x}) &= t^d \sum_{j\in \mathrm{Cl}(A_\alpha)} h^\delta(\alpha,j)N_j,\\ \mathrm{AV}^\delta_{(1.2)}(\alpha,\bar{x}) &= t^d \sum_{j\in \mathrm{Cl}(A_\alpha)} [A_\alpha : C_{A_\alpha}(j)]\, N_j.\end{aligned}$$

Since $\pm D_G(\bar{\rho})$ is an irreducible character, either all numbers N_j are non-negative or all numbers $-N_j$ are non-negative, and by (i) at least one N_j must be non-zero. So the expression

$$t^d \sum_{j\in \mathrm{Cl}(A_\alpha)} [A_\alpha : C_{A_\alpha}(j)]\, N_j$$

is a non-zero constant times t^d. Hence, in particular, its value at any q as in (iv) is non-zero. A slight modification of this argument also works for the other expression. Indeed, since $h^\delta(\alpha,j)$ gives a strictly positive integer when evaluated at any good prime power (namely the size of a conjugacy class), we conclude that the expression

$$t^d \sum_{j\in \mathrm{Cl}(A_\alpha)} h^\delta(\alpha,j)N_j$$

is a polynomial with the property that if we evaluate it at any good prime power then we obtain a strictly positive or a strictly negative number as a result. Again we are done. □

5 PROOF OF THEOREM 1.4

Let q be a power of a prime p and G be a connected reductive group defined over $\mathbb{F}_q$, with corresponding Frobenius map F. For the moment we make no assumption on p or on the center of G.

Lemma 5.1 *Let $G \subseteq G'$ be a regular embedding of G into a connected reductive group G' over $\mathbb{F}_q$ with a connected center and such that G, G' have the same derived subgroup. Then Problem 1.1 (respectively, Problem 1.2) has a positive solution for G if and only if it has a positive solution for G'.*

Proof. At first note that the embedding $G \subseteq G'$ defines a bijection between the F-stable unipotent classes of G and those of G'. Let C be any F-stable unipotent class of G, let ρ' be an irreducible character of ${G'}^F$, and let ρ be an irreducible component of the restriction of ρ' to G^F. By Clifford's Theorem the restriction of ρ' to G^F is a sum of irreducible characters of G^F which are of the form $\rho^x := \rho \circ c_x$, where c_x denotes the automorphism of G^F induced by conjugation with an element $x \in {G'}^F$.

It is clear that the function f_C defined with respect to G is $[{G'}^F : G^F]$ times the restriction of the corresponding function defined with respect to G'. Hence f_C is invariant under ${G'}^F$ and we have $f_C^x = f_C$ for all $x \in {G'}^F$.

Using the methods in (3.6) it can be easily seen that a similar statement also holds for the function f'_C on G^F. Hence it is also invariant under ${G'}^F$ and we have $(f'_C)^x = f'_C$ for all $x \in {G'}^F$.

We conclude that the scalar product of ρ^x with f_C (respectively, with f'_C) is the same as the scalar product of ρ with f_C (respectively, with f'_C). Using Clifford's Theorem in the above form, we see that the scalar product of ρ' with f_C (respectively, with f'_C) is a non-zero multiple of the scalar product of ρ with f_C (respectively, with f'_C). This implies the desired equivalence. $\square$

So from now on, we can assume that the center of G is connected. The next result shows that we can reduce to the case where G is simple modulo its center.

Lemma 5.2 *Let p be a fixed prime. Assume that Problem 1.1 (respectively, Problem 1.2) has a positive solution for all groups G which are defined over a finite field of characteristic p, which have a connected center and which are simple modulo their center. Then Problem 1.1 (respectively, Problem 1.2) has a positive solution for all groups defined over a finite field of characteristic p.*

Proof. Let G be any group defined over $\mathbb{F}_q$, where q is a power of p. By Lemma 5.1 we may assume that the center of G is connected. The following reasoning is almost entirely analogous to that in [11], (8.8).

We can find a surjective homomorphism $f\colon G' \to G$ of algebraic groups over $\mathbb{F}_q$ such that the center of G' is connected, the kernel of f is a central torus, and the derived subgroup of G' is semisimple and simply-connected. We claim that if Problem 1.1 (respectively, Problem 1.2) has a positive solution for G' then it also has a positive solution for G. Indeed, the map f induces a bijection between the unipotent classes of G' and G. Since the kernel of f is connected this bijection also works on the level of the finite groups, and we have $f({G'}^F) = G^F$. So, if ρ is an irreducible character of G^F then $\rho' := \rho \circ f$ is an irreducible character of ${G'}^F$. Furthermore, the function f_C (respectively, f'_C) lifts to the analogously defined function of ${G'}^F$. This implies the claim.

Hence we may now also assume that the derived group G_{der} of G is simply-connected.

Let us write $G_{\mathrm{der}} = R_{f_1}(G_1) \times \ldots \times R_{f_n}(G_n)$ where each G_i is a closed simple simply-connected subgroup and R_f denotes restriction of scalars from $\mathbb{F}_{q^f}$ to $\mathbb{F}_q$ (for some $f \geq 1$). We can embed each G_i regularly (over $\mathbb{F}_{q^{f_i}}$) into a connected reductive group G_i'' with a connected center and which is simple modulo its center. Let $G'' := R_{f_1}(G_1'') \times \ldots \times R_{f_n}(G_n'')$. Then we also have a regular embedding $G_{\mathrm{der}} \to G''$ (over $\mathbb{F}_q$).

Finally, as in [*loc. cit.*], there exists a connected reductive group G''' with connected center and defined over $\mathbb{F}_q$ and there exist regular embeddings $G \to G'''$, $G'' \to G'''$ (over $\mathbb{F}_q$) which are compatible with the regular embedding $G_{\mathrm{der}} \to G''$.

Now we can argue as follows. Using Lemma 5.1 twice we see that Problem 1.1 (respectively, Problem 1.2) has a positive solution for G if and only if this is the case for G''' if and only if this is the case for G''. Now G'' has a decomposition into a direct product of various factors of the form $R_{f_i}(G_i)$, and this leads to a similar decomposition on the level of the finite groups. Correspondingly, the irreducible characters of ${G''}^F$ are exterior tensor products of irreducible characters for the various factors, and it follows easily that Problem 1.1 (respectively, Problem 1.2) has a positive solution for G'' if this is the case for each factor G_i. Hence we are reduced to groups which have a connected center and are simple modulo their center. This completes the proof. $\square$

We are now ready for the proof of Theorem 1.4.

5.3 EXISTENCE OF UNIPOTENT SUPPORTS

Let G be as in the first sentence of this section, and assume that p is good.

Let us first show that Problem 1.1 has a positive solution (that is, the unipotent support of an irreducible character exists). By Lemmas 5.1 and 5.2 we may assume that G has a connected center and is simple modulo its center. Then we can apply the formalism of Section 4. Let ρ be an irreducible character of G^F contained in the series $\mathcal{E}_s$, say. Let $w_1 W_s$ and $\bar{X}(W_s, \gamma)$ as in (3.7a). By Proposition 4.4(i), there exists a unique $\alpha \in A$ with maximal possible value for d_α such that the average value polynomial $\mathrm{AV}^\delta_{(1.1)}(\alpha, \bar{x}_\rho)$ is non-zero (where $q \equiv \delta \bmod 3$). By Proposition 4.4(iv), we also have

$$\sum_{g \in C_\alpha^F} \rho(g) = \mathrm{AV}^\delta_{(1.1)}(\alpha, \bar{x}_\rho)(q) \neq 0.$$

Now let $\beta \in A$ be any element such that the average value of ρ on C_β^F is non-zero. Then, clearly, the corresponding average value polynomial itself is non-zero hence Proposition 4.4(i) implies that $\dim C_\beta = d_\beta \leq d_\alpha = \dim C_\alpha$ with equality only for $\alpha = \beta$. Hence the class C_α is the unipotent support of ρ.

A completely analogous argument shows that also Problem 1.2 has a positive solution, and Proposition 4.4(iii) proves that we obtain the same class as before. This proves part (a) in Theorem 1.4.

5.4 THE p-PARTS OF CHARACTER DEGREES

Let again G be as in the first sentence of this section, with p good. Now we turn to Theorem 1.4(b), that is, the problem concerning the p-part in the degree of an irreducible character ρ of G^F. We know already by (5.3) that ρ has a unipotent support, C say. Now we must show that the p-part in the degree of ρ is q^d where d is the dimension of the variety of Borel subgroups containing a fixed element in C. By the dimension formula in [3], Theorem 5.10.1, we have $d = (2N - \dim C)/2$ where N is the number of positive roots in the root system of G.

We can use a similar reasoning as before to reduce to the case where G has a connected center and is simple modulo its center. Indeed, let $G \subseteq G'$ be a regular embedding and ρ' an irreducible character of ${G'}^F$ whose restriction to G^F contains ρ as a constituent. Then ρ' also has unipotent support C (see Lemma 5.1). Since the index of G^F in ${G'}^F$ is certainly prime to p, Clifford's Theorem implies that the degree of ρ' is a multiple (coprime to p) of the degree of ρ. So the characters ρ and ρ' have the same p-part in their degrees and the dimensions of the unipotent supports are equal. Hence it is sufficient to consider groups G with a connected center. It is then also straightforward to check that the constructions in the proof of Lemma 5.2 behave well with respect to p-parts in character degrees and dimensions of unipotent supports. (This is certainly the case for the first reduction to groups G with a connected center and such that the derived group G_{der} is simply-connected; note that the remaining constructions just involve taking regular embeddings and direct products.)

Let us now assume that G has a connected center and is simple modulo its center. We use again the formalism of Section 4. Let $q \equiv \delta \bmod 3$ and ρ be an irreducible character of G^F contained in the series $\mathcal{E}_s$, say. Let $\bar{X}(W_s, \gamma)$ be the associated parameter set as in (3.7a), and $\bar{x} = \bar{x}_\rho$ be the element in this set corresponding to ρ. Let A be as in (4.1) and $\alpha_0 \in A$ be the unique element such that $d_{\alpha_0} = 0$ (so that C_{α_0} is the class of the trivial element in G). We define

$$\deg(\rho) := \mathrm{AV}^{\delta}_{(1.1)}(\alpha_0, \bar{x}_\rho) \in \mathbb{Q}[t].$$

Then, by the formula (4.4), the value of $\deg(\rho)$ at q is the degree of ρ. Let $a = a(\bar{x}) \geq 0$ such that t^a is the maximal power of t dividing $\deg(\rho)$. We claim that q^a is the p-part in the degree of ρ. Indeed, from the explicit description of the Fourier coefficients in [13], Chapter 4, and the formulae [13], (4.26.1) and (4.26.3), we deduce that there exists a positive integer d which is divisible by bad primes only and a monic polynomial $f \in \mathbb{Z}[t]$ such that $\deg(\rho) = (1/d)t^a f$ and $f \equiv \pm 1 \bmod t$. This implies our claim since p is good.

Thus, we have described the p-part in the degree of ρ purely in terms of our average value polynomials. On the other hand, we know by Proposition 4.4 and the argument in (5.3) that the unipotent support of ρ is also characterized purely in terms of the average value polynomials corresponding to our fixed $\bar{x} = \bar{x}_\rho$. Therefore, it will be sufficient to prove the following statement.

Given a δ-admissible coset $w_1 W'$ and $\bar{x} \in \bar{X}(W', \gamma)$ let $\alpha \in A$ be as in Proposition 4.4(i). Show that $a(\bar{x}) = (2N - d_\alpha)/2$.

Since this statement only concerns properties of the average value polynomials and dimensions of unipotent classes we can assume, without loss of generality, that q and p are large enough so that the results in [16] are applicable. Then the unipotent support $C = C_\alpha$ of our character ρ can also be characterized in terms of the map

$$\xi \colon \{\text{irreducible characters of } G^F\} \to \{F\text{-stable unipotent classes in } G\}$$

defined in [11], (13.4), or [16], (11.1). Indeed, by [16], Theorem 11.2, we have

$$C_\alpha = \xi(\rho') \quad \text{where} \quad \rho' = \pm D_G(\rho).$$

The above statement is an immediate consequence of the properties of the map ξ, as we will now check.

The first step in defining ξ is to associate with ρ' a so-called special conjugacy class in G^* (see [11], (13.2), for the precise definition). This is done as follows. With the character ρ' there is associated a family $\mathcal{F}$ of representations of W_s, and we let E_1 be the unique special representation in the family $\mathrm{sign} \otimes \mathcal{F}$ (cf. [11], (13.1.3)). By the Springer correspondence, we can associate with E_1 the class of a unipotent element $v \in C_{G^*}(s)$. Then the G^*-conjugacy class C' of the element sv is the desired special class in G^*. Next, Lusztig [11], (13.3), defines a map Φ from special classes in G^* to unipotent classes in G, and we have $\xi(\rho') = \Phi(C')$. The main property of the map Φ that we need is that it preserves the dimensions of classes. So we can conclude that

$$d_\alpha = \dim C_\alpha = \dim \xi(\rho') = \dim \Phi(C') = \dim C',$$

and it remains to check that $a(\bar{x}) = (2N - \dim C')/2$. Translating this back using the dimension formula in [3], Theorem 5.10.1, we see that we must show that $a(\bar{x})$ equals the dimension of the variety of Borel subgroups of $C_{G^*}(s)$ containing the unipotent element v. By [11], (13.1.1), the latter dimension is equal to the integer a_{E_1} associated with the special representation E_1 as in [11], (4.1). So, eventually, we see that we must show that

$$a_{E_1} = a(\bar{x}).$$

Now since $\rho' = \pm D_G(\rho)$ and $\mathcal{F}$ is the family associated with ρ', the results in [11], (8.6), imply that the family associated with ρ is $\mathrm{sign} \otimes \mathcal{F}$. But then the formula in [11], (4.26.3), just says that $a(\bar{x}) = a_{E_1}$, and we are done.

REFERENCES

[1] A.M. AUBERT, Some properties of character sheaves, preprint, Ecole Normale Supérieure, Paris, 1996.

[2] W.M. BEYNON AND N. SPALTENSTEIN, Green functions of finite Chevalley groups of type E_n $(n = 6, 7, 8)$, J. Algebra 88 (1984), 584–614.

[3] R.W. CARTER, Finite groups of Lie type: Conjugacy classes and complex characters. Wiley, 1985.

[4] D.I. DERIZIOTIS, On the number of conjugacy classes in finite groups of Lie type, Comm. in Algebra 13 (1985), 1019–1045.

[5] H. ENOMOTO, The characters of the finite Chevalley group $G_2(q), q = 3^f$, Japan. J. Math. 2 (1976), 191–248.

[6] M. GECK, Basic sets of Brauer characters of finite groups of Lie type III, Manuscripta Math. 85, (1994), 195–216.

[7] M. GECK AND G. MALLE, On the existence of a unipotent support for the irreducible characters of finite groups of Lie type, preprint (1996), submitted to Trans. Amer. Math. Soc.

[8] N. KAWANAKA, Generalized Gelfand-Graev representations of exceptional algebraic groups I, Invent. Math. 84 (1986), 575-616.

[9] G. LUSZTIG, Representations of finite Chevalley groups, C.B.M.S. Regional Conference Series in Math. 39, Amer. Math. Soc. (1978).

[10] G. LUSZTIG, Some problems in the representation theory of finite Chevalley groups, Proc. Symp. Pure Mat. 37 (1980), 313–317, Amer. Math. Soc.

[11] G. LUSZTIG, Characters of reductive groups over a finite field, Ann. Math. Studies 107, Princeton U. Press, 1984.

[12] G. LUSZTIG, Intersection cohomology complexes on a reductive group, Invent. Math. 75 (1984), 205–272.

[13] G. LUSZTIG, Character sheaves, Adv. Math. 56 (1985), 193–237; II, 57 (1985), 226–265; III, 57 (1985), 266–315; IV, 59 (1986), 1–63; V, 61 (1986), 103–155.

[14] G. LUSZTIG, On the character values of finite Chevalley groups at unipotents elements, Journal of Algebra, 104 (1986), 146–194.

[15] G. LUSZTIG, On the representations of reductive groups with disconnected centre, *in:* Orbites Unipotentes et Représentations, I. Groupes finis et Algèbres de Hecke, Astérisque 168 (1988).

[16] G. LUSZTIG, A unipotent support for irreducible representations, Adv. Math. 94 (1992), 139–179.

[17] T. SHOJI, Green functions of reductive groups over a finite field, Proc. Symp. Pure Math. 47 (1987), 289–302, Amer. Math. Soc.

[18] T. SHOJI, Character sheaves and almost characters of reductive groups, Advances in Math. 111 (1995), 244–313; II, 314–354.

[19] N. SPALTENSTEIN, On the generalized Springer correspondence for exceptional groups, Advanced Studies in Pure Math. 6 (1985), 317–338.

Meinolf Geck
Équipe des Groupes Finis
UFR de Mathématiques
Université Paris 7
2 Place Jussieu
F–75251 Paris

Chow Groups with Coefficients

Markus Rost

Received: October 18, 1996

Communicated by Alexander S. Merkurjev

Abstract. We develop a generalization of the classical Chow groups in order to have available some standard properties for homology theories: long exact sequences, spectral sequences for fibrations, homotopy invariance and intersections. The basis for our constructions is Milnor's K-theory.

1991 Mathematics Subject Classification: Primary 14C17.

Contents

Summary

The paper considers generalities for localization complexes for varieties. Examples of these complexes are given by the Gersten resolutions in various contexts, in particular in K-theory and in étale cohomology. The paper gives a general notion of coefficient systems for such complexes, the so called cycle modules. There are the corresponding "complexes of cycles with coefficients" and their homology groups, the "Chow groups with coefficients". For these some general constructions are developed: proper push-forward, flat pull-back, spectral sequences for fibrations, homotopy invariance and intersection theory.

If one specializes the material to the case of Milnor's K-theory as coefficient system, one obtains in particular an elementary development of intersections for the classical Chow groups. This treatment is somewhat different to former approaches. The main tool is still the deformation to the normal cone. The major difference is that homotopy invariance is not established alone for the Chow groups, but for the "cycle complex with coefficients in Milnor's K-Theory". This enables one to keep control in fibered situations. The proof of associativity of intersections is based on a doubled version of the deformation to the normal cone.

Conventions and Notations

We work over a ground field k and a base scheme $B \to \operatorname{Spec} k$. The word scheme means a localization of a separated scheme of finite type over k. (This includes schemes of finite type over a field finitely generated over k.) From Section 8 on all schemes are of finite type over a field. Moreover all schemes and morphisms are defined over B (with exceptions in Section 14). The letter M stands from Section 3 on for a cycle module. If not mentioned otherwise, it is defined over B (in Sections 3–5) or over X (in Sections 7–13).

For $x \in X$ we denote by $\dim(x, X)$ the dimension of the closure $\overline{\{x\}}$ of x in X and by $\operatorname{codim}(x, X)$ the dimension of the localization $X_{(x)}$. The set of points of X of dimension (resp. codimension) p is denoted by $X_{(p)}$ (resp. $X^{(p)}$). We make free use of some basic facts from commutative algebra and refer for this to (Hartshorne 1977; Matsumura 1980) and, in particular, to (Fulton 1984, App. A, App. B).

In Sections 6, 8, 9, and 11–13 we use the special notation $X \looparrowright Y$ for certain maps between the cycle complexes. This is explained in (3.8).

Introduction

The classical Chow groups $\mathrm{CH}_p(X)$ of p-dimensional cycles on a variety X may be defined as the cokernel of the divisor map

$$\coprod_{x\in X_{(p+1)}} \kappa(x)^* \xrightarrow{d} \coprod_{x\in X_{(p)}} \mathbb{Z}.$$

Here $X_{(p)}$ is the set of points of X of dimension p and $\kappa(x)$ is the residue class field of x. This paper studies complexes $C_*(X;M)$ of the following type:

$$\cdots \xrightarrow{d} \coprod_{x\in X_{(p+1)}} M\big(\kappa(x)\big) \xrightarrow{d} \coprod_{x\in X_{(p)}} M\big(\kappa(x)\big) \xrightarrow{d} \coprod_{x\in X_{(p-1)}} M\big(\kappa(x)\big) \xrightarrow{d} \cdots .$$

Here M is what we call a cycle module. This is a functor $F \to M(F)$ on fields to abelian groups equipped with four structural data (the even ones: restriction and corestriction; the odd ones: multiplication with K_1 and residue maps for discrete valuations). Moreover there is imposed a list of certain rules and axioms. A particular example of a cycle module is $M = K_*$, given by Milnor's (or Quillen's) K-ring

$$K_*F = \mathbb{Z} \oplus F^* \oplus K_2F \oplus \cdots .$$

Other examples are provided by Galois cohomology, specifically

$$M(F) = \coprod_{n\geq 0} H^n(F; D\otimes \mu_r^{\otimes n})$$

with D a Galois module over a ground field k with $\operatorname{char} k$ prime to r.

The complex $C_*(X;M)$ is called the chain complex of cycles on X and its homology groups $A_p(X;M)$ are called the Chow groups of X (with coefficients in M).

The Chow groups $A_p(X;M)$ enclose various familiar objects. The classical Chow group $\mathrm{CH}_p(X)$ is a direct summand of $A_p(X;K_*)$. The E^2-terms of the local-global spectral sequences in étale cohomology and in Quillen's K-theory are of type $A_p(X;M)$. For proper smooth X of dimension d the group $A_d(X;M)$ is a birational invariant—the "M-valued" analogue of unramified Galois cohomology.

The paper develops some basic constructions for the cycle complexes $C_*(X;M)$ and the Chow groups $A_*(X;M)$ for schemes X of finite type over a field. There are proper push-forward, flat pull-back and homotopy invariance. Moreover intersection theory is available: for regular imbeddings and morphisms to smooth varieties there is a pull-back map. Finally for a morphism $\pi\colon X \to Z$ there is a spectral sequence

$$E^2_{p,q} = A_p(Z; A_q[\pi;M]) \Longrightarrow A_{p+q}(X;M).$$

Here the $A_q[\pi;M]$ are certain cycle modules obtained from taking homology in the fibers. All the mentioned functorial behavior extends for appropriate fiber diagrams to the cycle modules $A_q[\pi;M]$ and the spectral sequences.

The constructions are carried out on complex level in a pointwise manner. The treatment has some parallels to a standard development of homology of CW-complexes. This analogy should not be taken too serious, but may give a first impression about the sort of technicalities. In this picture our "cells" are just all points of the variety in question. The patching data for the "cells" are given by the (geometric) valuations on the residue class field of one point having center in another point. The appropriate local coefficient systems are the cycle modules. However, the nature of these coefficient systems is more complicated than in topology. First of all, their ground ring is provided by Milnor's K-theory of fields. Moreover, besides the usual functorial behavior, there is need for transfer maps (basically because one has to deal with non algebraically closed fields) and there are residue maps for valuations (to give passage from one point of a variety to its specializations).

The material of this paper grew out from considerations concerning the bijectivity of the norm residue homomorphism and Hilbert's Satz 90 for Milnor's K_n. There the computation of the Chow groups of certain norm varieties and quadrics plays an important role. As a general technique (see also Karpenko and Merkurjev 1991) we used a spectral sequence for morphisms $\pi\colon X \to Z$ relating the Chow groups of the total space to something like "the Chow groups of the base with coefficients in the Chow groups of the fibers"; moreover these spectral sequences should be compatible with intersection operations. The goal of the paper was to present an appropriate framework in a fairly direct manner.

With the remarks following, we have tried to draw the line of development of the paper. In the discussion of intersection theory, we restrict for simplicity to typical situations and with Milnor's K-theory as coefficient system, although the actual treatment is more general.

Even if one is interested in classical Chow groups alone, one is led to consider some more general versions of Chow groups. To start with a simple situation, let $Y \subset X$ be a closed subvariety. Then there is an exact sequence

$$\mathrm{CH}_p(Y) \to \mathrm{CH}_p(X) \to \mathrm{CH}_p(X \setminus Y) \to 0.$$

For concrete computations as well as for general considerations, there appears the problem to extend this sequence to the left in a reasonable way by a sort of higher variants of Chow groups. Similarly, let $\pi\colon X \to Z$ be a morphism of varieties and try to relate the Chow groups of X with the Chow groups of Z and of the fibers. When working within the classical Chow groups alone, there will be no good answer in general.

In this paper the approach to these problems is provided by Milnor's K-theory. For a variety X one forms for $n \in \mathbb{Z}$ the complex[◇] $C_*(X;n)$ with

$$C_p(X;n) = \coprod_{x \in X_{(p)}} K_{n+p}\kappa(x)$$

[◇] The complex of cycles with coefficients in Milnor's K-theory to be considered later splits up as a direct sum $C_*(X;K_*) = \coprod_n C_*(X;n)$ according to the grading of Milnor's K-ring.

where $K_n F$ is Milnor's n-th K-group of a field F. The homology groups of the complex $C_*(X;n)$ are denoted by $A_p(X;n)$. For $n=-p\le 0$ it ends up with

$$\cdots \xrightarrow{d} \coprod_{x\in X_{(p+2)}} K_2\kappa(x) \xrightarrow{d} \coprod_{x\in X_{(p+1)}} K_1\kappa(x) \xrightarrow{d} \coprod_{x\in X_{(p)}} K_0\kappa(x) \longrightarrow 0$$

and one has $\mathrm{CH}_p(X) = A_p(X;-p)$.

Then for a subvariety $Y\subset X$ there is a long exact sequence

$$\cdots \to A_{p+1}(X\setminus Y;n) \to A_p(Y;n) \to A_p(X;n) \to A_p(X\setminus Y;n) \to \cdots .$$

Moreover let $\pi\colon X\to Z$ be a morphism. The filtration of the set $X_{(p)}$ given by the dimension of the image gives rise to a filtration of the complex $C_*(X;n)$. The corresponding E^1-spectral sequence looks like

$$E^1_{p,q} = \coprod_{z\in Z_{(p)}} A_q(X_z;n+p) \Longrightarrow A_{p+q}(X;n) \tag{1}$$

with $X_z = X\times_Z \operatorname{Spec}\kappa(z)$.

A major problem in intersection theory is to produce for a regular imbedding $f\colon X'\to X$ a pull-back map $f^\bullet$ on the Chow groups having the geometric meaning of intersecting cycles on X with X'. (For a general account on intersections we refer to Fulton 1984)

These maps are in the actual context of type

$$f^\bullet\colon A_p(X;n) \to A_{p-d}(X';n+d)$$

with $d=\operatorname{codim}(f)$.

In the paper the maps $f^\bullet$ are defined by first constructing homomorphisms of complexes

$$I(f)\colon C_*(X;n)\to C_{*-d}(X';n+d)$$

and then passing to homology. In a fibered situation (that is f lies over some map $Z'\to Z$ with appropriate smoothness conditions), the maps $I(f)$ can be chosen to respect the filtrations, thereby inducing homomorphisms on the corresponding spectral sequences.

As the reader might guess, the maps $I(f)$ cannot be defined canonically in terms of f. Namely, $I(f)$ gives in particular a lift of the classical pull-back map $f^\bullet\colon \mathrm{CH}_p(X)\to \mathrm{CH}_{p-d}(X')$ to the cycle groups. But if a cycle W on X does not meet X' properly, there is in general no way to define $W\cap X'$ by a canonical cycle.

It may be surprising that one can handle with such pull-back maps $I(f)$ on complex level in a reasonable way. Therefore we will discuss here the nature of these maps in some detail.

When working with the complexes $C_*(X;n)$, it turns out that the necessary constructions can be described in terms of four basic operations. These, called the "four basic maps", are of the following type.

For a morphism $f\colon X \to Y$, there is a push-forward map

$$f_*\colon C_p(X;n) \to C_p(Y;n).$$

For a morphism $g\colon X \to Y$ with fiber dimension s, there is a pull-back map

$$g^*\colon C_p(Y;n) \to C_{p+s}(X;n-s).$$

Moreover there is "multiplication with K_1": for a global unit a on X, there is a map

$$\{a\}\colon C_p(X;n) \to C_p(X;n+1)$$

given by pointwise multiplication with $a(x) \in \kappa(x)^* = K_1\kappa(x)$.

Finally for a closed immersion there is a canonical "boundary map"

$$\partial\colon C_p(X \setminus Y;n) \to C_{p-1}(Y;n).$$

All these maps are defined in a pointwise manner. If f is proper and g is flat, the maps f_* and g^* commute with the differentials of the complexes. One uses f_* also for open immersions f and g^* also for closed immersions g (then f_* and g^* are just the corresponding projections and don't commute with the differentials). The maps $\{a\}$ and ∂ anti-commute with the differentials.

In fact, the four basic maps are enough to define intersections on complex level: by their very definition, the maps $I(f)$ are sums of compositions of the four basic maps. For the construction of the $I(f)$, the first major tool is the deformation to the normal cone. This yields a canonical "deformation map"

$$C_*(X;n) \to C_*(N;n)$$

where N is the normal cone of f. The next step is to define for a vector bundle $\pi\colon V \to X$ of dimension d a homotopy inverse

$$C_*(V;n) \to C_{*-d}(X;n+d)$$

to the pull-back map π^*. It is at this place where one needs some extra noncanonical choices. The choice to be made is (at most) that of what we call a "coordination" of π. This is a stratification of X together with bundle trivializations on the strata.

In the end there is a canonical procedure which starts from the choice of a coordination of the normal bundle of f and yields a map $I(f)$ as desired, defined in terms of the four basic maps. Different choices lead to homotopic maps $I(f)$, with the homotopies again expressible in terms of the four basic maps. In a fibered situation, one may arrange things to end up with filtration preserving maps $I(f)$. Once having made the necessary choices, the construction is quite functorial. For example, it is compatible with respect to base change and localization. In order to establish functoriality (namely $I(f \circ f')$ should be homotopic to $I(f') \circ I(f)$, if necessary under a filtration preserving homotopy), we use a kind of doubled deformation space.

The viewpoint of the paper is to put the four basic maps in the center. In particular the maps $\{a\}$, ∂ are treated as if they were a kind of morphisms in their own right, of equal rank as the more familiar push-forward and pull-back maps. This has at least technical advantages. For example, in order to check various compatibilities concerning the maps $f^\bullet$, it is very convenient to reduce to a separate treatment of the four basic maps.

The reader may ask why we insist to stay on complex level although one is interested mainly in the Chow groups. Over some range this is quite natural from the material. However, the proof of homotopy invariance with respect to vector bundles is much simpler for the Chow groups (using the spectral sequences) than for the cycle complexes themselves (where one has to construct explicit homotopy inverses).

The major motive for keeping the complex level throughout was to keep control on the filtrations in fibered situations.

Besides this, we hope that our method is of some interest concerning questions for correspondences between arbitrary varieties. To give an example let $f\colon \widetilde{X} \to X$ be a proper birational morphism with X smooth. Then there are pull-back maps

$$I(f)\colon C_*(X;n) \to C_*(\widetilde{X};n)$$

similar to the $I(f)$ above. The $I(f)$ are unique up to homotopy and have the standard push-forward map f_* as left inverse. In particular, $I(f)$ identifies $C_*(X;n)$ as a subcomplex of $C_*(\widetilde{X};n)$. In the case of a blow up in a point x, the choice to be made in the construction of $I(f)$ is (at most) that of a system of parameters around x.

We think of the maps $I(f)$ as a sort of generalized correspondences. One can make this more precise in a further development which we call bivariant theory of cycles. There the four basic maps find their place as morphisms of varieties in an appropriate differential category and (the homotopy classes of) the maps $I(f)$ appear rather as morphisms in a category of varieties admitting products, than just as homomorphisms of complexes (as in this paper).

The motive of introducing a general notion of coefficient systems for cycles appears when looking at the spectral sequence (1). Its E^2-terms are the homology groups of complexes of type

$$\cdots \xrightarrow{d} \coprod_{z\in Z_{(p)}} A_q(X_z;n+1) \xrightarrow{d} \coprod_{z\in Z_{(p-1)}} A_q(X_z;n) \xrightarrow{d} \cdots .$$

We interpret this by saying that the collection of functors (with $n \in \mathbb{Z}$)

$$A_q[\pi;n]\colon F \mapsto A_q(X \times_Z \operatorname{Spec} F;n),$$

defined on fields F over Z, appear as new coefficient systems. This process of creating coefficient systems may in fact be iterated.

Therefore it seems convenient to have available some appropriate general notion of coefficient systems. The class considered in this paper is provided by the notion of what we call cycle modules. Its definition is formal and somewhat ad hoc. The important thing for us is, that it contains standard functors like Milnor's (or Quillen's) K-theory and Galois cohomology (as indicated above), that it is closed under processes like $M \to A_q[\pi;M]$ and that it allows intersection theory. Anyway it might be of at least heuristic interest, that many general constructions (intersections, also the proof of acyclicity for smooth local rings) can be based on pure formal properties—at least if one starts from Milnor's K-theory of fields.

Milnor's K-theory is the fundamental base of the whole paper. This was at first suggested by our original problem, Hilbert's Satz 90 for Milnor's K_n. Besides this, Milnor's K-theory seems to give the minimal framework needed in order to express the considerations on intersections discussed above. By the way, it seems likely that the general method works also with Milnor's K-theory replaced by the Witt ring of quadratic forms of fields of characteristic different from 2.

Milnor's K-theory has a simple definition in terms of generators and relations. Despite this fact, it is by no means a simple and well understood functor. Already to define the norm homomorphisms takes some effort. An even more serious and in general an unsolved problem is for example the computation of the torsion in Milnor's K-groups. These problems are related with Hilbert's Satz 90 (Merkurjev and Suslin 1982, 1986) and are part of a broader picture (Beilinson conjectures, motivic cohomology). In this context there appear other and more general higher versions of the classical Chow groups than the groups $A_p(X;n)$ based on Milnor's K-theory, namely motivic cohomology (Bloch's higher Chow groups and Suslin's singular homology) and also K-cohomology (Bloch 1986; Quillen 1973; Suslin and Voevodsky 1996). Milnor's K-theory forms a central part of motivic cohomology and of Quillen's K-theory. In fact, in the smooth case there are natural maps from the motivic cohomology of X to the groups $A_*(X;n)$ and from $A_*(X;n)$ to the K-cohomology of X, both of which are isomorphisms in some low degrees. On the other hand, motivic cohomology and Quillen's K-theory give rise to cycle modules in our sense. (In the case of Quillen's K-theory this is made more precise in Sections 1, 2 and 5). These functors are definitely necessary for a full understanding of Milnor's K-theory. For the purpose of this paper however, it turned out to be enough to rely on elementary properties of Milnor's K-theory.

The paper may be roughly divided in four parts. In Sections 1–2 the notion of cycle modules is defined. Here we have lent some weight to a discussion of the axioms. In Sections 3–5 the cycle complexes, the Chow groups and their basic functorial behavior are established; Section 6 is a side remark concerning the acyclicity of Gersten-type resolutions. Sections 7–8 treat the spectral sequences. Sections 9–14 are concerned with intersection theory.

I am indebted to Inge Meier for typesetting a first version of this paper.

1. Cycle Premodules

Cycle premodules are roughly said functors on fields which have transfer, are modules over Milnor's K-theory, are equipped with residue maps for discrete valuations and satisfy the "usual rules". The definition is quite formal. It forms the local dimension 1 part of the notion of cycle modules. A major difference to cycle modules is that cycle premodules do not have to obey laws involving an infinite number of valuations like the sum formula for $\mathbb{P}^1$.

Cycle premodules are defined by a list of data and rules. These are just usual properties, quite familiar to standard examples. Equivalently, one may define cycle premodules as the additive functors on a certain category which has an explicit description in terms of Milnor's K-theory and valuations (see Remark 1.10). This point of view is perhaps more satisfying. It tells that our list of data and rules is in a sense a complete list. However, it would take some effort to establish the composition rule in the category and we omit therefore a detailed discussion. Moreover, in order to establish certain functors as cycle premodules, it is more convenient to refer to the explicit lists of properties.

The viewpoint of the four basic maps mentioned in the introduction would at first lead to functors $F \to M(F)$, such that each $M(F)$ is a module over the tensor algebra TF^*. However, the existence of norm maps and the homotopy property leads one to pass to modules over Milnor's K-ring (see Remark 2.7).

We first recall basic facts from Milnor's K-theory. Let F be a field. By definition Milnor's K-ring (Milnor 1970) of F is[◇]

$$K_*F = TF^*/J$$

where F^* is the multiplicative group of F, TF^* is the tensor algebra of F^* as abelian group and J is the two-sided ideal of TF^* generated by the set

$$\{\, a \otimes b \mid a, b \in F^*,\ a + b = 1 \,\}.$$

The standard grading on TF^* induces a grading

$$K_*F = \coprod_{n \geq 0} K_nF.$$

K_nF is the n-th Milnor's K-group of F. By definition $K_0F = \mathbb{Z}$ and $K_1F = F^*$. The elements of K_nF represented by tensors $a_1 \otimes \cdots \otimes a_n$, $a_i \in F^*$, are called symbols and denoted by $\{a_1, \ldots, a_n\}$. The group law in K_nF is written additively, e.g., $\{ab\} = \{a\} + \{b\}$. There are the rules $\{a, -a\} = 0$ and $\{a, b\} + \{b, a\} = 0$, see (Milnor 1970). In particular, K_*F is an anti-commutative ring with respect to the natural $\mathbb{Z}/2$-grading.

For a homomorphism of fields $\varphi\colon F \to E$ there is the ring homomorphism

$$\varphi_*\colon K_*F \to K_*E,$$
$$\varphi_*(\{a_1, \ldots, a_n\}) = \{\varphi(a_1), \ldots, \varphi(a_n)\}.$$

[◇] In the literature one often uses the notation $K_*^{\mathrm{M}}F$ for Milnor's K-ring, while K_*F stands for Quillen's K-ring.

If φ is finite, there is the norm homomorphism

$$\varphi^*\colon K_*E \to K_*F.$$

φ^* preserves the $\mathbb{Z}$-grading. Its component $\mathbb{Z} \to \mathbb{Z}$ in degree 0 is multiplication with $\deg \varphi = [E : F]$. In degree 1 it is the usual norm map $N_\varphi\colon E^* \to F^*$ for the finite field extensions. φ^* has been defined by Bass and Tate (1972) with respect to a choice of generators of E over F; it is in fact independent of such a choice (Kato 1980). For a characterization of φ^* see the remark after Theorem 1.4.

For a valuation $v\colon F^* \to \mathbb{Z}$ we denote by $\mathcal{O}_v$, $\mathfrak{m}_v$, $\kappa(v)$ its ring, maximal ideal and residue class field, respectively. For nontrivial v there is the residue homomorphism $\partial_v\colon K_*F \to K_*\kappa(v)$, see (Milnor 1970). ∂_v is of degree -1. It has the characterizing properties

$$\begin{aligned} \partial_v(\{\pi, u_1, \dots, u_n\}) &= \{\bar{u}_1, \dots, \bar{u}_n\}, \\ \partial_v(\{u_1, \dots, u_n\}) &= 0 \end{aligned}$$

for a prime π of v and for v-units u_i with residue classes $\bar{u}_i \in \kappa(v)^*$. Define

$$\begin{aligned} &s_v^\pi\colon K_*F \to K_*\kappa(v), \\ &s_v^\pi(x) = \partial_v(\{-\pi\} \cdot x). \end{aligned}$$

s_v^π is a ring homomorphism and is characterized by

$$\begin{aligned} s_v^\pi(\{u_1, \dots, u_n\}) &= \{\bar{u}_1, \dots, \bar{u}_n\}, \\ s_v^\pi(\{\pi, u_1, \dots, u_n\}) &= 0. \end{aligned}$$

Rules between the maps φ_*, φ^*, ∂_v and the multiplicative structure of K_* are comprised below in Theorem 1.4.

Let B be a scheme over a field k (recall our conventions). In the following we mean by a field over B a field F together with a morphism $\operatorname{Spec} F \to B$ such that F is finitely generated over k. By a valuation over B we mean a discrete valuation v of rank 1 together with a morphism $\operatorname{Spec} \mathcal{O}_v \to B$ such that v is of geometric type over k. The latter means that $\mathcal{O}_v$ is the localization of an integral domain of finite type over k in a regular point of codimension 1. Alternatively, valuations of geometric type may be characterized by: $k \subset \mathcal{O}_v$, the quotient field F and the residue class field $\kappa(v)$ are finitely generated over k and $\operatorname{tr.deg}(F|k) = \operatorname{tr.deg}(\kappa(v)|k) + 1$.

This geometric setting is convenient for our later purposes. We impose its restrictive conditions from the beginning in order to keep things straight. For some purposes one may consider also arbitrary fields and valuations (discrete, of rank 1 and eventually not equicharacteristic) over an arbitrary scheme B.

In the following, the letters φ, ψ stand for homomorphisms of fields over B and all maps between various $M(F)$, $M(E)$, ... are understood as homomorphisms of graded abelian groups.

(1.1) DEFINITION. *Let $\mathcal{F}(B)$ be the class of fields over B. A cycle premodule M consists of an object function $M\colon \mathcal{F}(B) \to \mathcal{A}$ to the class of abelian groups together with a $\mathbb{Z}/2$-grading $M = M_0 \oplus M_1$ or a $\mathbb{Z}$-grading $M = \coprod_n M_n$ and with the following data* D1–D4 *and rules* R1a–R3e.

D1: For each $\varphi\colon F \to E$ there is $\varphi_*\colon M(F) \to M(E)$ of degree 0.
D2: For each finite $\varphi\colon F \to E$ there is $\varphi^*\colon M(E) \to M(F)$ of degree 0.
D3: For each F the group $M(F)$ is equipped with a left K_*F-module structure denoted by $x \cdot \rho$ for $x \in K_*F$ and $\rho \in M(F)$. The product respects the gradings: $K_nF \cdot M_m(F) \subset M_{n+m}(F)$.
D4: For a valuation v on F there is $\partial_v\colon M(F) \to M\big(\kappa(v)\big)$ of degree -1.

For a prime π of v on F we put

$$s_v^\pi\colon M(F) \to M\big(\kappa(v)\big),$$
$$s_v^\pi(\rho) = \partial_v(\{-\pi\} \cdot \rho).$$

R1a: For $\varphi\colon F \to E$, $\psi\colon E \to L$ one has $(\psi \circ \varphi)_* = \psi_* \circ \varphi_*$.
R1b: For finite $\varphi\colon F \to E$, $\psi\colon E \to L$ one has $(\psi \circ \varphi)^* = \varphi^* \circ \psi^*$.
R1c: Let $\varphi\colon F \to E$, $\psi\colon F \to L$ with φ finite. Put $R = L \otimes_F E$. For $p \in \operatorname{Spec} R$ let $\varphi_p\colon L \to R/p$, $\psi_p\colon E \to R/p$ be the natural maps. Moreover let l_p be the length of the localized ring $R_{(p)}$. Then

$$\psi_* \circ \varphi^* = \sum_p l_p \cdot (\varphi_p)^* \circ (\psi_p)_*.$$

R2: For $\varphi\colon F \to E$, $x \in K_*F$, $y \in K_*E$, $\rho \in M(F)$, $\mu \in M(E)$ one has (with φ finite in the projection formulae R2b and R2c):
R2a: $\varphi_*(x \cdot \rho) = \varphi_*(x) \cdot \varphi_*(\rho)$,
R2b: $\varphi^*\big(\varphi_*(x) \cdot \mu\big) = x \cdot \varphi^*(\mu)$,
R2c: $\varphi^*\big(y \cdot \varphi_*(\rho)\big) = \varphi^*(y) \cdot \rho$.

R3a: Let $\varphi\colon E \to F$ and let v be a valuation on F which restricts to a nontrivial valuation w on E with ramification index e. Let $\bar\varphi\colon \kappa(w) \to \kappa(v)$ be the induced map. Then

$$\partial_v \circ \varphi_* = e \cdot \bar\varphi_* \circ \partial_w.$$

R3b: Let $\varphi\colon F \to E$ be finite and let v be a valuation on F. For the extensions w of v to E let $\varphi_w\colon \kappa(v) \to \kappa(w)$ be the induced maps. Then

$$\partial_v \circ \varphi^* = \sum_w \varphi_w^* \circ \partial_w.$$

R3c: Let $\varphi\colon E \to F$ and let v be a valuation on F which is trivial on E. Then

$$\partial_v \circ \varphi_* = 0.$$

R3d: Let φ, v be as in R3c, let $\bar\varphi\colon E \to \kappa(v)$ be the induced map and let π be a prime of v. Then

$$s_v^\pi \circ \varphi_* = \bar\varphi_*.$$

R3e: For a valuation v on F, a v-unit u and $\rho \in M(F)$ one has

$$\partial_v(\{u\} \cdot \rho) = -\{\bar u\} \cdot \partial_v(\rho).$$

□

The maps φ_*, φ^* are called the restriction and corestriction homomorphisms, respectively. We use the notations $\varphi_* = r_{E|F}$, $\varphi^* = c_{E|F}$ if there is no ambiguity.
Note that R2c with $y = 1 \in K_0E$ gives

R2d: For finite φ: $F \to E$ one has

$$\varphi^* \circ \varphi_* = (\deg \varphi) \cdot \mathrm{id}.$$

Moreover R1c implies

R2e: For finite totally inseparable φ: $F \to E$ one has

$$\varphi_* \circ \varphi^* = (\deg \varphi) \cdot \mathrm{id}.$$

We consider $M(F)$ also as a right K_*F-module via

$$\rho \cdot x = (-1)^{nm} x \cdot \rho$$

for $x \in K_nF$ and $\rho \in M_m(F)$.
The maps ∂_v are called the residue homomorphisms and the maps s_v^π are called the specialization homomorphisms. It is easy to check that R3e implies

R3f: For a valuation v on F, $x \in K_nF$, $\rho \in M(F)$ and a prime π of v one has

$$\begin{aligned}\partial_v(x \cdot \rho) &= \partial_v(x) \cdot s_v^\pi(\rho) + (-1)^n s_v^\pi(x) \cdot \partial_v(\rho) + \{-1\} \cdot \partial_v(x) \cdot \partial_v(\rho),\\ s_v^\pi(x \cdot \rho) &= s_v^\pi(x) \cdot s_v^\pi(\rho).\end{aligned}$$

If π' is another prime and u is the v-unit with $\pi' = \pi u$, then

$$s_v^{\pi'}(x) = s_v^\pi(x) - \{\bar{u}\} \cdot \partial(x).$$

From this and the rule R3c it follows in particular that the rule R3d holds for every prime π.

More remarks concerning these formulae and the residue homomorphisms in general are given below.

All relevant cycle premodules M known to us are $\mathbb{Z}$-graded with $M_n = 0$ for $n < 0$. Within the general theory however there is need only for a $\mathbb{Z}/2$-grading and we will understand this case if not mentioned otherwise.

A morphism f: $B' \to B$ defines a transformation $\mathcal{F}(B') \to \mathcal{F}(B)$ and the restriction of a cycle premodule M over B to $\mathcal{F}(B')$ is a cycle premodule over B'. It will be sometimes denoted by f^*M but mostly by M as well. If $B = \operatorname{Spec} R$ is affine, we call a cycle premodule over B a cycle premodule over R. If R is a field, we speak of a constant cycle premodule. The reference to the base B will be often dropped.

(1.2) DEFINITION. *A pairing $M \times M' \to M''$ of cycle premodules over B is given by bilinear maps for each F in $\mathcal{F}(B)$*

$$\begin{gathered}M(F) \times M'(F) \to M''(F),\\ (\rho, \mu) \mapsto \rho \cdot \mu\end{gathered}$$

which respect the gradings and which have the properties P1–P3 *stated below.*

A ring structure on a cycle premodule M is a pairing $M \times M \to M$ which induces on each $M(F)$ an associative and anti-commutative ring structure.

P1: For $x \in K_*F$, $\rho \in M(F)$, $\mu \in M'(F)$ one has
P1a: $(x \cdot \rho) \cdot \mu = x \cdot (\rho \cdot \mu)$,
P1b: $(\rho \cdot x) \cdot \mu = \rho \cdot (x \cdot \mu)$.

P2: For $\varphi\colon F \to E$, $\eta \in M(F)$, $\nu \in M(E)$, $\rho \in M'(F)$, $\mu \in M'(E)$ one has (with φ finite in P2b, P2c)
P2a: $\varphi_*(\eta \cdot \rho) = \varphi_*(\eta) \cdot \varphi_*(\rho)$,
P2b: $\varphi^*\big(\varphi_*(\eta) \cdot \mu\big) = \eta \cdot \varphi^*(\mu)$,
P2c: $\varphi^*\big(\nu \cdot \varphi_*(\rho)\big) = \varphi^*(\nu) \cdot \rho$.

P3: For a valuation v on F, $\eta \in M_n(F)$, $\rho \in M'(F)$ and a prime π of v one has

$$\partial_v(\eta \cdot \rho) = \partial_v(\eta) \cdot s_v^\pi(\rho) + (-1)^n s_v^\pi(\eta) \cdot \partial_v(\rho) + \{-1\} \cdot \partial_v(\eta) \cdot \partial_v(\rho).$$

□

Note that P3 implies

$$s_v^\pi(\eta \cdot \rho) = s_v^\pi(\eta) \cdot s_v^\pi(\rho).$$

(1.3) Definition. *A homomorphism $\omega\colon M \to M'$ of cycle premodules over B of even resp. odd type is given by homomorphisms*

$$\omega_F\colon M(F) \to M'(F)$$

which are even resp. odd and which satisfy (with the signs corresponding to even resp. odd type)

(1) $\varphi_* \circ \omega_F = \omega_E \circ \varphi_*$,
(2) $\varphi^* \circ \omega_E = \omega_F \circ \varphi^*$,
(3) $\{a\} \cdot \omega_F(\rho) = \pm\omega_F\left(\{a\} \cdot \rho\right)$,
(4) $\partial_v \circ \omega_F = \pm\omega_{\kappa(v)} \circ \partial_v$.

A unit a on B provides a simple example of a homomorphism of odd type, namely $\{a\}\colon M \to M$ given by $\{a\}_F(\rho) = \{a_F\} \cdot \rho$ where $a_F \in F^*$ is the restriction of a.

The cycle premodules over B together with the notion of homomorphism of Definition 1.3 form an ($\mathbb{Z}/2$-graded) abelian category.

(1.4) Theorem. *Milnor's K-theory K_* together with the data*

$$\varphi_*,\ \varphi^*,\ \text{multiplication},\ \partial_v$$

is a $\mathbb{Z}$-graded cycle premodule over any field k. With its multiplication, K_* is a cycle premodule with ring structure. □

This statement is a compact form of results in (Bass and Tate 1972; Kato 1980; Milnor 1970); we omit a detailed deduction.

Theorem 1.4 holds also in the setting of arbitrary fields and valuations (discrete of rank 1 and with a restriction in R3b, see Remark 1.8 below).

Given the rings K_*F for each F in $\mathcal{F}(\operatorname{Spec} k)$, the maps φ_*, φ^* and ∂_v are uniquely determined by R1b, R1c, P2, P3 and

(1) $\varphi_*(1) = 1,$
(2) $\varphi_*(\{a\}) = \{\varphi(a)\},$
(3) $\varphi^*(1) = \deg\varphi\cdot 1,$
(4) $\varphi^*(\{a\}) = \{N(\varphi(a))\},$
(5) $\partial_v(1) = 0,$
(6) $\partial_v(\{a\}) = v(a),$
(7) $\partial_v(\{a,b\}) = \{(-1)^{v(a)v(b)}b^{v(a)}a^{-v(b)} \bmod \mathfrak{m}_v\}.$

Here v denotes a normalized valuation: $v(\pi)=1$.

This statement is trivial for the maps φ_* and ∂_v; for the uniqueness of the maps φ^* see in particular (Bass and Tate 1972, p. 40).

The multiplication maps of the K_*F-module structures on $M(F)$ for each F give rise to a pairing of cycle premodules

$$K_*\times M\to M.$$

Here the axioms P1, P2, and P3 follow from D3, R2, and R3f.

In order to establish a cycle premodule it is convenient to use the following reduction.

(1.5) LEMMA. *For the validity of* R3d *it suffices* (*under presence of the other rules of* Definition 1.1) *to require* R3d *for the case* $E=\kappa(v)$.

Proof: By R1a the rule R3d holds for E if it holds for some extension E' of E with $E'\subset\mathcal{O}_v$. Moreover by R3a we may replace $\mathcal{O}_v$ by any unramified extension $\mathcal{O}'_v$ with the same residue class field (we don't want to pass to the henselization $\lim\mathcal{O}'_v$, since our fields should be finitely generated over k). Now by lifting a transcendence base of $\kappa(v)$ over E to $\mathcal{O}_v$ we may assume that $\kappa(v)$ is finite over E. Moreover we may assume that E is algebraically closed in any $\mathcal{O}'_v$ as above. Then $\kappa(v)$ is totally inseparable over E. Suppose $p=\operatorname{char}F>0$. We argue by induction on $[\kappa(v):E]$. Let $a\in E^*$ such that $E_1=E(\sqrt[p]{a})$ is contained in $\kappa(v)$ but not in any $\mathcal{O}'_v$. Then the extension v_1 of v to $F_1=F(\sqrt[p]{a})$ has ramification index p, has the same residue class field and $[\kappa(v_1):E_1]<[\kappa(v):E]$. Using R2c, R3b, R3c and R3e it is now easy to see that R3d holds for the pair (v,E) if it holds for the pair (v_1,E_1) (use the fact that the norm of a prime for v_1 is a prime for v). □

The rest of this section will not be used later within the general theory. However the following remarks may be of at least heuristic interest and we will refer to them partially in later side-remarks.

(1.6) REMARK. There is the following point of view concerning R3f. See also (Bass and Tate 1972; Milnor 1970, remark at the end of p. 323).

For a valuation $v\colon F^*\to\mathbb{Z}$ let

$$K_*(v)=K_*F\;/\;\{1+\mathfrak{m}_v\}\cdot K_*F.$$

Consider the ring homomorphisms

$$\tilde{p}\colon K_*F \to K_*(v),$$
$$i\colon K_*\kappa(v) \to K_*(v)$$

given by projection resp. by the formula

$$i(\{\bar{u}_1, \dots, \bar{u}_n\}) = \tilde{p}(\{u_1, \dots, u_n\})$$

for v-units u_i. There is an exact sequence

$$0 \longrightarrow K_*\kappa(v) \xrightarrow{i} K_*(v) \xrightarrow{\partial} K_*\kappa(v) \longrightarrow 0$$

with $\partial_v = \partial \circ \tilde{p}$. Any prime π gives rise to a section $y \mapsto \tilde{p}(\{\pi\}) \cdot i(y)$ of ∂.

We put

$$M(v) = K_*(v) \otimes_{K_*\kappa(v)} M\big(\kappa(v)\big).$$

Then there is an exact sequence

$$0 \longrightarrow M\big(\kappa(v)\big) \xrightarrow{i} M(v) \xrightarrow{\partial} M\big(\kappa(v)\big) \longrightarrow 0,$$

and the splittings above give for every π a decomposition of $K_*\kappa(v)$-modules

$$M(v) = M\big(\kappa(v)\big) \oplus M\big(\kappa(v)\big).$$

We define

$$p\colon M(F) \to M(v),$$
$$p(\rho) = 1 \otimes s_v^\pi(\rho) + \tilde{p}(\{\pi\}) \otimes \partial_v(\rho).$$

Note that p is independent of the choice of π. One has $\partial_v = (\partial \otimes 1) \circ p$.

Now R3f may be reformulated by saying that p is a module homomorphism over the ring homomorphism $\tilde{p}$. Similarly one may understand P3 via pairings

$$M(v) \otimes_{K_*(v)} M'(v) \to M''(v).$$

(1.7) Remark. A particular consequence of R3e is the fact that the subgroup

$$\{1 + \mathfrak{m}_v\} \cdot M(F)$$

is killed whenever one passes to $M\big(\kappa(v)\big)$. This seems to be a reasonable condition from a geometric point of view. However note that the continuous Steinberg symbol $K_2\mathbb{Q} \to \mathbb{Z}/2$ corresponding to the 2-adic valuation on $\mathbb{Q}$ (Milnor 1971, § 11) maps $\{5, 2\}$ to the nontrivial element.

(1.8) Remark. If one wants to consider arbitrary valuations (discrete and of rank 1), one has to require in R3b that the integral closure of $\mathcal{O}_v$ in E is finite over $\mathcal{O}_v$. This condition holds for geometric and for complete valuation rings, see (Serre 1968). By looking at completions and using R1c and R3a one may then derive for arbitrary valuations a formula

$$\partial_v \circ \varphi^* = \sum_w l_w \cdot \varphi_w^* \circ \partial_w$$

with certain integers l_w. This remark applies in particular to Milnor's K-theory.

(1.9) LEMMA. *In the situation of* R3a *let π be a prime of v, let τ be a prime of w and let u be the v-unit with $\pi^e = \tau u$. Then*

$$s_v^\pi \circ \varphi_* = \bar{\varphi}_* \circ s_w^\tau - \{\bar{u}\} \cdot \bar{\varphi}_* \circ \partial_w.$$

Proof: First note that the validity of the statement does not depend on the choices of π and τ. Moreover, if $E \subset K \subset F$ is an intermediate field, we may restrict to consider the extensions $K|E$ and $F|K$.

If $F|E$ is unramified ($e = 1$), we may take $\pi = \tau$ and the claim follows from R2a and R3a.

After lifting a transcendence base of $\kappa(v)$ over $\kappa(w)$ to $\mathcal{O}_v$ we may therefore assume that F is finite over E.

If $e = [F : E]$ (case of total ramification, see Serre 1968, Chap. I, § 6), we may take $\tau = -N_\varphi(-\pi)$; then $\bar{u} = 1$ and the claim follows from R2c and R3b. We have now already covered totally inseparable extensions.

For a separable finite extension $F|E$, let $L|E$ be a Galois extension containing F, fix some extension of v to L and let $D(L|F) \subset D(L|E) \subset \mathrm{Gal}(L|E)$ be the decomposition groups.

Then $F' = L^{D(L|F)}$ is unramified over F with the same residue class field; by R3a we are reduced to consider the extension $F'|E$. The field $E' = L^{D(L|E)}$ is contained in F'; since it is unramified over E, we know the claim for $E'|E$ and we are reduced to consider $F'|E'$. Let $K = L^U$ where

$$U = \{\, g \in \mathrm{Gal}(L|E') \mid g \text{ acts trivially on } \kappa(v) \,\}$$

is the inertia group. Then $K|E'$ is unramified and $F'|K$ is totally ramified. □

(1.10) REMARK. The rules and Lemma 1.9 show that every composite of maps between various $M(F)$ given by the data D1–D4 is a sum of composites of the form

$$\psi^* \circ (x \cdot \underline{\quad}) \circ \varphi_* \circ \partial_{v_r} \circ \cdots \circ \partial_{v_1} \circ (y \cdot \underline{\quad}).$$

This kind of normal form for composites can be made more precise as follows. There is a category $\widetilde{\mathcal{F}}$ with objects the class of arbitrary fields and with morphism groups

$$\mathrm{Hom}(F, E) = \coprod_v \coprod_H K_* H \widehat{\otimes}_{K_*\kappa(v)} K_*(v).$$

Here v runs through the valuations on F with value groups $\mathbb{Z}^r$ with lexicographical order and with $r \geq 0$. The groups $K_*(v)$ are defined exactly as above for $r = 1$. Moreover H runs through those composites of $\kappa(v)$ and E which are finite over E (and $\widehat{\otimes}$ denotes the graded tensor product).

Restricting to the class $\mathcal{F}(B)$ and to geometric higher rank valuations one obtains a category $\widetilde{\mathcal{F}}(B)$. The cycle premodules over B may be then characterized as the additive functors on $\widetilde{\mathcal{F}}(B)$. In this alternative definition all the rules including Lemma 1.9 are hidden in the composition law of $\widetilde{\mathcal{F}}$.

(1.11) Remark. — *Galois cohomology as cycle premodule.* Any torsion étale sheaf on B (with the torsion prime to $\operatorname{char} k$) gives rise via Galois cohomology to a cycle premodule over B. For simplicity we restrict here to the case $B = \operatorname{Spec} k$ with k a field and to finite Galois modules over k. For generalities of Galois cohomology we refer to (Serre 1968, 1994; Shatz 1972).

Let $\bar{k}$ be a separable closure of k, let r be prime to $\operatorname{char} k$, let $\mu_r \subset \bar{k}^*$ be the group of r-th roots of unity and let D be a finite continous $\operatorname{Gal}(\bar{k}|k)$-module of exponent r. For a field F over k let $\bar{F}$ be a separable closure containing $\bar{k}$. Then μ_r and D are $\operatorname{Gal}(\bar{F}|F)$-modules via $\operatorname{Gal}(\bar{F}|F) \to \operatorname{Gal}(\bar{k}|k)$. Put

$$\widetilde{H}^*(F; D) = \coprod_{n \geq 0} H^n(F; D \otimes \mu_r^{\otimes n}).$$

Here we use for a finite Galois module C the notation

$$H^n(F; C) = H^n(\operatorname{Gal}(\bar{F}|F); C) = \varinjlim H^n(\operatorname{Gal}(L|F); C)$$

where L runs through the finite Galois subextensions of $\bar{F}|F$ such that $\operatorname{Gal}(\bar{F}|L)$ acts trivially on C.

$\widetilde{H}^*(F; \mathbb{Z}/r)$ is a ring and $\widetilde{H}^*(F; D)$ is a module over $\widetilde{H}^*(F; \mathbb{Z}/r)$ via cup products.

The object function $H^*[D]$ on $\mathcal{F}(k)$ given by $H^*[D](F) = \widetilde{H}^*(F; D)$ is in a natural way a $\mathbb{Z}$-graded cycle premodule over k. This statement is just a collection of well-known properties of Galois cohomology. In the following we restrict ourselves to a description of the data D1–D4. The rules follow from standard properties of the cohomology of finite groups and from standard ramification theory.

D1 and D2: For $\varphi\colon F \to E$ let $\bar{\varphi}\colon \bar{F} \to \bar{E}$ be some extension over $\bar{k}$ and let $\tilde{\varphi}\colon \operatorname{Gal}(\bar{E}|E) \to \operatorname{Gal}(\bar{F}|F)$ be the induced map. Define φ_* as the usual restriction homomorphism induced from $\tilde{\varphi}$. For finite φ define φ^* as the usual transfer homomorphism induced from $\tilde{\varphi}$ times the degree of inseparability $[E : E \cap \bar{\varphi}(\bar{F})]$ (cf. Serre 1992).

D3: The K_*F-module structure on $\widetilde{H}^*(F; D)$ is given by cup products and the norm residue homomorphism

$$h_F\colon K_*F/rK_*F \to \widetilde{H}^*(F; \mathbb{Z}/r).$$

h_F is the $\mathbb{Z}$-graded ring homomorphism which in degree 1 is given by the Kummer isomorphism $F^*/(F^*)^r \to H^1(F; \mu_r)$. For the rule $h_F(\{a\}) \smile h_F(\{1-a\}) = 0$ see for example (Tate 1976) or Remark 2.7.

D4: Let E be the completion of F with respect to v. Then there is a natural exact sequence

$$1 \to I \to \operatorname{Gal}(\bar{E}|E) \to \operatorname{Gal}(\bar{\kappa}|\kappa) \to 1$$

where I is the inertia group. Put $D_n = D \otimes \mu_r^{\otimes n}$ and consider the corresponding Hochschild-Serre spectral sequences

$$E_2^{p,q} = H^p(\kappa; H^q(I; D_n)) \Longrightarrow H^{p+q}(E; D_n).$$

The cohomology of the inertia group I is given by $H^0(I;D_n)=D_n$, $H^1(I;D_n)=\operatorname{Hom}(\mu_r,D_n)=D_{n-1}$ and $H^q(I;D_n)=0$ for $q\geq 2$ (Serre 1968, Chap. IV). Hence the spectral sequences give rise to homomorphisms

$$\tilde{\partial}_v\colon H^n(E;D_n)\to H^{n-1}(\kappa;D_{n-1}).$$

Composing with $H^n(F;D_n)\to H^n(E;D_n)$ defines the desired maps[◊]

$$\partial_v\colon H^n[D](F)\to H^{n-1}[D](\kappa).$$

(1.12) Remark. — *Quillen's K-theory as cycle premodule.* We denote by $K'_*F=\coprod_n K'_nF$ Quillen's K-ring of a field F. Hereby we understand the definition $K'_nF=\pi_{n+1}\big(\mathrm{BQ}\,\mathit{Mod}(F)\big)$ of (Quillen 1973) with the product as defined in (Grayson 1978). (Here $\mathit{Mod}(F)$ is the category of finite dimensional F-modules. For generalities of Quillen's K-theory see also Grayson 1976; Srinivas 1991.)

The object function $F\to K'_*F$ defines a $\mathbb{Z}$-graded cycle premodule with ring structure over any field k. Its data are given as follows.

D1 and D2: One takes the pull-back map $\bar{\varphi}^*$ resp. the push-forward map $\bar{\varphi}_*$ of (Quillen 1973, § 7) where $\bar{\varphi}\colon \operatorname{Spec}E\to\operatorname{Spec}F$ is the morphism corresponding to φ.

D3: One uses the natural homomorphism $\omega\colon K_*F\to K'_*F$ from Milnor's to Quillen's K-theory. To define ω, one may refer to $K'_nF=\pi_n(\mathrm{BGL}(F)^+)$ and the computations $\pi_1(\mathrm{BGL}(F)^+)=H_1(\mathrm{GL}(F),\mathbb{Z})=K_1F$, $\pi_2(\mathrm{BGL}(F)^+)=H_2(\mathrm{E}(F);\mathbb{Z})=K_2F$ (Matsumoto's theorem, see Milnor 1971). Another possibility is to define directly a homomorphism $\omega_1\colon F^*\to K'_1F$ and then to check the rule $\omega_1(\{a\})\cdot\omega_1(\{1-a\})=0$ using the arguments of Remark 2.7.

D4: One uses the connecting map of the long exact localization sequence for $\mathcal{O}_v$ (Quillen 1973, § 7).

The verification of the rules is omitted. It is a lengthy but straightforward exercise to deduce them from (Grayson 1978; Quillen 1973).

[◊] According to the conventions made for the cup product and the spectral sequence, one may have different signs in the product rules for the differentials. This affects rule R3e, so if necessary, one should replace ∂_v by an appropriate sign (depending alone on n).

2. Cycle Modules

In this section we define the notion of a cycle module and derive important properties: the homotopy property for $\mathbb{A}^1$ and the sum formula for proper curves. Moreover we give a simplification of the axioms for a constant cycle module over a perfect field. The axioms of a cycle module are basic for all further considerations. Therefore we have included discussions on various related properties to a much larger extent than is actually needed in the following sections.

Throughout the section, M denotes a cycle premodule over some scheme B (recall our conventions).

For a scheme X over B we write $M(x) = M\big(\kappa(x)\big)$ for $x \in X$. The generic point of an irreducible scheme X is denoted by ξ or ξ_X. If X is normal, then for $x \in X^{(1)}$ the local ring of X at x is a valuation ring; let $\partial_x \colon M(\xi_X) \to M(x)$ be the corresponding residue homomorphism.

For x, $y \in X$ we define

$$\partial^x_y \colon M(x) \to M(y)$$

as follows. Let $Z = \overline{\{x\}}$. If $y \notin Z^{(1)}$, then $\partial^x_y = 0$. Otherwise let $\tilde{Z} \to Z$ be the normalization and put

$$\partial^x_y = \sum_{z|y} c_{\kappa(z)|\kappa(y)} \circ \partial_z \tag{2.1.0}$$

with z running through the finitely many points of $\tilde{Z}$ lying over y.

(2.1) Definition. *A cycle module M over B is a cycle premodule M over B which satisfies the following conditions* (FD) *and* (C).

(FD): Finite support of divisors. Let X be a normal scheme and $\rho \in M(\xi_X)$. Then $\partial_x(\rho) = 0$ for all but finitely many $x \in X^{(1)}$.

(C): Closedness. Let X be integral and local of dimension 2. Then

$$0 = \sum_{x \in X^{(1)}} \partial^x_{x_0} \circ \partial^\xi_x \colon M(\xi_X) \to M(x_0)$$

where ξ_X is the generic and x_0 is the closed point of X.

Many remarks and definitions of Section 1 are understood accordingly for cycle modules. For example a homomorphism of cycle modules is a homomorphism of the underlying cycle premodules.

Of course (C) has sense only under presence of (FD) which guarantees finiteness in the sum. More generally, note that if (FD) holds, then for any X, $x \in X$ and $\rho \in M(x)$ one has $\partial^x_y(\rho) = 0$ for all but finitely many $y \in X$.

If X is integral and (FD) holds for X, we put

$$d = (\partial^\xi_x)_{x \in X^{(1)}} \colon M(\xi_X) \longrightarrow \coprod_{x \in X^{(1)}} M(x).$$

In the following, F denotes a field over B and $\mathbb{A}^1_F = \operatorname{Spec} F[u]$ is the affine line over F with function field $F(u)$. Proofs of Proposition 2.2 and Theorem 2.3 are given after Remark 2.6.

(2.2) Proposition. *Let M be a cycle module over B. Then the following properties* (H) *and* (RC) *hold for all fields F over B.*

(H): Homotopy property for $\mathbb{A}^1$. The sequence

$$0 \longrightarrow M(F) \xrightarrow{r} M\big(F(u)\big) \xrightarrow{d} \coprod_{x\in\mathbb{A}^1_{F(0)}} M(x) \longrightarrow 0$$

is an exact complex (with $r = r_{F(u)|F}$).

(RC): Reciprocity for curves. Let X be a proper curve over F. Then

$$M(\xi_X) \xrightarrow{d} \coprod_{x\in X_{(0)}} M(x) \xrightarrow{c} M(F)$$

is a complex: $c\circ d = 0$ (with $c = \sum c_{\kappa(x)|F}$).

The properties (FD), (C), (H), (RC) are all what we need in further sections. Axiom (FD) enables one to write down the differentials d of the complexes $C_*(X;M)$, axiom (C) guarantees that $d\circ d = 0$, property (H) yields the homotopy invariance of the Chow groups $A_*(X;M)$ and finally property (RC) is needed to establish proper push-forward. For the material from Section 3 on the reader may take (H) and (RC) just as additional axioms of cycle modules and skip without much harm everything after Remark 2.6 below.

For another example of the fundamental role of axioms like (RC) in formal definitions of functors on fields see also (Somekawa 1990).

For integral X we put

$$A^0(X;M) = \ker d = \bigcap_{x\in X^{(1)}} \ker\partial^\xi_x \ \subset\ M(\xi_X).$$

One may think of $A^0(X;M)$ as the group of "unramified M-valued functions" on X.

(2.3) Theorem. *Let M be a cycle premodule over a perfect field k. Then M is a cycle module over k if and only if the following properties* (FDL) *and* (WR) *hold for all fields F over k.*

(FDL): Finite support of divisors on the line. Let $\rho \in M\big(F(u)\big)$. Then $\partial_v(\rho) = 0$ for all but finitely many valuations v of $F(u)$ over F.

(WR): Weak reciprocity. Let ∂_∞ be the residue map for the valuation of $F(u)|F$ at infinity. Then

$$\partial_\infty\big(A^0(\mathbb{A}^1_F;M)\big) = 0.$$

One implication here is obvious from Proposition 2.2, since trivially (FD) $\Rightarrow$ (FDL) and (RC) $\Rightarrow$ (WR). Conditions (FDL) and (WR) are comparatively weak: they deal only with the affine line and involve no corestriction maps. For nonconstant cycle modules (i.e., B is not the spectrum of a field) we don't know any similar simplification of the axioms (FD) and (C).

Further properties of cycle modules are

(Co): Continuity. Let X be smooth and local and let $Y \to X$ be the blow up in the unique closed point x_0. Then

$$A^0(X;M) \subset A^0(Y;M).$$

In other words, if v is the valuation corresponding to the exceptional fiber over x_0, then

$$\partial_v\big(A^0(X;M)\big) = 0.$$

(E): Evaluation. In the situation of (Co) there is a unique homomorphism

$$\text{ev: } A^0(X;M) \to M(x_0)$$

("evaluation at x_0") such that

$$r_{\kappa(v)|\kappa(x_0)} \circ \text{ev} = s_v^\pi \mid A^0(X;M)$$

for any prime π of v.

The validity of these two properties will follow from the construction of the pull-back maps $f^\bullet$: $A^0(X;M) \to A^0(Z;M)$ for morphisms f: $Z \to X$ in Section 12. Namely, the inclusion of (Co) is given by $f^\bullet$ with f: $Y \to X$ the blow up. Moreover in (E) one has ev $= f^\bullet$ with f: $\operatorname{Spec}\kappa(x_0) \to X$ the inclusion. See also Remark 2.8 below.

(2.4) Remark. A basic example of a cycle module over any field k is Milnor's K-ring K_*. Axiom (FD) follows as for classical divisors. For (H) see (Milnor 1970). The validity of (RC) for $X = \mathbb{P}^1$ is intrinsic to the definition of the norm homomorphisms in (Bass and Tate 1972). Kato (1986) has used (RC) to prove (C) by passing to completions.

(2.5) Remark. The cycle premodules $H^*[D]$ and K'_* of Remarks 1.11 and 1.12 are cycle modules. Axioms (FD) and (C) are contained in (Bloch and Ogus 1974) and in (Quillen 1973, § 7, Sect. 5), respectively.

For $H^*[D]$ one may use here alternatively Theorem 2.3 and Tsen's Theorem as follows (see also Serre 1992). (FDL) follows from the fact that every finite extension of $F(u)$ is ramified only in finitely many places of $F(u)|F$. One has trivially (H) $\Rightarrow$ (WR). If F is separably closed, then (H) follows from Tsen's Theorem (i.e., $H^q(F(u);\mu_r) = 0$ for $q \geq 2$; see for example Shatz 1972) and the Kummer isomorphism $K_1F(u)/r = H^1(F(u);\mu_r)$. To deduce (H) for arbitrary F one applies the Hochschild-Serre spectral sequence for the extension $\bar{F}(u)|F(u)$.

(2.6) Remark. Probably the considerations of this section (and of the whole paper) may be developed in characteristic $\neq 2$ also for a version of cycle modules which are modules over the Witt ring of quadratic forms instead over Milnor's K-ring. A transferring would be not at all formal because the residue maps for the Witt ring depend on choices of parameters.

In the following proofs of Proposition 2.2 and Theorem 2.3 we use the notations $\mathbb{A}^1 = \operatorname{Spec} F[u]$, $\mathbb{A}^2 = \operatorname{Spec} F[s,t]$ and $Z = \mathbb{A}^2_{(<s,t>)}$, the localization of $\mathbb{A}^2$ at 0. Moreover y, $z \in Z^{(1)} \subset \mathbb{A}^{2(1)}$ denote the points with parameters s, t, respectively. We proceed in several steps.

Step 1: (FD) + (C) $\Rightarrow$ (WR). Given $\rho \in A^0(\mathbb{A}^1; M)$ put

$$\eta = \{t\} \cdot \rho(t/s) \in M\big(F(s,t)\big)$$

or, more precisely, $\eta = \{t\} \cdot \varphi_*(\rho)$ with $\varphi\colon F(u) \to F(s,t)$, $\varphi(u) = t/s$. Using R2 and R3 one finds

$$\begin{aligned}
&\partial_x(\eta) = 0 \text{ for } x \in Z^{(1)} \setminus \{y,z\},\\
&\partial_y(\eta) = -\{t\} \cdot r_{\kappa(y)|F}\big(\partial_\infty(\rho)\big),\\
&\partial_z(\eta) = \partial_z\big(\varphi_*(\{u\}\cdot\rho) + \{s\}\cdot\varphi_*(\rho)\big)\\
&\qquad = r_{\kappa(z)|F} \circ \partial_0(\{u\}\cdot\rho) - \{s\} \cdot r_{\kappa(z)|F}\big(\partial_0(\rho)\big).
\end{aligned}$$

(C) and $\partial_0(\rho) = 0$ give

$$0 = \sum_{x \in Z^{(1)}} \partial_0^x \circ \partial_x(\eta) = \partial_0^y \circ \partial_y(\eta) = -\partial_\infty(\rho). \qquad \square$$

Step 2: (FDL) + (WR) $\Rightarrow$ (H). Note that $d \circ r = 0$ by R3c. Moreover any specialization map for an F-rational point on $\mathbb{P}^1$ is a left inverse to r by R3d.

Surjectivity of d: For a closed point $x \in \mathbb{A}^1$ let

$$\Phi^x\colon M(x) \to M\big(F(u)\big),$$
$$\Phi^x(\mu) = c_{\kappa(x)(u)|F(u)}\big(\{u - u(x)\} \cdot r_{\kappa(x)(u)|\kappa(x)}(\mu)\big)$$

and let

$$\Phi = \sum_x \Phi^x\colon \coprod_{x \in \mathbb{A}^1_{(0)}} M(x) \longrightarrow M\big(F(u)\big).$$

Then $d \circ \Phi = \mathrm{id}$ by R3b–R3e.

Exactness at $M(F(u))$: Given $\rho \in A^0(\mathbb{A}^1; M)$, put

$$\eta = \{t\} \cdot \big(\rho(u+t) - \rho(u)\big) \in M(F(u)(t)).$$

More precisely: Let $E = F(u)$, let $i, \varphi\colon E \to E(t)$ be the homomorphisms over F with $i(u) = u$, $\varphi(u) = u + t$ and put $\eta = \{t\} \cdot \big(\varphi_*(\rho) - i_*(\rho)\big)$.

We compute $\partial_w(\eta)$ for the valuations w of $E(t)$ over E. One finds easily $\partial_w(\eta) = 0$ for $w \neq 0, \infty$. But also $\partial_0(\eta) = 0$ by R3d, since the valuation at $t = 0$ restricts trivially under i and φ and since the induced homomorphisms $E \to \kappa(0)$ coincide. Hence (WR) tells $\partial_\infty(\eta) = 0$. On the other hand one has

$$\begin{aligned}
\partial_\infty\big(\{t\}\cdot i_*(\rho)\big) &= -\rho,\\
\partial_\infty\big(\{t\}\cdot\varphi_*(\rho)\big) &= \partial_\infty\big((\{t/(u+t)\} + \{u+t\})\cdot\varphi_*(\rho)\big)\\
&= -\{\overline{t/(u+t)}\}\cdot\partial_\infty\big(\varphi_*(\rho)\big) + \partial_\infty\big(\varphi_*(\{u\}\cdot\rho)\big)\\
&= 0 + r_{E|F}\big(\partial_\infty(\{u\}\cdot\rho)\big)
\end{aligned}$$

by making particular use of R3d and R3e (note that $t/(u+t)$ has residue class $1 \in \kappa(\infty)$). So

$$\rho = r_{E|F}\big(\partial_\infty(\{u\}\cdot\rho)\big) \in r_{E|F}\big(M(F)\big). \qquad \square$$

Step 3: (FD) + (H) $\Rightarrow$ (RC). There is a finite morphism $X \to \mathbb{P}^1$ over F. Using this and R3b one reduces to the case $X = \mathbb{P}^1$. Then it suffices to check

$$\sum_{u \in \mathbb{P}^1_{(0)}} c_{\kappa(u)|F} \circ \partial_u \circ \Phi^x = 0$$

for Φ^x as in Step 2. This equation follows from the computation $d \circ \Phi = \mathrm{id}$ and

$$\partial_\infty \circ \Phi^x = -c_{\kappa(x)|F},$$

a consequence of R3b and R3d. $\square$

The proof of Proposition 2.2 is now complete. We next consider the nontrivial implication of Theorem 2.3. We will refer at some places to Sections 3 and 4, but only in a mild way. Note that (H) is available by Step 2.

Step 4: (FDL) $\Rightarrow$ (FD) for $X = \mathbb{A}^n$. Let $p_i \colon \mathbb{A}^n \to \mathbb{A}^{n-1}$ be the n standard projections. Then

$$\mathbb{A}^{n(1)} \subset \bigcup_i p_i^{-1}(\xi)$$

where ξ is the generic point of $\mathbb{A}^{n-1}$. (FD) follows from (FDL) applied to $F = \kappa(\xi)$. $\square$

Step 5: (FD) for $X = \mathbb{A}^2$ + (WR) $\Rightarrow$ (C) for $X = Z$. As in Step 2 we have

$$M(\xi_Z) = M\big(F(s)\big) + \sum_{x \in \left(\mathbb{A}^1_{F(s)}\right)_{(0)}} \Phi^x\big(M(x)\big)$$

with

$$\Phi^x(\mu) = c_{\kappa(x)(t)|F(s,t)}\big(\{t - t(x)\} \cdot r_{\kappa(x)(t)|\kappa(x)}(\mu)\big).$$

(C) holds obviously on $M\big(F(s)\big)$. Let us verify (C) on the image of Φ^x for fixed x. By $d \circ \Phi = \mathrm{id}$ in Step 2 we are reduced to check

$$\partial_0^y \circ \partial_y \circ \Phi^x = -\partial_0^x.$$

Let v run through the valuations on $\kappa(x)(t)$ which restrict on $F[s,t]$ to the valuation with parameter s (and corresponding to y). Let $\bar{v}$ be the restriction of v to $\kappa(x)$ and let $c(\bar{v}) \in \overline{\{x\}}$ be the center of $\bar{v}$. If $c(\bar{v}) \neq 0$, then $t(x)$ is a $\bar{v}$-unit with residue $t(c(\bar{v}))$. Suppose $0 \in \overline{\{x\}}$ and let R_x be the residue class ring of x localized at 0. The valuations v with $c(\bar{v}) = 0$ restrict in a one-to-one manner to the valuations w of $\kappa(x)$ with $R_x \subset \mathcal{O}_w$. For these one has $\bar{v}(t(x)) > 0$ (since $t(x)$ is nilpotent in R_x/sR_x).

In the following, u runs through $\overline{\{x\}} \cap \overline{\{y\}} \setminus \{0\}$. One finds

$$\begin{aligned}
\partial_y \circ \Phi^x(\mu) &= \sum_v c_{\kappa(v)|\kappa(y)} \circ \partial_v\big(\{t - t(x)\} \cdot r_{\kappa(x)(t)|\kappa(x)}(\mu)\big) \\
&= -\sum_{v, c(\bar v) \neq 0} c_{\kappa(v)|\kappa(y)}\Big(\{t - t\big(c(\bar v)\big)\} \cdot r_{\kappa(v)|\kappa(\bar v)} \circ \partial_{\bar v}(\mu)\Big) \\
&\qquad - \{t(y)\} \cdot \sum_{v, c(\bar v) = 0} c_{\kappa(v)|\kappa(y)} \circ r_{\kappa(v)|\kappa(\bar v)} \circ \partial_{\bar v}(\mu) \\
&= -\sum_u c_{\kappa(u)(t)|\kappa(y)}\big(\{t - t(u)\} \cdot r_{\kappa(u)(t)|\kappa(u)} \circ \partial^x_u(\mu)\big) \\
&\qquad - \{t(y)\} \cdot r_{\kappa(y)|F} \circ \partial^x_0(\mu).
\end{aligned}$$

Since $t(u) \neq 0$ the sum vanishes under ∂^y_0 and we are done by R3d. □

Step 6: Reduction of (C) to the case $\kappa(x_0) \subset O_X$. Let X be as in (C) and write $X = \operatorname{Spec} R$. Lift a transcendence base of $\kappa(x_0)$ over k to elements $t_i \in R$ and put $K = k(t_1, \ldots, t_n) \subset R$. Then $\kappa(x_0)$ is a finite extension of K. Since k is perfect, we may take here a transcendence base such that $\kappa(x_0)$ is separable over K. Let $X' = \operatorname{Spec} R \otimes_K \kappa(x_0)$, let $u \in X'$ be the canonical lift of x_0 and let X'' be the localization of X' in u. We assume that (FD) holds for X and X''. Consider the pull-back along the flat map $X'' \to X$, see Section 3. The induced map $M(x_0) \to M(u)$ is injective, since x_0 and u have the same residue class fields. An application of R3a and R3b (see Proposition 4.6.2) shows that (C) holds for X if (C) holds for X''. □

We know now in particular that (C) holds for every localization of $\mathbb{A}^2$ in some closed point.

Step 7: Proof of (FD). There exists a generically finite separable rational map $X \to \mathbb{A}^n_k$. All but finitely many $x \in X^{(1)}$ correspond to points of $\mathbb{A}^{n(1)}$. The argument of Step 4 yields a reduction to a plane curve X over some field K. So consider the case $X = \overline{\{x\}}$ for some $x \in \mathbb{A}^2_{(1)}$. We may assume that X maps dominantly to $\operatorname{Spec} F[s]$ so that Φ^x as in Step 5 is defined. Put $\eta = \Phi^x(\rho)$. We have $\partial_x(\eta) = \rho$. Moreover $\partial_u(\eta) \neq 0$ only for finitely many u (Step 4). The closure of $u \neq x$ meets X only in finitely many points. Now, since (C) holds for every localization of $\mathbb{A}^2$, we have $\partial^\xi_w(\rho) = 0$ for all but finitely many $w \in X^{(1)} \subset \mathbb{A}^2_{(0)}$. □

Step 8: Proof of (C). By Step 6 we may assume that $F = \kappa(x_0)$ is contained in O_X. Choose a closed (2-dimensional) subscheme $Y \subset \mathbb{P}^n_F$ such that X is the localization of Y in a (F-rational) point y. We consider the generic projection from $\mathbb{P}^n_F$ to $\mathbb{P}^2_F$. More precisely: let T be the Grassmannian of 3-codimensional linear subspaces of $\mathbb{P}^n_F$, let $E = F(T)$, let $H \subset \mathbb{P}^n_E$ be the tautological subspace and let π: $\mathbb{P}^n_E \setminus H \to \mathbb{P}^2_E$ be a linear projection. Then $H \cap Y_E = \varnothing$ and π restricts to a proper map p: $Y_E \to \mathbb{P}^2_E$. Let $D = p^{-1}\big(p(y)\big)$. Then D is the intersection of Y_E with a generic 2-codimensional linear subspace passing through y. Hence

$$D \setminus \{y\} \subset Y_E \setminus Y.$$

In particular $D \cap Y_{(0)} = \{y\}$. Now we consider flat pull-back along the base change $q\colon Y_E \to Y$ followed by the push-forward along p, see Section 3. One finds (see (1) and (2) of Proposition 4.6) for $\rho \in M(\xi_X)$:

$$r_{E|F}\Big(\sum_{x\in X^{(1)}} \partial^x_{x_0} \circ \partial^\xi_x(\rho)\Big) = \sum_{u\in U^{(1)}} \partial^u_{p(u)} \circ \partial_u\big(p_* \circ q^*(\rho)\big)$$

where U is the localization of $\mathbb{P}^2_E$ in $p(y)$. The right hand side vanishes by Step 5 and $r_{E|F}$ is injective since $E|F$ is a rational extension. $\square$

We conclude with some more considerations concerning the axioms of cycle modules. These have been included here more for illustration than for application. In order not to be too tiring, we have taken here the freedom to be a bit vague about our actual assumptions.

(2.7) Remark. In the datum D3 of cycle premodules there is hidden a strong rule, namely the relation $\{a, 1-a\} = 0$ of Milnor's K-theory. The main justification within this paper for using Milnor's K-theory is that it works well. Asking naively, one may try to weaken D3 by requiring only the existence of bilinear pairings

$$K_1F \times M(F) \to M(F),$$
$$\big(\{a\}, \rho\big) \mapsto \{a\} \cdot \rho$$

and restricting to $x \in K_1E$, $y \in K_1F$ in R2. Then $M(F)$ would be a TF^*-module.

However, if one wants to develop a geometric theory, one is in the end led to pass to modules over Milnor's K-theory. A reasoning for this is given by the following little game. It refers in a mild way to the rules of cycle premodules and to a part of the homotopy property (H).

Let $\rho \in M(F)$, let L be an overfield of F, let $u \in L \setminus \{0, 1\}$ and consider

$$\eta(u) = \{u\} \cdot \big(\{1-u\} \cdot r_{L|F}(\rho)\big) \in M(L).$$

Our aim is to conclude $\eta(u) = 0$ for the case $L = F$. Assuming reasonable specialization maps, this follows from the generic case $L = F(u)$ with u a variable. To treat this case, our strategy is to argue that $\eta(u)$ is unramified on the whole affine line. Then, by homotopy invariance, $\eta(u)$ is constant. An extra argument finally shows $\eta(u) = 0$.

To be specific first a little calculation (which provides by the way already the divisibility of $\eta(u)$ referring only to the existence of norm maps and the projection formula). Let $L' = F(u')$ be the function field in the variable u' and let $L = F(u) \subset L'$ with $u = u'^n$. Then $1 - u = N_{L'|L}(1-u')$ and the projection formulae R2b and R2c give

$$\begin{aligned}\eta(u) &= \{u\} \cdot \big(\{N_{L'|L}(1-u')\} \cdot r_{L|F}(\rho)\big)\\ &= \{u\} \cdot c_{L'|L}\big(\{1-u'\} \cdot r_{L'|F}(\rho)\big)\\ &= c_{L'|L}\Big(\{u'^n\} \cdot \big(\{1-u'\} \cdot r_{L'|F}(\rho)\big)\Big)\\ &= n \cdot c_{L'|L}\big(\eta(u')\big).\end{aligned}$$

We want to conclude that $\partial_v\bigl(\eta(u)\bigr) = 0$ for all finite places v of $L|F$. This is quite natural to assume as long as u and $1-u$ are v-units. For the place at $u=0$ (and similarly at $u=1$) one may argue as follows.

Let $\alpha = \partial_0\bigl(\eta(u)\bigr) \in M(F)$ be the residue for the valuation of $L|F$ at $u=0$. Similarly let $\alpha' = \partial_0\bigl(\eta(u')\bigr)$, now with respect to the valuation of $L'|F$ at $u'=0$. A change of variables $u \to u'$ shows $\alpha = \alpha'$. But the above computation and rule R3b yields $\alpha = n\cdot\alpha'$. Taking $n=2$ gives $\alpha = 0$.

Now (H) tells that $\eta(u)$ comes from $M(F)$, that is $\eta(u) = r_{L|F}(\eta)$ for some $\eta \in M(F)$. Naturality with respect to the homomorphisms $L \to L$ over F with $u \to -u$ and $u \to u^2$ gives

$$\eta(u) = \eta(-u) = \eta(u^2).$$

On the other hand one has

$$\eta(u^2) = 2\eta(u) + 2\eta(-u),$$

just by linearity. One concludes $3\eta(u) = 0$. But then $3\eta(u') = 0$ as well and the above computation for $n=3$ tells $\eta(u) = 0$.

(2.8) Remark. As already mentioned, the properties (Co) and (E) of cycle modules follow from the material in Section 12. The considerations there use the deformation to the normal cone and homotopy inverses. But things simplify considerably if one may pass to the limits $X = \operatorname{Spec} k[[t_1,\dots,t_n]]$. In the following we consider the case $X = \operatorname{Spec} k[[s,t]]$, tacitly assuming that our cycle modules are defined on an appropriate category of schemes. In fact we could have taken also more general schemes as basis for our notions, say excellent schemes over a perfect field (however one should then be careful with Theorem 2.3).

(C) $\Rightarrow$ (Co): Let $E = F(u)$ and $T = \operatorname{Spec} E[[s,t]]$. Given $\rho \in A^0(X;M)$ put

$$\eta = \{t - us\}\cdot r_{E((s,t))|F((s,t))}(\rho) \in M(\xi_T).$$

One may then calculate

$$\sum_{z\in T^{(1)}} \partial_0^z \circ \partial_z(\eta) = \partial_v(\rho).$$

Hence $\partial_v(\rho) = 0$ by (C).

(C) $\Rightarrow$ (E): By the last argument we may use (Co). It follows that for $\rho \in A^0(X;M)$ the value of $s_v^\pi(\rho)$ is independent of the choice of the prime π. Since $\kappa(v) = F(t/s)$ is rational, one is by (H) reduced to check

$$\partial_w \circ s_v^\pi(\rho) = 0$$

for all valuations w of $F(t/s)$ over F except the one with $w(t/s) = -1$. Every w defines a point in the exceptional fiber of the blow up $Y \to X$. One calculates for the w in question

$$0 = \sum_{y\in Y^{(1)}} \partial_w^y \circ \partial_y(\{s\}\cdot\rho) = \partial_w \circ \partial_v(\{s\}\cdot\rho).$$

Finally note $s_v^\pi(\rho) = \partial_v(\{s\}\cdot\rho)$ for the choice $\pi = -s$.

(2.9) Remark. In the case of a constant cycle premodule one may derive (C) from (Co) under presence of (FD). This tells that axiom (C) appears naturally in our framework if we require the existence of pull-back maps $f^\bullet$. As in Remark 2.8 we consider here the case $X = \operatorname{Spec} k[[s,t]]$.

(Co) $\Rightarrow$ (C): To derive (C) from (Co) for constant M and for $X = \operatorname{Spec} k[[s,t]]$ one argues first similarly as for Step 5 above as follows. Let y be the point with parameter s and define for $x \in X^{(1)} \setminus \{y\}$:

$$\Phi^x \colon M(x) \to M(\xi_X),$$
$$\Phi^x(\mu) = r_{F((s,t))|F((s))(t)} \circ c_{\kappa(x)(t)|F((s))(t)}\big(\{t - t(x)\} \cdot r_{\kappa(x)(t)|\kappa(x)}(\mu)\big).$$

As in Step 5 one has $\partial_x \circ \Phi^x = \mathrm{id}$ and $\partial_z \circ \Phi^x = 0$ for $z \neq x$, y; moreover one finds $\partial_y \circ \Phi^x = -\{t\} \cdot r_{F((t))|F} \circ \partial_0^x$. This shows that (C) holds on the image of the Φ^x.

In order to verify (C) for $\tilde\rho \in M(\xi_X)$ we may arrange things such that $\partial_y(\tilde\rho) = 0$. We are reduced to check (C) for

$$\rho = \tilde\rho - \sum_{x \neq y} \Phi^x \circ \partial_x(\tilde\rho).$$

Using the above computations one finds

$$\partial_x(\rho) = \begin{cases} 0 & \text{for} \quad x \neq y, \\ \{t\} \cdot r_{\kappa(y)|F}(\theta) & \text{for} \quad x = y \end{cases}$$

for some $\theta \in M(F)$.

We must show $\theta = 0$. Put $E = F(r)$ and—written in a somewhat sloppy form—

$$\eta = \rho(rs, rt) - \rho(s,t) - \{s, r\} \cdot \theta \in M\big(E((s,t))\big).$$

One computes $\eta \in A^0(\operatorname{Spec} E[[s,t]]; M)$ and

$$\partial_v(\eta) = -\{r\} \cdot \theta \in M\big(E(s/t)\big).$$

(Co) gives $\{r\} \cdot \theta = 0$ in $M\big(F(r, s/t)\big)$. Applying appropriate specialization and residue maps shows $\theta = 0$.

3. The Four Basic Maps

The purpose of this section is to introduce the cycle complexes and all the types of operations on them needed further on (except the cross products to be defined in Section 14).

Let M be a cycle module over X, let N be a cycle module over Y and let $U \subset X$, $V \subset Y$ be subsets. For a homomorphism

$$\alpha\colon \coprod_{x \in U} M(x) \longrightarrow \coprod_{y \in V} N(y)$$

we write $\alpha^x_y\colon M(x) \to N(y)$ for the components of α.

(3.1) Change of coefficients. Let $\omega\colon M \to N$ be a homomorphism of cycle modules over X and let $U \subset X$ be a subset. We put

$$\omega_\#\colon \coprod_{x \in U} M(x) \longrightarrow \coprod_{x \in U} N(x)$$

where $(\omega_\#)^x_x = \omega_{\kappa(x)}$ and $(\omega_\#)^x_y = 0$ for $x \neq y$.

(3.2) Cycle complexes. For a cycle module M over X and an integer p let

$$C_p(X; M) = \coprod_{x \in X_{(p)}} M(x).$$

We define

$$d = d_X\colon C_p(X; M) \to C_{p-1}(X; M)$$

by $d^x_y = \partial^x_y$ with ∂^x_y as in (2.1.0). This definition has sense by axiom (FD).

(3.3) Lemma. $d_X \circ d_X = 0$.

Proof: One has to check $(d \circ d)^x_z = 0$ for $x \in X_{(p+1)}$, $z \in X_{(p-1)}$. This is trivial if $z \notin \overline{\{x\}}$. Otherwise let Y be the localization of $\overline{\{x\}}$ in z. Since our schemes are catenary, we have $X_{(p)} \cap Y = Y_{(1)}$ and $\dim Y = 2$. Now apply axiom (C) to Y. $\square$

The complex $C_*(X; M) = (C_p(X; M), d_X)_{p \geq 0}$ is called the *complex of cycles on X with coefficients in M*.

When developing a theory of cycles, first natural questions are the following. Given a proper morphism $f\colon X \to Y$, what is the push-forward map f_* on cycles? Or, given a flat morphism $g\colon Y \to X$, what is the pull-back map g^* on cycles? In fact, we will define such maps. However these questions are not our guiding point of view. We rather fix schemes X, Y and numbers p, q and then ask: what is the class of maps

$$C_p(X; M) \to C_q(Y; M)$$

which we should consider? Our answer is then motivated by what we want to do with the complexes, namely developing intersection theory etc. This leads to the "four basic maps" as defined in (3.4)–(3.7).

The definitions of the basic maps "multiplication with K_1" and "boundary maps" in (3.6) and (3.7) are easy to understand. However our way of introducing push-forward and pull-back maps as in (3.4) and (3.5) deserves some words of comment. It turns out that these maps (denoted by f_* and $[\mathcal{A}, g, s]$) are sums of compositions of maps of simpler type, namely push-forward maps f_* for proper morphisms f, pull-back maps g^* for flat morphisms g and the projections i^* and inclusions j_* corresponding to closed (or open) subvarietes (see 3.10). This fact (which we will not prove) seems to be however only of heuristic interest. In fact it would be a nuisance if we had to consider at each step such a reduction of the language expressing the maps between the cycle complexes.

(3.4) Push-forward. For a morphism $f\colon X \to Y$ of schemes of finite type over a field we define

$$f_*\colon C_p(X; M) \to C_p(Y; M)$$

as follows. If $y = f(x)$ and if $\kappa(x)$ is finite over $\kappa(y)$, then $(f_*)^x_y = c_{\kappa(x)|\kappa(y)}$. Otherwise $(f_*)^x_y = 0$.

(3.5) Pull-back. Our main interest is to define the particular types of pull-back maps as considered in (3.5.5) below. In our general definition in (3.5.3) we define pull-back maps $C_p(X; M) \to C_q(Y; M)$ associated to any morphism $g\colon Y \to X$ of relative dimension $\leq q - p$. Moreover we use coherent sheaves $\mathcal{A}$ on Y as modifiers of the arising multiplicities. This construction gives great technical flexibility and is useful in Section 4.

(1) For a morphism $g\colon Y \to X$ let

$$s(g) = \max\left\{ \dim(y, Y) - \dim\bigl(g(y), X\bigr) \mid y \in Y \right\}.$$

Moreover let $Y_x = Y \times_X \operatorname{Spec} \kappa(x)$ for $x \in X$.

Note that if $x \in X_{(p)}$, $y \in Y_{(q)}$, $g(y) = x$ and $s(g) \leq q - p$, then necessarily $y \in Y_x^{(0)}$.

(2) Let $g\colon Y \to X$ be a morphism and let $\mathcal{A}$ be a coherent sheaf on Y. For $x \in X$ and $y \in Y_x^{(0)}$ we define an integer

$$[\mathcal{A}, g]^x_y \in \mathbb{Z}$$

as follows. The localization $Y_{x,(y)}$ of Y_x in y is the spectrum of an artinian ring R with only residue class field $\kappa(y)$. Let $\tilde{\mathcal{A}}$ be the pull-back of $\mathcal{A}$ via $Y_{x,(y)} \to Y_x \to Y$ and define $[\mathcal{A}, g]^x_y = l_R(\tilde{\mathcal{A}})$ as the length of $\tilde{\mathcal{A}}$ considered as R-module (for the notion of length and further properties we refer to Fulton 1984, App. A).

(3) Fix $s \in \mathbb{Z}$. Let $g\colon Y \to X$ be a morphism with $s(g) \le s$ and let $\mathcal{A}$ be a coherent sheaf on Y. We define homomorphisms

$$[\mathcal{A}, g, s]\colon C_p(X; M) \to C_{p+s}(Y; M)$$

by

$$[\mathcal{A}, g, s]_y^x = \begin{cases} [\mathcal{A}, g]_y^x \cdot r_{\kappa(y)|\kappa(x)} & \text{if } g(y) = x, \\ 0 & \text{otherwise.} \end{cases}$$

Here $\kappa(x)$ is considered as a subfield of $\kappa(y)$ via g.

(4) Let F be a field, let $g\colon Y \to \operatorname{Spec} F$ be a morphism and let

$$0 \to \mathcal{A}' \to \mathcal{A} \to \mathcal{A}'' \to 0$$

be an exact sequence of coherent sheaves over Y. Then

$$[\mathcal{A}', g, s] - [\mathcal{A}, g, s] + [\mathcal{A}'', g, s] = 0.$$

This follows from the additivity of length with respect to short exact sequences.

(5) For some particularly interesting cases we use the following notations. Let $F \to E$ be a homomorphism of fields, let X be of finite type over F and let $g\colon Y = X \times_{\operatorname{Spec} F} \operatorname{Spec} E \to X$ be the base change. Then we put $g^* = [\mathcal{O}_Y, g, 0]$.

A morphism $g\colon Y \to X$ of schemes of finite type over a field is said to have (constant) relative dimension s if all fibers are either empty or equidimensional of dimension s. In this case we write $\dim(g) = s$ and put

$$g^* = [\mathcal{O}_Y, g, \dim(g)].$$

Particular cases are here open and closed immersions (with $s = 0$).

(3.6) Multiplication with units. For global units $a_1, \ldots, a_n \in O_X^*$ we define homomorphisms

$$\{a_1, \ldots, a_n\}\colon C_p(X; M) \to C_p(X; M)$$

by

$$\{a_1, \ldots, a_n\}_y^x(\rho) = \begin{cases} \{a_1(x), \ldots, a_n(x)\} \cdot \rho & \text{for } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

This definition turns $C_p(X; M)$ into a module over the tensor algebra of O_X^*. If X is defined over some field F, then $C_p(X; M)$ becomes via $F^* \subset O_X^*$ a module over K_*F.

(3.7) Boundary maps. Let X be of finite type over a field, let $i\colon Y \to X$ be a closed immersion and let $j\colon U = X \setminus Y \to X$ be the inclusion of the open complement. We will refer to (Y, i, X, j, U) as a boundary triple and define

$$\partial = \partial_Y^U\colon C_p(U; M) \to C_{p-1}(Y; M)$$

by taking for ∂_y^x the definition in (2.1.0) with respect to X. The map ∂_Y^U is called the boundary map associated to the boundary triple, or just the boundary map for the closed immersion $i\colon Y \to X$.

We conclude this section with a few notations and remarks concerning the four basic maps.

(3.8) Generalized correspondences. We introduce the notation

$$\alpha\colon X \bullet\!\!\to Y$$

to denote homomorphisms

$$\alpha\colon C_*(X;M) \to C_*(Y;M)$$

which are sums of composites of the four basic maps f_*, g^*, $\{a\}$ and ∂ for schemes of finite type over a field.

This notation is made for the sake of simplification. It also stresses the fact that we think of the maps in question rather as a sort of morphisms of varieties than just maps of complexes associated to every M. As mentioned in the introduction, this can be made more precise in a further development. (The differential d_X is not subject to this notation convention—we rather think of d_X as a part of the inner structure of X. Similarly for homomorphisms induced by a change of coefficients.)

(3.9) Gradings. The $\mathbb{Z}/2$-gradings on M induces a $\mathbb{Z}/2$-grading on $C_*(X;M)$ by

$$C_p(X;M,n) = \coprod_{x \in X_{(p)}} M_{n+p}(x)$$

with $n \in \mathbb{Z}/2$. Suppose $\alpha\colon X \bullet\!\!\to Y$ respects this grading in the sense that

$$\alpha\big(C_*(X;M,n)\big) \subset C_*(Y;M,n+r)$$

for some $r \in \mathbb{Z}/2$. In this case we write $\operatorname{sgn}(\alpha) = (-1)^r$. One has $\operatorname{sgn}(f_*) = \operatorname{sgn}(g^*) = +1$, $\operatorname{sgn}(\{a_1,\ldots,a_n\}) = (-1)^n$ and $\operatorname{sgn}(\partial) = -1$. Moreover we put

$$\delta(\alpha) = d \circ \alpha - \operatorname{sgn}(\alpha) \cdot \alpha \circ d.$$

Then

$$\begin{aligned} \operatorname{sgn}\big(\delta(\alpha)\big) &= -\operatorname{sgn}(\alpha), \\ \delta \circ \delta(\alpha) &= 0, \\ \delta(\alpha \circ \beta) &= \delta(\alpha) \circ \beta + \operatorname{sgn}(\alpha) \cdot \alpha \circ \delta(\beta). \end{aligned}$$

All the maps α to be considered will respect the $\mathbb{Z}/2$-grading. Moreover, if M is $\mathbb{Z}$-graded, then the α will respect the corresponding $\mathbb{Z}$-gradings on the complexes. Additionally they respect the natural $\mathbb{Z}$-gradings given by dimension. So if M is $\mathbb{Z}$-graded, there is an underlying $\mathbb{Z} \times \mathbb{Z}$-grading (see also Section 5). In the general treatment however there is need only for the $\mathbb{Z}/2$-grading.

(3.10) BOUNDARY TRIPLES. Let (Y, i, X, j, U) be a boundary triple. The set theoretic union $X_{(p)} = Y_{(p)} \cup U_{(p)}$ yields a natural decomposition

$$C_p(X; M) = C_p(Y; M) \oplus C_p(U; M) \tag{3.10.1}$$

of abelian groups. Here the complex $C_*(Y; M)$ is a subcomplex of the complex $C_*(X; M)$ with $C_*(U; M)$ as quotient complex. The maps i_*, j_* and i^*, j^* are the corresponding inclusions and projections, respectively. In a formal way, the situation is described by the following formulae:

$$\begin{aligned}
&\partial = i^* \circ d_X \circ j_*,\\
&i^* \circ i_* = \mathrm{id}_Y, \quad j^* \circ j_* = \mathrm{id}_U,\\
&i^* \circ j_* = 0, \quad j^* \circ i_* = 0,\\
&i_* \circ i^* + j_* \circ j^* = \mathrm{id}_X,\\
&\delta(j_*) = i_* \circ \partial, \quad \delta(i_*) = -\partial \circ j^*,\\
&\partial = i^* \circ \delta(j_*) = -\delta(i^*) \circ j_*,\\
&\delta(i_*) = 0, \quad \delta(j^*) = 0, \quad \delta(\partial) = 0.
\end{aligned}$$

Later on we will make free use of these simple rules, in particular in Sections 6 and 9. The canonical decomposition (3.10.1) is the source of our formal treatment of intersection theory on complex level.

4. COMPATIBILITIES

In this section we establish the basic compatibilities for the maps considered in the last section. All arguments are simple in nature or at least familiar to cycle theories. They are basically of local nature. As usual the treatment of flat pull-back causes most of the technicalities.

Rules among the maps of (3.4)–(3.7) are formulated in (4.1)–(4.5). Proposition 4.6 is concerned with the compatibility with the differentials. The compatibilities with change of coefficients are obvious and we don't make a point of them here and further.

(4.1) PROPOSITION.
(1) *For* $f: X \to Y$, $f': Y \to Z$ *as in* (3.4) *one has* $(f' \circ f)_* = f'_* \circ f_*$.
(2) *Let* $g: Y \to X$ *and* $g': Z \to Y$ *be morphisms. Let* $s \geq s(g)$ *and* $s' \geq s(g')$ *and let* $\mathcal{A}$, $\mathcal{A}'$ *be coherent sheaves on* Y, Z, *respectively, with* $\mathcal{A}'$ *flat over* Y. *Then* $s + s' \geq s(g \circ g')$ *and*

$$[g'^*\mathcal{A} \otimes_{\mathcal{O}_Z} \mathcal{A}', g \circ g', s + s'] = [\mathcal{A}', g', s'] \circ [\mathcal{A}, g, s].$$

In particular

$$(g \circ g')^* = g'^* \circ g^*$$

for g, g' *as in* (3.5.5) *with* g' *flat.*

(3) *Consider a pull-back diagram*

$$\begin{array}{ccc} U & \xrightarrow{g'} & Z \\ \downarrow{\scriptstyle f'} & & \downarrow{\scriptstyle f} \\ Y & \xrightarrow{g} & X \end{array}$$

with f and f' as in (3.4). *Let $s \geq s(g)$, $s(g')$ and let $\mathcal{A}$ be a coherent sheaf on Y. Then*

$$[\mathcal{A}, g, s] \circ f_* = f'_* \circ [f'^*\mathcal{A}, g', s].$$

In particular

$$g^* \circ f_* = f'_* \circ g'^*$$

for g as in (3.5.5).

Proof: (1) is immediate from the definitions and R1a.

Proof of (2): The inequality is obvious. Let $x \in X$, $y \in Y_x$, $z \in Z_y$ with $\dim(y, Y) = \dim(x, X) + s$, $\dim(z, Z) = \dim(y, Y) + s'$. We have to check

$$[g'^*\mathcal{A} \otimes_{\mathcal{O}_Z} \mathcal{A}', g \circ g']^x_z = [\mathcal{A}', g']^y_z \cdot [\mathcal{A}, g]^x_y.$$

We may assume $X = \operatorname{Spec} \kappa(x)$ and $Y = \operatorname{Spec} R$ with R as in (3.5.2). By devisage using the flatness of $\mathcal{A}'$ over R and (3.5.4) we may reduce to the case $\mathcal{A} = \kappa(y)$. Now the claim is trivial.

Proof of (3): Let $\delta = [\mathcal{A}, g, s] \circ f_* - f'_* \circ [f'^*\mathcal{A}, g', s]$. We have to show $\delta^z_y = 0$ for $z \in Z_{(p)}$ and $y \in Y_{(p+s)}$.

This obvious if $f(z) \neq g(y)$. Otherwise let $x = f(z) = g(y)$. Our assumptions give $\dim(x, X) \leq \dim(z, Z)$ and $\dim(x, X) \geq \dim(y, Y) - s(g) \geq p$; hence $\dim(x, X) = \dim(z, Z) = p$ and $\kappa(z)$ is finite over $\kappa(x)$.

Let $u \in U_z$ be a maximal point of the fiber over z. Our assumptions give $\dim(u, U) \geq \dim(y, Y) = p + s$ and $\dim(u, U) \leq \dim(z, Z) + s(g') \leq p + s$; hence $u \in U_{(p+s)}$. This shows that δ^z_y remains unchanged if we replace X by $\operatorname{Spec} \kappa(x)$, Z by $\operatorname{Spec} \kappa(z)$, Y by $Y_{x,(y)} = \operatorname{Spec} R$ (see 3.5.2) and p, s by 0. Then f is finite and flat. Hence f' is flat and by devisage using (3.5.4) we may assume $\mathcal{A} = \kappa(y)$ as R-module. But then it suffices to consider the case $Y = \operatorname{Spec} \kappa(y)$ and the claim follows from rule R1c. □

(4.2) Lemma. *Let $f\colon Y \to X$ be as in* (3.4).

(1) *If a is a unit on X, then*

$$f_* \circ \{f^*(a)\} = \{a\} \circ f_*.$$

(2) *Let f be finite and flat and let a be a unit on Y. Then*

$$f_* \circ \{a\} \circ f^* = \{\tilde{f}_*(a)\}.$$

Here $\tilde{f}_*\colon O^*_Y \to O^*_X$ is the standard transfer map.

Proof: (1) is immediate from R2b. For (2) we may assume $X = \operatorname{Spec} F$ with F a field. Then for $y \in Y$ let l_y be the length of $\mathcal{O}_{y,Y}$. By R2c we have

$$f_* \circ \{a\} \circ f^* = \sum_y l_y \cdot c_{\kappa(y)|F}\big(\{a(y)\}\big)$$

and the claim follows. □

(4.3) LEMMA. *Let a be a unit on X.*
(1) *For $g\colon Y \to X$ as in* (3.5.5) *one has*

$$g^* \circ \{a\} = \{g^*a\} \circ g^*.$$

(2) *For a boundary triple (Y, i, X, j, U) one has*

$$\partial^U_Y \circ \{j^*(a)\} = -\{i^*(a)\} \circ \partial^U_Y.$$

Proof: (1) follows from R2a and (2) from R2b and R3e. □

Let $h\colon X \to X'$ be a morphism of schemes of finite type over a field and let $Y' \hookrightarrow X'$ be a closed immersion. Consider the induced diagram given by $U' = X' \setminus Y'$ and pull-back:

(4.4.0)
$$\begin{array}{ccccc} Y & \hookrightarrow & X & \hookleftarrow & U \\ \downarrow{\scriptstyle \bar{h}} & & \downarrow{\scriptstyle h} & & \downarrow{\scriptstyle \bar{\bar{h}}} \\ Y' & \hookrightarrow & X' & \hookleftarrow & U' \end{array}$$

(4.4) PROPOSITION.
(1) *If h is proper, then*

$$\bar{h}_* \circ \partial^U_Y = \partial^{U'}_{Y'} \circ \bar{\bar{h}}_*.$$

(2) *If h is flat* (*of constant relative dimension*), *then*

$$\bar{h}^* \circ \partial^{U'}_{Y'} = \partial^U_Y \circ \bar{\bar{h}}^*.$$

Proof: Immediate from Proposition 4.6.1 and 4.6.2 below. □

(4.5) LEMMA. *Let $g\colon Y \to X$ be a smooth morphism of schemes of finite type over a field of constant fiber dimension* 1, *let $\sigma\colon X \to Y$ be a section to g and let $t \in O_Y$ be a global parameter defining the subscheme $\sigma(X)$. Moreover let $\tilde{g}\colon Y \setminus \sigma(X) \to X$ be the restriction of g and let ∂ be the boundary map associated to σ. Then*

$$\partial \circ \tilde{g}^* = 0 \quad \text{and} \quad \partial \circ \{t\} \circ \tilde{g}^* = (\mathrm{id}_X)_*.$$

Proof: One reduces to $X = \operatorname{Spec} E$ and applies R3c and R3d. □

(4.6) PROPOSITION.
(1) *For proper $f\colon X \to Y$ as in* (3.4) *one has*

$$d_Y \circ f_* = f_* \circ d_X.$$

(2) *Let $g\colon Y \to X$ be a morphism and let $\mathcal{A}$ be a coherent sheaf on Y flat over X. Then*

$$d_Y \circ [\mathcal{A}, g, s] = [\mathcal{A}, g, s] \circ d_X$$

for $s \geq s(g)$. In particular

$$g^* \circ d_X = d_Y \circ g^*$$

for flat g as in (3.5.5).

(3) *For a unit a on X one has*

$$d_X \circ \{a\} = -\{a\} \circ d_X.$$

(4) *For a boundary triple (Y, i, X, j, U) one has*

$$d_Y \circ \partial_Y^U = -\partial_Y^U \circ d_U.$$

Proof: (3) follows as Lemma 4.3.2 and (4) follows from Lemma 3.3.

Proof of (1): Let $\delta(f_*) = d_Y \circ f_* - f_* \circ d_X$. We have to show $\delta(f_*)_y^x = 0$ for $x \in X_{(p)}$ and $y \in Y_{(p-1)}$. Let $z = f(x)$ and $q = \dim(z, Y)$. If $y \notin \overline{\{z\}}$, the claim is obvious. If $y = z$, we first replace Y by $\operatorname{Spec} \kappa(y)$ and then X by $\overline{\{x\}}$. This is the case of a proper curve over a field considered in (RC) of Section 2. If $y \in \overline{\{z\}}$ and $y \neq z$, we must have $q = p$ and $\kappa(x)$ is finite over $\kappa(z)$. We may assume $Y = \overline{\{z\}}$ and $X = \overline{\{x\}}$. Consider the diagram

$$\begin{array}{ccc} \tilde{X} & \xrightarrow{g} & X \\ \Big\downarrow{\scriptstyle \tilde{f}} & & \Big\downarrow{\scriptstyle f} \\ \tilde{Y} & \xrightarrow{h} & Y \end{array}$$

where g and h are the normalizations. Let $\tilde{x} \in \tilde{X}$ and $\tilde{z} \in \tilde{Y}$ be the generic points (lying over x and z, respectively). We have $\delta(g_*) \mid M(\tilde{x}) = 0$ by the very definition of the differentials; similarly $\delta(h_*) \mid M(\tilde{z}) = 0$. This and 4.1.1 show

$$\begin{aligned} \delta(f_*) \circ g_* \mid M(\tilde{x}) &= (d_Y \circ h_* \circ \tilde{f}_* - f_* \circ g_* \circ d_{\tilde{X}}) \mid M(\tilde{x}) \\ &= h_* \circ \delta(\tilde{f}_*) \mid M(\tilde{x}). \end{aligned}$$

Since $g_* \mid M(\tilde{x})$ is an isomorphism onto $M(x)$ we are reduced to show $\delta(\tilde{f}_*)_{\tilde{y}}^{\tilde{x}} = 0$ for $\tilde{y} \in \tilde{Y}_{(p-1)}$. Let $\tilde{u} \in \tilde{X}$ be a point over $\tilde{y}$. We have $\tilde{u} \in \tilde{X}^{(1)}$. Now $\delta(\tilde{f}_*)_{\tilde{y}}^{\tilde{x}} = 0$ follows from rule R3b, the properness of $\tilde{f}$ and the fact that the local rings of $\tilde{y}$ and of all the preimages $\tilde{u}$ are valuation rings.

Proof of (2): Let $\delta = d_Y \circ [\mathcal{A}, g, s] - [\mathcal{A}, g, s] \circ d_X$. We have to show $\delta_y^x = 0$ for $x \in X_{(p)}$, $y \in Y_{(p+s-1)}$. Put $z = g(y)$. If $z \notin \overline{\{x\}}$, the claim is obvious. If $z = x$, then for $u \in Y_x$ all valuations on $\kappa(u)$ with center y are trivial on $\kappa(x)$; the claim follows from rule R3c. We are now reduced to the case $z \in \overline{\{x\}}$, $z \neq x$. Then $z \in X_{(p-1)}$ since $\dim(z, X) \geq \dim(y, Y) - s = p - 1$. We may assume $X = \overline{\{x\}}$. Moreover by Propositions 4.1.3 and 4.6.1 we may additionally assume that X is normal. Let $U = \left\{ u \in Y_x^{(0)} \mid y \in \overline{\{u\}} \right\}$. Then

$$\delta_y^x = \sum_{u \in U} \partial_y^u \circ [\mathcal{A}, g, s]_u^x - [\mathcal{A}, g, s]_y^z \circ \partial_z^x.$$

We may replace X and Y by its localizations in z and y, respectively. Then $X = \operatorname{Spec} R$ with R a valuation ring, $Y = \operatorname{Spec} S$ with S local of dimension ≤ 1 and $U = Y_{(1)}$.

In this case we have by definition

$$\delta^x_y = \sum_{u \in U} l_{S_{(u)}}(\mathcal{A}_{(u)}) \cdot \partial^u_y \circ r_{\kappa(u)|\kappa(x)} - l_S(\mathcal{A}/\pi\mathcal{A}) \cdot r_{\kappa(y)|\kappa(z)} \circ \partial^x_z$$

where $S_{(u)}$, $\mathcal{A}_{(u)}$ are the localizations at u and where π is a prime element of R.

For $u \in U$ let $\tilde{S}_u$ be the normalization of S/u. For a $\tilde{S}_u$-module H of finite length we define

$$L_u(H) = \sum_w l_{\tilde{S}_{(w)}}(H_{(w)}) \cdot [\kappa(w) : \kappa(y)]$$

where w runs through the maximal prime ideals of $\tilde{S}_u$ and where $\tilde{S}_{(w)}$ resp. $H_{(w)}$ are the localizations of $\tilde{S}_u$ resp. H at w. We claim that $L_u(H)$ is the length of H as S-module:

$$L_u(H) = l_S(H).$$

To prove this use devisage to reduce to the trivial case $H = \kappa(w)$ for some w.

Moreover we have

$$l_S(\tilde{S}_u/\pi\tilde{S}_u) = l_S(S/u + \pi S).$$

This follows from the fact that the cokernel and the (trivial) kernel of $S/u \to \tilde{S}_u$ have finite S-length and π is a nonzero divisor of S/u and $\tilde{S}_u$ (see Fulton 1984, Lemma A.2.4).

We have for fixed u:

$$\begin{aligned}
\partial^u_y \circ r_{\kappa(u)|\kappa(x)} &= \sum_w c_{\kappa(w)|\kappa(y)} \circ \partial^u_w \circ r_{\kappa(u)|\kappa(x)} \\
&= \sum_w l_{\tilde{S}_{(w)}}(\tilde{S}_{(w)}/\pi\tilde{S}_{(w)}) \cdot c_{\kappa(w)|\kappa(y)} \circ r_{\kappa(w)|\kappa(z)} \circ \partial^x_z \\
&= L_u(\tilde{S}_u/\pi\tilde{S}_u) \cdot r_{\kappa(y)|\kappa(z)} \circ \partial^x_z.
\end{aligned}$$

Here we used the definition of ∂^u_y and R3a, R1b, R2d.

Putting things together one finds that $\delta^x_y = 0$ follows from

$$l_S(\mathcal{A}/\pi\mathcal{A}) = \sum_{u \in U} l_{S_{(u)}}(\mathcal{A}_{(u)}) \cdot l_S(S/u + \pi S).$$

This formula is exactly the formula of (Fulton 1984, Lemma A.2.7) because the map $\mathcal{A} \to \mathcal{A}$, $a \mapsto \pi a$ is injective by the flatness of $\mathcal{A}$ over R. $\square$

5. Cycle Complexes and Chow Groups

This section contains notations and a few remarks and examples. In Section 3 we have introduced for a cycle module M over X the complexes

$$C_p(X;M)=\coprod_{x\in X_{(p)}} M(x)$$

with differentials

$$d=d_X\colon C_p(X;M)\to C_{p-1}(X;M).$$

Sometimes it is convenient to use the codimension index instead of the dimension index. We put

$$C^p(X;M)=\coprod_{x\in X^{(p)}} M(x)$$

and define

$$d=d_X\colon C^p(X;M)\to C^{p+1}(X;M)$$

again by $d^x_y=\partial^x_y$ with ∂^x_y as in (2.1.0). Similar as in Lemma 3.3 one finds $d\circ d=0$.

The choice between the dimension and codimension index depends on the matter. Our basic interest is in schemes X of finite type over a field F. In this case the dimension setting is in general appropriate, since then the dimension of a point x is an absolute notion independent of the ambient space: $\dim(x,X)=\mathrm{tr.deg}(\kappa(x)|F)$.

If X is in addition equidimensional of dimension d, then $X^{(p)}=X_{(d-p)}$ and $C^p(X;M)=C_{d-p}(X;M)$. Then we will freely switch between the two notions if it is convenient (in particular if we consider intersections in the smooth case). The codimension setting will also be used for certain schemes not necessarily of finite type over a field, e.g., for spectra of local rings. In this case we understand the material of Section 4 to be transferred from the dimension to the codimension setting via finite type models.

In practice, all M have a $\mathbb{Z}$-grading and one likes to keep track on it. We put for $\mathbb{Z}$-graded M

$$C_p(X;M,n)=\coprod_{x\in X_{(p)}} M_{n+p}(x),$$

$$C^p(X;M,n)=\coprod_{x\in X^{(p)}} M_{n-p}(x).$$

with $n\in\mathbb{Z}$. Then there are decompositions of complexes

$$C_*(X;M)=\coprod_n C_*(X;M,n),$$

$$C^*(X;M)=\coprod_n C^*(X;M,n).$$

(In the introduction we have used the notation $C_*(X;n)=C_*(X;K_*,n)$).

The *Chow group of p-dimensional cycles with coefficients in M* is defined as the p-th homology group of the complex $C_*(X;M)$ and denoted by $A_p(X;M)$. Similarly we define $A^p(X;M)$, $A_p(X;M,n)$ and $A^p(X;M,n)$ according to the notations used for the complexes.

The homomorphisms f_* for proper f, g^* for flat g, $\{a_1,\dots,a_n\}$, ∂^U_Y and $\omega_\#$ of Section 3 (anti-)commute with the differentials (see Proposition 4.6). The induced maps on the (co-)homology groups will be denoted by the same letters. The compatibilities of (4.1)–(4.5) carry over (for proper f, f' and flat g).

It is obvious from (3.10) that for a boundary triple (Y,i,X,j,U) there is the long exact sequence

$$\cdots \xrightarrow{\partial} A_p(Y;M) \xrightarrow{i_*} A_p(X;M) \xrightarrow{j_*} A_p(U;M) \xrightarrow{\partial} A_{p-1}(Y;M) \xrightarrow{i_*} \cdots.$$

We conclude by mentioning a few examples. $H^*[D]$ and K'_* denote the cycle modules given by Galois cohomology and Quillen's K-theory as considered in Sections 1–2.

(5.1) REMARK. — *Classical Chow groups.* We understand here $\mathrm{CH}_p(X)$ as the group of p-cycles modulo rational equivalence as defined in (Fulton 1984, Sect. 1.3; denoted by $A_p(X)$). From this definition[◇] it is obvious that

$$A_p(X;K_*,-p) = \mathrm{CH}_p(X).$$

For the Chow group $\mathrm{CH}^p(X)$ of p-codimensional cycles (for smooth irreducible X say) our notations give

$$A^p(X;K_*,p) = \mathrm{CH}^p(X).$$

(5.2) REMARK. — *Unramified cohomology.* For a proper smooth variety X over a field k and a cycle module M over k, the group

$$A^0(X;M) \subset M\big(k(X)\big)$$

is a birational invariant of the field extension $k(X)|k$ (see Corollary 12.10). A well-known example here is the unramified Brauer group of $k(X)|k$. Its n-torsion subgroup is in our notations given by $A^0(X;H^*[\mu_n^{\otimes -1}],2)$.

(5.3) REMARK. — *Relations with local-global spectral sequences.* In Quillen's K-theory as well as in étale cohomology there are spectral sequences given by codimension of support (see Quillen 1973, Sect. 5; Bloch and Ogus 1974). The corresponding E_1-terms together with the d^1-differentials may be identified with the complexes $C^*(X;K'_*,n)$ and $C^*(X;H^*[D],n)$. The corresponding E_2-terms are of the form $E_2^{p,q} = A^p(X;K'_*,-q)$ and $E_2^{p,q} = A^p(X;H^*[D\otimes\mu_r^{\otimes -q}],q)$, respectively (where $r\cdot D = 0$).

[◇] Namely the definition of $\mathrm{CH}_p(X)$ mentioned in the first sentence of the introduction.

(5.4) Remark. — *The map from Milnor's to Quillen's K-theory.* The natural homomorphisms $K_*F \to K'_*F$ form a homomorphism of cycle modules. It is an isomorphism in degrees ≤ 2. Moreover for a valuation v on F one has $\partial_v(K_3F) = \partial_v(K'_3F)$, see (Merkurjev and Suslin 1987). It follows that the induced homomorphisms

$$A_p(X; K_*, n) \to A_p(X; K'_*, n)$$

are bijective for $n + p \leq 2$.

6. Acyclicity for Smooth Local Rings

The following observations have been included to underpin the notion of cycle modules. They are not needed in further sections. M is a cycle module over a field k.

(6.1) Theorem. *Let X be smooth and semi-local. Then*

$$A^p(X; M) = 0 \quad \textit{for} \quad p > 0.$$

This theorem is known in Quillen's K-theory (Quillen 1973, § 7, Theorem 5.11) and in étale cohomology (Bloch and Ogus 1974) and has been proved by O. Gabber for Milnor's K-theory. The main step in these proofs is sometimes called "Quillen's trick" and carries over to cycle modules as well. Here we follow essentially this method but with a simplification due to I. Panin.

Let V be a vector space over k and let $\mathbb{A}(V)$ be the associated affine space. For a linear subspace W of V let

$$\pi_W \colon \mathbb{A}(V) \to \mathbb{A}(V/W),$$
$$\pi_W(v) = v + W$$

be the projection.

(6.2) Lemma. *Let $X \subset \mathbb{A}(V)$ be an equidimensional closed subvariety with* $\dim X = d$ *and let $Y \subset X$ be a closed subvariety with* $\dim Y < d$. *Moreover let $S \subset Y$ be a finite subset such that X is smooth in S. Then for a generic $(d-1)$-codimensional linear subspace W of V the following conditions hold.*

(1) *The restriction*

$$\pi_W \mid Y : Y \to \mathbb{A}(V/W)$$

is finite.

(2) *The restriction*

$$\pi_W \mid X : X \to \mathbb{A}(V/W)$$

is locally around S smooth of relative dimension 1.

Proof: (Panin) We extend the situation to the projective closure $\mathbb{A}(V) \subset \mathbb{P}(V \oplus k)$ with $\mathbb{P}(V) \subset \mathbb{P}(V \oplus k)$ as hyperplane at infinity. Let

$$\bar{\pi}_W \colon \mathbb{P}(V \oplus k) \setminus \mathbb{P}(W) \to \mathbb{P}(V/W \oplus k),$$
$$\bar{\pi}_W([v,t]) = [v + W, t]$$

be the projection. $\bar{\pi}_W$ is an affine bundle over $\mathbb{P}(V/W \oplus k)$ which extends the affine bundle π_W over $\mathbb{A}(V/W)$.

Let $\overline{Y} \subset \mathbb{P}(V \oplus k)$ be the closure of Y and let $Y_\infty = \overline{Y} \cap \mathbb{P}(V)$. Then $\dim Y_\infty < d-1$. Hence for generic $(d-1)$-codimensional W we have $Y_\infty \cap \mathbb{P}(W) = \varnothing$. Therefore there is the map

$$\bar{\pi}_W \mid \overline{Y} \colon \overline{Y} \to \mathbb{P}(V/W \oplus k).$$

This map is finite since it is proper and since $\bar{\pi}_W$ is an affine bundle. This shows (1) because $\pi_W \mid Y$ is the pull-back of $\bar{\pi}_W \mid \overline{Y}$ along $\mathbb{A}(V) \subset \mathbb{P}(V \oplus k)$.

For condition (2) one just needs to guarantee that π_W maps for each $s \in S$ the tangent space $T_s X \subset V$ of X in s epimorphically onto the tangent space V/W of $\mathbb{A}(V/W)$ in $\pi_W(s)$. This is again an open condition for W. □

This lemma is very close to (Quillen 1973, § 7, Lemma 5.12.) and suffices for all applications I know. The existence of such a space W is not clear over finite ground fields and needs some extra discussion. However, if one is in the end interested in (co-)homology groups, there is usually no problem with replacing the ground field k by a rational extension $k(t_1, \ldots, t_r)$. In this case one may take for W for example the tautological subspace of V defined over the function field of the Grassmannian of $(d-1)$-codimensional subspaces of V. In our situation we refer here to the following remark.

(6.3) Lemma. *Let X be a variety over k and let $g \colon X_{k(t)} \to X$ be the base change. Then*

$$g^* \colon A_q(X; M) \to A_q(X_{k(t)}; M)$$

is injective.

This lemma will be become obvious in the next section where we show that $F \to A_q(X_F; M)$ is a cycle module. Then g^* is just the restriction map $r_{k(t)|k}$ for this cycle module and any specialization s_v^π at a rational point of $\mathbb{P}^1$ yields a left inverse. (What one really uses here is Lemma 4.5 with $Y = X \times \mathbb{P}^1$ and Proposition 4.6.2.)

(6.4) Proposition. *Let X be a smooth variety over a field and let $Y \subset X$ be a closed subscheme of codimension ≥ 1. Then for any finite subset $S \subset Y$ there is an open neighborhood X' of S in X such that the map*

$$i_* \colon A_*(Y \cap X'; M) \to A_*(X'; M)$$

is the trivial map. Here $i \colon Y \cap X' \to X'$ is the inclusion.

Proof: We may assume that X is affine. By Lemma 6.2 we find a diagram (at least after replacing k by a rational extension)

$$\begin{array}{ccc} Y & \xrightarrow{i} & X \\ & \searrow \quad \swarrow & \\ & A & \end{array}$$

with $Y \to A$ finite and with $X \to A$ smooth of relative dimension 1 in S.

Put $Z = Y \times_A X$ and consider the diagram

$$\begin{array}{ccc} & Z & \\ {}^{g}\swarrow\!\!\nearrow_{\sigma} & & \searrow^{\pi} \\ Y & \xrightarrow{i} & X \end{array}$$

where g and π are the projections and σ is the diagonal. Note that π is finite, g is smooth of relative dimension 1 in S and that σ is a section to g and a lift of the immersion i. Moreover after a localization to an open subset $X' \subset X$ containing S we may assume that there is a global parameter $t \in O_Z$ defining the closed subscheme $\sigma(Y)$.

Let (Y, σ, Z, j, Q) be the boundary triple given by σ (with $Q = Z \setminus \sigma(Y)$) and let $\tilde{g}$: $Q \to Y$ be the restriction of g. Now consider the composite

$$H \colon Y \overset{\tilde{g}^*}{\bullet\!\!\longrightarrow} Q \overset{\{t\}}{\bullet\!\!\longrightarrow} Q \overset{j_*}{\bullet\!\!\longrightarrow} Z \overset{\pi_*}{\bullet\!\!\longrightarrow} X.$$

One finds

$$\delta(H) = \pi_* \circ \sigma_* \circ \partial^Q_Y \circ \{t\} \circ \tilde{g}^* = \pi_* \circ \sigma_* = i_*$$

by Lemma 4.5. Therefore i_* is nullhomotopic. $\square$

Proof of Theorem 6.1: We may assume that X is connected. Put $d = \dim X$. Consider pairs (U, S) where U is a smooth d-dimensional variety of finite type over k and $S \subset U$ is a finite subset such that X is the localization of U in S. Then

$$C^p(X; M) = \varinjlim_{(U,S)} C^p(U; M).$$

Moreover

$$C^p(U; M) = C_{d-p}(U; M) = \varinjlim_{Y} C_{d-p}(Y; M)$$

where Y runs over the closed p-codimensional subsets of U. Hence

$$A^p(X; M) = \varinjlim_{(U,S)} A^p(U; M) = \varinjlim_{(U,S)} \varinjlim_{Y} A_{d-p}(Y; M).$$

But Proposition 6.4 tells that $A_{d-p}(Y; M) \to A^p(U; M) \to A^p(U'; M)$ is the trivial map for small enough $U' \subset U$. $\square$

In the smooth case we sheafify cycle modules as follows. For a smooth variety X let $\mathcal{M}_X$ be the Zariski sheaf on X given by

$$\mathcal{M}_X(U) = A^0(U; M) \subset M(\xi_X)$$

for open subsets U of X.

(6.5) COROLLARY. *For a smooth variety X over k there are natural isomorphisms*

$$A^p(X; M) = H^p_{\mathrm{Zar}}(X; \mathcal{M}_X).$$

Proof: Define the Zariski sheaves $\mathcal{C}^p$ on X by

$$\mathcal{C}^p(U) = C^p(U; M).$$

Then there is a complex of sheaves

$$0 \longrightarrow \mathcal{M}_X \longrightarrow \mathcal{C}^0 \xrightarrow{d} \mathcal{C}^1 \xrightarrow{d} \cdots .$$

The complex is exact. This holds at $\mathcal{M}_X$ and at $\mathcal{C}^0$ by the very definitions. Theorem 6.1 implies exactness at positive dimensions. The corollary follows, since the $\mathcal{C}^p$ are flasque. □

The resolution of $\mathcal{M}_X$ considered in this proof has nice functorial properties. Namely, we will define for morphisms $f\colon Y \to X$ maps of complexes (Section 12)

$$I(f)\colon C^p(X; M) \to C^p(Y; M)$$

and, under presence of a ring structure for M, a pairing of complexes (Section 14)

$$C^*(X; M) \times C^*(X; M) \to C^*(X; M).$$

These are functorial with respect to localizations. Therefore the isomorphisms of Corollary 6.5 are compatible with pull-backs and with products.

The following example is a nice illustration of Corollary 6.5. Let X be smooth and define the Zariski sheaf $\mathcal{K}_n$ on X by

$$\mathcal{K}_n(U) = A^0(U; K_*, n) \subset K_n k(X).$$

The sheaf $\mathcal{K}_n$ has a comparatively simple definition: it just refers to the definition of Milnor's K-groups for fields and of the residue maps for valuations. Corollary 6.5 yields the following interpretation of the classical Chow groups on a smooth variety:

$$\mathrm{CH}^p(X) = H^p_{\mathrm{Zar}}(X; \mathcal{K}_p). \tag{6.6}$$

The same result holds with Milnor's K-theory replaced by Quillen's K-theory. The corresponding sheaf $\mathcal{K}'_n$ coincides with the sheaf induced from the presheaf $U \to K^Q_n(U)$ where $K^Q_n(U)$ denotes the n-th Quillen's K-group of the category of vector bundles on U. In this context (6.6) is known as Bloch's formula (see Quillen 1973, Thm. 5.19; Grayson 1978).

Another special case of Corollary 6.5 for $M = K_*$ is

$$A_0(X; K_*, n) = H^d_{\mathrm{Zar}}(X; \mathcal{K}_{n+d}),$$

with $d = \dim X$. This interpretation of the "Chow groups of zero cycles on X with coefficients in K_n" was obtained already in (Kato 1986).

7. The Cycle Modules $A_q[\rho; M]$

In this section we show that new cycle modules can be obtained from the Chow groups of the fibers of a morphism. It was in fact this process of forming local coefficient systems for cycles which motivated the notion of cycle modules.

Let $\rho: Q \to B$ be a morphism of finite type and let M be a cycle module over Q. For fields F over B let $Q_F = Q \times_B \operatorname{Spec} F$. We define an object function $A_q[\rho, M]$ on $\mathcal{F}(B)$ by

$$A_q[\rho; M](F) = A_q(Q_F; M).$$

Our aim is to show that $A_q[\rho, M]$ is in a natural way a cycle module over B.

All the properties of cycle modules except axiom (C) hold already on complex level, i.e., for the groups $C_q(Q_F; M)$. It is appropriate to establish first the corresponding object function as a cycle premodule.

So let $\widehat{M}$ be the object function on $\mathcal{F}(B)$ defined by

$$\widehat{M}(F) = C_q(Q_F; M).$$

We first describe its data as a cycle premodule. These will be denoted by $\widehat{\varphi}_*$, $\widehat{\varphi}^*$, $\widehat{\partial}_v$, $\widehat{r}_{E|F}$, $\widehat{c}_{E|F}$, etc. in order to distinguish them from the data φ_*, φ^*, ∂_v, etc. of M.

For a homomorphism of fields $\varphi: F \to E$ let $\bar{\varphi}: Q_E \to Q_F$ be the induced morphism. We define the data D1 and D2 by

$$\widehat{\varphi}_* = \bar{\varphi}^*: C_q(Q_F; M) \to C_q(Q_E; M),$$
$$\widehat{\varphi}^* = \bar{\varphi}_*: C_q(Q_E; M) \to C_q(Q_F; M).$$

For D3 we take the K_*F-module structure on $C_q(Q_F; M)$ described in (3.6). To establish D4 put $\widetilde{Q}_v = Q \times_B \operatorname{Spec} \mathcal{O}_v$. It has over $\operatorname{Spec} \mathcal{O}_v$ the generic fiber Q_F and the special fiber $Q_{\kappa(v)}$. Define

$$\widehat{\partial}_v: C_q(Q_F; M) \to C_q(Q_{\kappa(v)}; M)$$

by $(\widehat{\partial}_v)^x_y = \partial^x_y$ with ∂^x_y as in (2.1.0) with respect to the scheme $\widetilde{Q}_v$.

(7.1) Theorem. *Together with these data, $\widehat{M}$ is a cycle premodule over B.*

Proof: All the required properties follow from the rules and axioms for M and from Section 4.

Below we consider R3a in detail. Here is a sketch for the other (less complicated) cases:
for R1a use (4.1.2); for R1b use (4.1.1);
for R1c use (4.1.3) and a length consideration;
for R2a use (4.3.1); for R2b use (4.2.1); for R2c use R1c and R2c;
for R3b use (4.6.1); for R3c use R3c;
for R3d use (1.5) and R3d; for R3e use R2b and R3e.

Proof of R3a: Let $g_\xi\colon Q_F \to Q_E$ and $g_0\colon Q_{\kappa(v)} \to Q_{\kappa(w)}$ be the projections. We have to show that the following diagram is commutative:

$$\begin{array}{ccc} C_q(Q_F; M) & \xrightarrow{\widehat{\partial}_v} & C_q(Q_{\kappa(v)}; M) \\ \Big\uparrow{\scriptstyle g_\xi^*} & & \Big\uparrow{\scriptstyle e\cdot g_0^*} \\ C_q(Q_E; M) & \xrightarrow{\widehat{\partial}_w} & C_q(Q_{\kappa(w)}; M). \end{array}$$

We want to apply Proposition 4.6.2 to the projection $g\colon \widetilde{Q}_v \to \widetilde{Q}_w$. The pull-back of g along $\operatorname{Spec} E \to \operatorname{Spec} \mathcal{O}_w$ is g_ξ. Let

$$\bar{g}_0\colon \bar{Q}_{\kappa(v)} = \widetilde{Q}_v \times_{\operatorname{Spec} \mathcal{O}_w} \operatorname{Spec} \kappa(w) \to Q_{\kappa(w)}$$

be the pull-back of g along $\operatorname{Spec} \kappa(w) \to \operatorname{Spec} \mathcal{O}_w$. Note that $\bar{Q}_{\kappa(v)}$ and $Q_{\kappa(v)}$ have the same reductions and therefore the same cycle groups.

We claim $\bar{g}_0^* = e \cdot g_0^*$. Let $R = \mathcal{O}_v \otimes_{\mathcal{O}_w} \kappa(w)$. Note that g_0, $\bar{g}_0$ are the pull-backs along $Q_{\kappa(w)} \to \operatorname{Spec} \kappa(w)$ of the morphisms $\operatorname{Spec} \kappa(v) \to \operatorname{Spec} \kappa(w)$, $\operatorname{Spec} R \to \operatorname{Spec} \kappa(w)$, respectively. The claim follows from $e = l_R(R)$ and a standard length consideration.

It remains to show that the diagram commutes with $e \cdot g_0^*$ replaced by $\bar{g}_0^*$. This follows (cum grano salis, see the following remark) from Proposition 4.6.2.

Remark. When applying here Proposition 4.6 in a formal way, there appears an artificial problem caused by the fact that the dimension index does not behave perfectly well for schemes over local rings like $\widetilde{Q}_v$. However, note that to check a commutativity like $\partial_v \circ g_\xi^* = \bar{g}_0^* \circ \partial_w$ it suffices to restrict to the components corresponding to points $x \in Q_{E(q)}$ with $\overline{\{x\}} \cap \big(Q_{\kappa(w)}\big)_{(q)} \neq \varnothing$. For these points one has $x \in (\widetilde{Q}_w)_{(q+1)}$ by the dimension inequality (Matsumura 1980, p. 85). A similar remark applies to $\widetilde{Q}_v$. Therefore the desired identity follows from $d \circ g^* = g^* \circ d$ on $C_{q+1}(\widetilde{Q}_w; M)$. One may avoid these considerations by looking more closely to the proof of Proposition 4.6. □

We have to relate the differentials for the cycle premodule $\widehat{M}$ to the differentials for the cycle module M.

Let $X \to B$ be a scheme over B and let $\widetilde{X} = Q \times_B X$. Then for x, $y \in X$ there is the map

$$\widehat{\partial}^x_y\colon \widehat{M}(x) \to \widehat{M}(y)$$

according to (2.1.0). By definition this is a map

$$\widehat{\partial}^x_y\colon C_q(Q_{\kappa(x)}; M) \to C_q(Q_{\kappa(y)}; M)$$

between cycle groups with coefficients in M.

(7.2) PROPOSITION. *Let $\tilde{x}$, $\tilde{y} \in \widetilde{X}$ be points lying over x, $y \in X$, respectively, and suppose $\tilde{x} \in \big(Q_{\kappa(x)}\big)_{(q)}$ and $\tilde{y} \in \big(Q_{\kappa(y)}\big)_{(q)}$. Denote by $(\widehat{\partial}^x_y)^{\tilde{x}}_{\tilde{y}}$ the component of $\widehat{\partial}^x_y$ with respect to $\tilde{x}$ and $\tilde{y}$. Then*

$$(\widehat{\partial}^x_y)^{\tilde{x}}_{\tilde{y}} = \partial^{\tilde{x}}_{\tilde{y}}\colon M(\tilde{x}) \to M(\tilde{y}).$$

Proof: We may assume $\tilde{y} \in \overline{\{\tilde{x}\}}^{(1)}$, since otherwise both sides are trivial. The dimension inequality (Matsumura 1980, p. 85) shows then $y \in \overline{\{x\}}^{(1)}$. Let v run through the valuations of $\kappa(x)$ with center y in X. Moreover let w run through the valuations on $\kappa(\tilde{x})$ with center $\tilde{y}$ in $\tilde{X}$. The restriction of any w to $\kappa(x)$ is one of the valuations v. Let $\tilde{w} \in Q_{\kappa(v)}$ be the center of w in $\widetilde{X} \times_X \operatorname{Spec} \mathcal{O}_v$. Now the claim follows from

$$\begin{aligned}
(\widehat{\partial}^x_y)^{\tilde{x}}_{\tilde{y}} &= (\sum_v \widehat{c}_{\kappa(v)|\kappa(y)} \circ \widehat{\partial}_v)^{\tilde{x}}_{\tilde{y}} \\
&= \sum_v \sum_{w|v} (\widehat{c}_{\kappa(v)|\kappa(y)})^{\tilde{w}}_{\tilde{y}} \circ (\widehat{\partial}_v)^{\tilde{x}}_{\tilde{w}} \\
&= \sum_v \sum_{w|v} c_{\kappa(\tilde{w})|\kappa(\tilde{y})} \circ c_{\kappa(w)|\kappa(\tilde{w})} \circ \partial_w \\
&= \sum_w c_{\kappa(w)|\kappa(\tilde{y})} \circ \partial_w = \partial^{\tilde{x}}_{\tilde{y}}. \qquad \square
\end{aligned}$$

It follows from Proposition 4.6 that the data of the cycle premodules $\widehat{M}$ (for various q) commute resp. anti-commute with the differentials of the complexes $C_*(Q_F; M)$. Passing to homology we obtain data D1–D4 for the object functions $A_q[\rho; M]$.

(7.3) Theorem. *Together with these data, $A_q[\rho; M]$ is a cycle module over B.*

Proof: The rules for the data of the cycle premodule $A_q[\rho; M]$ are immediate from the rules for $\widehat{M}$. Moreover axiom (FD) for M and Proposition 7.2 show that (FD) holds for $\widehat{M}$—consequently also for $A_q[\rho; M]$. It remains to verify axiom (C).

Consider the maps[◇]

$$C_q(Q_{\kappa(\xi)}) \xrightarrow{\Theta} C_{q-1}(Q_{\kappa(\xi)}) \oplus \coprod_{x \in X^{(1)}} C_q(Q_{\kappa(x)}) \oplus C_{q+1}(Q_{\kappa(x_0)}) \xrightarrow{\Theta} C_q(Q_{\kappa(x_0)})$$

defined by $\Theta^z_y = \partial^z_y$ with ∂^z_y as in (2.1.0) with respect to the scheme $Q \times_B X$.

By Proposition 7.2 we are reduced to show $\Theta \circ \Theta = 0$. It suffices to check $(\Theta \circ \Theta)^z_y = 0$ for $z \in \big(Q_{\kappa(\xi)}\big)_{(q)}$ and $y \in \big(Q_{\kappa(x_0)}\big)_{(q)}$ with $y \in \overline{\{z\}}^{(2)}$ (here $\overline{\{z\}}$ is the closure of z in $\widetilde{X}$). The dimension inequality (Matsumura 1980, p. 85) shows

$$Z^{(1)} \subset \big(Q_{\kappa(\xi)}\big)_{(q-1)} \cup \bigcup_x \big(Q_{\kappa(x)}\big)_{(q)} \cup \big(Q_{\kappa(x_0)}\big)_{(q+1)}$$

with $Z = \overline{\{z\}}_{(y)}$. We are done by axiom (C) for M. $\square$

In the following proposition we formulate some functorial properties of the construction $\rho \to A_q[\rho; M]$. Let

(7.4.0)
$$\begin{tikzcd} Y \arrow[rr, "h"] \arrow[dr, "\eta"'] & & X \arrow[dl, "\rho"] \\ & B & \end{tikzcd}$$

◇ Here the indication of the cycle module M has been dropped.

be a commutative diagram with η and ρ of finite type and let M be a cycle module over X. For a field F over B let

$$h_F\colon Y_F \to X_F$$

be the morphism induced by h.

(7.4) PROPOSITION. *The following transformations are homomorphisms of cycle modules over B:*

(1) *For proper h let*

$$[h_*]\colon A_q[\eta; M] \to A_q[\rho; M]$$

with $[h_*]_F = (h_F)_*$.

(2) *For flat h of relative dimension s let*

$$[h^*]\colon A_q[\rho; M] \to A_{q+s}[\eta; M]$$

with $[h^*]_F = [h_F^*]$.

(3) *For a global unit a on X let*

$$[\{a\}]\colon A_q[\rho; M] \to A_q[\rho; M]$$

with $[\{a\}]_F = \{a \mid X_F\}$.

(4) *For a boundary triple (Y, i, X, j, U) let*

$$[\partial] = [\partial_Y^U]\colon A_q[\rho \circ j; M] \to A_{q-1}[\rho \circ i; M]$$

with $[\partial]_F$ the boundary map for $Y_F \to X_F$.

Proof: One has to check the compatibility with D1–D4. This follows for (1) from (4.1.3), (4.1.1), (4.2.1) and (4.6.1); for (2) from (4.1.2), (4.1.3), (4.3.1) and (4.6.2); for (3) from (4.2.1), (4.3.1), the anti-commutativity of K_* and (4.6.3); for (4) from (4.6.1), (4.6.2), (4.6.3) and (C). □

Let $\rho\colon Q \to B$ be flat and not necessarily of finite type. One may then define cycle modules $A^q[\rho; M]$ with

$$A^q[\rho; M](F) = A^q(Q_F; M).$$

To establish these cycle modules one proceeds analogous to the $A_q[\rho; M]$ above. Alternatively one may reduce to the consideration of the $A_q[\rho; M]$ as follows. If ρ is of finite type, one may assume that it is of constant dimension s. In this case one has $A^q[\rho; M] = A_{s-q}[\rho; M]$. For the general case note that (at least locally with respect to B) one has $A^q(Q_F; M) = \varinjlim A^q(Q'_F; M)$ where $\rho'\colon Q' \to B$ runs through the flat finite type models of ρ.

8. Fibrations

In this section we consider the spectral sequence associated to a morphism and formulate some basic functorial properties. A first application yields the homotopy property for vector bundles.

From now on all schemes are assumed to be of finite type over a field and M is (with exceptions in Section 14) a cycle module over X.

For a morphism ρ: $X \to X'$ we put

$$C_{p,l}(\rho) = \coprod_{x \in X_{(p,l)}} M(x) \subset C_l(X; M)$$

where

$$X_{(p,l)} = \left\{ x \in X_{(l)} \mid \dim\bigl(\rho(x), X'\bigr) \leq p \right\}.$$

Then

$$\cdots \subset C_{p-1,*}(\rho) \subset C_{p,*}(\rho) \subset \cdots \subset C_*(X; M)$$

is a finite filtration of $C_*(X; M)$ by subcomplexes. This filtration has the subquotients

$$\coprod_{u \in X'_{(p)}} C_*(X_{\kappa(u)}; M).$$

Let $\bigl(E^n_{p,q}(\rho)\bigr)_n$ be the associated spectral sequence (see e.g. Hilton and Stammbach 1971). The differential for X restricts on $C_*(X_{\kappa(u)}; M)$ to the differential for $X_{\kappa(u)}$. Therefore

$$E^1_{p,q}(\rho) = \coprod_{u \in X'_{(p)}} A_q(X_{\kappa(u)}; M).$$

(8.1) Proposition. *The differential $d^1_{p,q}$ of this spectral sequence equals the differential $d_{X'}$ for the cycle module $A_q[\rho; M]$.*

Proof: For $u' \in X'_{(p)}$, $y' \in X'_{(p-1)}$ we have to check equality of the corresponding components of $d^1_{p,q}$ and $d_{X'}$:

$$\bigl(d^1_{p,q}\bigr)^{u'}_{y'} = (d_{X'})^{u'}_{y'} : A_q(X_{\kappa(u')}; M) \to A_q(X_{\kappa(y')}; M).$$

The map $(d^1_{p,q})^{u'}_{y'}$ is by definition induced from the map

$$\Theta\colon C_q(X_{\kappa(u')}; M) \to C_q(X_{\kappa(y')}; M)$$

where $\Theta^u_y = \partial^u_y$ for u, $y \in X$ lying over u', y', respectively.
The claim follows from Proposition 7.2. □

(8.2) COROLLARY. *There is a convergent spectral sequence*

$$E^2_{p,q}(\rho) = A_p(X'; A_q[\rho; M]) \Longrightarrow A_{p+q}(X; M).$$

If X' is equidimensional and ρ is flat, then there is a convergent spectral sequence

$$E_2^{p,q}(\rho) = A^p(X'; A^q[\rho; M]) \Longrightarrow A^{p+q}(X; M). \qquad \square$$

Here the second statement follows from the first by a formal switch to codimension index. In this codimension setting one may drop the finite type hypotheses.

(8.3) REMARK. We will use the following dictions. Let $\rho\colon X \to X'$, $\eta\colon Y \to Y'$ be morphisms. Then $\alpha\colon X \bullet\!\!\to Y$ is called filtration preserving (with respect to ρ, η) of degree (r,t), if

$$\alpha\big(C_{p,l}(\rho)\big) \subset C_{p+r,l+t}(\eta).$$

If $\delta(\alpha) = 0$ (see Sec. 3 for the definition of δ), then α is homomorphism of filtered complexes and induces maps (denoted by the same letter)

$$\alpha\colon E^n_{p,q}(\rho) \to E^n_{p+r,q+t-r}(\eta).$$

Two filtration preserving maps $\alpha, \beta\colon X \bullet\!\!\to Y$ of degree (r,t) are called homotopic, $\alpha \simeq \beta$, if there is a filtration preserving $H\colon X \bullet\!\!\to Y$ of degree $(r+1, t+1)$ such that $\alpha - \beta = \delta(H)$. If $\delta(\alpha) = \delta(\beta) = 0$ and $\alpha \simeq \beta$, then the induced maps on the E^2-terms coincide. If the homotopy H can be chosen of degree $(r, t+1)$, then already the induced maps on the E^1-terms coincide. This follows from a little calculation working for arbitrary filtered complexes.

Let

$$\begin{array}{ccc} Y & \xrightarrow{f} & X \\ \big\downarrow{\scriptstyle \eta} & & \big\downarrow{\scriptstyle \rho} \\ Y' & \xrightarrow{f'} & X' \end{array} \tag{8.4.0}$$

be a commutative diagram of morphisms. The following statement is trivial.

(8.4) LEMMA.

(1) *One has*

$$f_*\big(C_{p,l}(\eta)\big) \subset C_{p,l}(\rho).$$

(2) *Suppose f has relative dimension t and let $s \geq s(f')$, see* (3.5.1). *Then*

$$f^*\big(C_{p,l}(\rho)\big) \subset C_{p+s,l+t}(\eta).$$

(3) *If a is a unit on X, then*

$$\{a\} \cdot \big(C_{p,l}(\rho)\big) \subset C_{p,l}(\rho).$$

(4) *Let* (Y, i, X, j, U) *be a boundary triple. Then*

$$\partial_Y^U\bigl(C_{p,l}(\rho\circ j)\bigr) \subset C_{p,l-1}(\rho\circ i).$$

(5) *For the diagram* (4.4.0) *one has*

$$\partial_Y^U\bigl(C_{p,l}(\bar{\bar{h}})\bigr) \subset C_{p-1,l-1}(\bar{h}). \qquad \square$$

Let

$$\begin{array}{ccccc} Y & \xrightarrow{\hat f} & \widehat{Y} & \xrightarrow{\bar f} & X \\ \downarrow{\scriptstyle \eta} & & \downarrow{\scriptstyle \hat\rho} & & \downarrow{\scriptstyle \rho} \\ Y' & = & Y' & \xrightarrow{f'} & X' \end{array} \tag{8.5.0}$$

be the natural decomposition of diagram (8.4.0) with $\widehat{Y} = Y' \times_{X'} X$ and $f = \bar f \circ \hat f$. We call the diagram (8.4.0) a flat square if $\hat f$ and f' are flat of some constant relative dimensions. This holds then also for f.

We use the natural identification

$$A_q[\hat\rho; (\bar f)^* M] = (f')^* A_q[\rho; M]$$

of cycle modules over Y'.

(8.5) Proposition.

(1) *If* f *and* f' *are proper, then the map*

$$f_*\colon E^2_{p,q}(\eta) \to E^2_{p,q}(\rho)$$

corresponding to (8.4.1) *equals the composite*

$$A_p(Y'; A_q[\eta; M]) \xrightarrow{[\hat f_*]_\#} A_p(Y'; A_q[\rho; M]) \xrightarrow{f'_*} A_p(X'; A_q[\rho; M]).$$

(2) *Suppose the square* (8.4.0) *is flat and put* $r = \dim(f')$, $s = \dim(\hat f)$. *Then the map*

$$f^*\colon E^2_{p,q}(\rho) \to E^2_{p+r,q+s}(\eta)$$

corresponding to (8.4.2) *equals the composite*

$$A_p(X'; A_q[\rho; M]) \xrightarrow{(f')^*} A_{p+r}(Y'; A_q[\rho; M]) \xrightarrow{[\hat f^*]_\#} A_{p+r}(Y'; A_{q+s}[\eta; M]).$$

(3) *For a global unit* a *on* X *the map*

$$\{a\}\colon E^2_{p,q}(\rho) \to E^2_{p,q}(\rho)$$

corresponding to (8.4.3) *equals* $[\{a\}]_\#$.

(4) *The map*
$$\partial\colon E^2_{p,q}(\rho\circ j)\to E^2_{p,q-1}(\rho\circ i)$$
corresponding to (8.4.4) *equals* $[\partial^U_Y]_\#$.
(5) *The map*
$$\partial\colon E^2_{p,q}(\bar{\bar h})\to E^2_{p-1,q}(\bar h)$$
corresponding to (8.4.5) *equals the map*
$$\partial^{U'}_{Y'}\colon A_p(U';A_q[h;M])\to A_{p-1}(Y';A_q[h;M]).$$

Proof: (3) is trivial. (5) follows from Proposition 8.1. In (1) and (2) one may suppose either $f=\hat f$ or $f=\bar f$.

Proof of (1) for $f=\hat f$: Here $X'=Y'$ and the map
$$f_*\colon C_l(Y;M)\to C_l(X;M)$$
is the family of maps
$$(f_{\kappa(u)})_*\colon C_q(Y_{\kappa(u)};M)\to C_q(X_{\kappa(u)};M)$$
with $u\in X'_{(p)}$ and $p+q=l$. On the other hand
$$[\hat f_*]_\#\colon C_p(X';A_q[\eta;M])\to C_p(X';A_q[\rho;M])$$
is componentwise induced by the maps $(f_{\kappa(u)})_*$.

Proof of (1) for $f=\bar f$: Here we have a pull-back diagram, $Y=Y'\times_{X'}X$. We consider the maps induced by
$$f_*\colon C_l(Y;M)\to C_l(X;M)$$
on the E^1-terms. These are maps (with $p+q=l$)
$$\coprod_{y'\in Y'_{(p)}}C_q(X_{\kappa(y')};M)\longrightarrow\coprod_{x'\in X'_{(p)}}C_q(X_{\kappa(x')};M).$$
Their components are the corestrictions $c_{\kappa(y)|\kappa(x)}$ with $y\in\big(X_{\kappa(y')}\big)_{(q)}$, $y'\in Y'_{(p)}$ and $f(y)\in\big(X_{\kappa(x')}\big)_{(q)}$, $x'=f'(y')\in X_{(p)}$. Here $\kappa(y')$ is necessarily finite over $\kappa(x')$, since both fields have the same transcendence degree. Therefore the maps on the E^1-terms are given by the maps
$$(f'_{y'})_*\colon C_q(X_{\kappa(y')};M)\to C_q(X_{\kappa(f'(y'))};M)$$
with $y'\in Y'_{(p)}$ such that $\kappa(y')|\kappa(f'(y'))$ is finite and where
$$f'_{y'}\colon X_{\kappa(y')}\to X_{\kappa(f'(y'))}$$
is the associated finite morphism. On the other hand
$$f'_*\colon C_p(Y';A_q[\rho;M])\to C_p(X';A_q[\rho;M])$$
is induced exactly by the maps $f'_{y'}$.

Proof of (2) for $f = \hat{f}$: One argues as for (1) and notes that

$$f^*\colon C_l(X;M) \to C_{l+s}(Y;M)$$

is the family of maps

$$(f_{\kappa(u)})^*\colon C_q(X_{\kappa(u)};M) \to C_{q+s}(Y_{\kappa(u)};M)$$

with $u \in X'_{(p)}$ and $p+q=l$.

Proof of (2) for $f = \bar{f}$: The map

$$f^*\colon C_l(X;M) \to C_{l+r}(Y;M)$$

is the family of maps

$$[\mathcal{O}_Y, f]_y^x \cdot r_{\kappa(y)|\kappa(x)}\colon M(x) \to M(y)$$

with $y \in Y_{(l+r)}$, $x \in X_{(l)}$ and $f(y) = x$.

The map

$$(f')^*\colon C_p(X';A_q[\rho;M]) \to C_{p+r}(Y';A_q[\rho;M])$$

is the family of maps induced by the maps

$$[\mathcal{O}_{Y'}, f']_{y'}^{x'} \cdot (f'_{y'})^*\colon C_q(X_{\kappa(x')};M) \to C_q(X_{\kappa(y')};M)$$

with $y' \in Y'_{(p+r)}$, $x' \in X'_{(p)}$, $f'(y') = x'$ and where $f'_{y'}\colon X_{\kappa(y')} \to X_{\kappa(x')}$ is the natural map. Moreover $(f'_{y'})^*$ is the family of maps

$$[\mathcal{O}_X \otimes_{\mathcal{O}_{X'}} \kappa(y'), f'_{y'}]_y^x \cdot r_{\kappa(x)|\kappa(y)}\colon M(x) \to M(y)$$

with $y \in Y_{(p+q+r)}$ lying over $x \in X_{(p+q)}$ and over y'.

The claim amounts to show for such points y', x', y, x the equality

$$l\big((\kappa(x) \otimes_{\mathcal{O}_{X'}} \mathcal{O}_{Y'})_{(y)}\big) = l\big((\kappa(x') \otimes_{\mathcal{O}_{X'}} \mathcal{O}_{Y'})_{(y')}\big) \cdot l\big((\kappa(x) \otimes_{\mathcal{O}_{X'}} \kappa(y'))_{(y)}\big).$$

For this see (Fulton 1984, A.4.1).

Proof of (4): The map

$$\partial_Y^U\colon C_l(U;M) \to C_{l-1}(Y;M)$$

is on the subquotients of the filtrations given by the family of maps

$$\partial_u\colon C_q(U_{\kappa(u)};M) \to C_{q-1}(Y_{\kappa(u)};M)$$

with $u \in X'_{(p)}$, $p+q=l$ and where ∂_u is the boundary map for the closed immersion $Y_{\kappa(u)} \to X_{\kappa(u)}$. On the other hand

$$[\partial_Y^U]_\#\colon C_p\,(X';A_q[\rho \circ j;M]) \to C_p\,(X';A_q[\rho \circ i;M])$$

is componentwise induced by the maps ∂_u. $\square$

By an affine bundle of dimension n we mean a bundle $\pi\colon V \to X$ which is locally on X isomorphic to $X \times \mathbb{A}^n \to X$ with affine transition maps. (In applications we are mainly interested in the special case of vector bundles.)

A first application of the spectral sequence is

(8.6) PROPOSITION. *Let $\pi\colon V \to X$ be an affine bundle of dimension n. Then*

$$\pi^*\colon A_p(X;M) \to A_{p+n}(V;M)$$

is bijective for all p. If X is equidimensional, then

$$\pi^*\colon A^p(X;M) \to A^p(V;M)$$

is bijective for all p.

Here again the second statement follows from the first and one may drop in the codimension setting the finite type hypothesis.

Proof: By Corollary 8.2 and Proposition 8.5.2 applied to $Y' = X' = X$, $Y = V$, $f = \hat{f} = \pi$, all we need to show is

$$A_q[\pi;M] = 0 \quad \text{for} \quad q \neq n$$

and that

$$[\pi^*]\colon M = A_0[\mathrm{id}_X;M] \to A_n[\pi;M]$$

is an isomorphism of cycle modules over X.

Therefore we are reduced to the case $X = \operatorname{Spec} F$. Then V is a trivial bundle, $V = \mathbb{A}^n_F$. For $n = 1$ the claim is (H) of Section 2. So we know Proposition 8.6 for line bundles over an arbitrary base. But then the case $V = \mathbb{A}^n_F$ follows by induction on n. $\square$

9. Homotopy

We have just observed the homotopy property for affine bundles. In this section we show that this fact can be made more precise on cycle level by means of a homotopy inverse.

A homomorphism $\alpha\colon X \bullet\!\!\rightarrow Y$ with $\delta(\alpha) = 0$ is called a strong homotopy equivalence if there is $r\colon Y \bullet\!\!\rightarrow X$ and $H\colon Y \bullet\!\!\rightarrow Y$ such that

$$\delta(r) = 0, \tag{9.0.1}$$

$$r \circ \alpha = \mathrm{id}, \tag{9.0.2}$$

$$H \circ \alpha = 0, \tag{9.0.3}$$

$$\delta(H) = \mathrm{id} - \alpha \circ r. \tag{9.0.4}$$

The pair (r, H) will be called h-data for α.

Let $\pi\colon V \to X$ be an affine bundle. We will show that $\pi^*\colon X \bullet\!\!\rightarrow V$ is a strong homotopy equivalence. A crucial point here is the treatment of the case $V = X \times \mathbb{A}^1$. The general case is then more or less clear in view of the decomposition of the cycle complexes corresponding to boundary triples. We give here explicit formulas in order to make clear compatibility with base change and filtrations.

By a *coordination* $\tau = (X_i, \tau_i)$ of an affine bundle $\pi\colon V \to X$ of dimension n we mean a sequence $\varnothing = X_0 \subset X_1 \subset \cdots \subset X_k = X$ of closed subsets of X together with trivializations

$$\tau_i\colon V \mid (X_i \setminus X_{i-1}) \to (X_i \setminus X_{i-1}) \times \mathbb{A}^n.$$

(We use the notation $V \mid U = V \times_W U$ for $U \subset W$ and a scheme $V \to W$ over W). Coordinations clearly exist since X is noetherian. For a morphism $f\colon Y \to X$ we denote by $f^*\tau$ the induced coordination on the pull-back bundle f^*V.

In the following we construct in several steps h-data

$$r(\tau)\colon C_p(V; M) \to C_{p-n}(X; M),$$
$$H(\tau)\colon C_p(V; M) \to C_{p+1}(V; M)$$

for π^* depending on a coordination τ.

(9.1) The case $V = X \times \mathbb{A}^1$. h-data (r, H) for $\pi^*\colon X \bullet\!\!\rightarrow X \times \mathbb{A}^1$ are given by the composites

$$r\colon X\times\mathbb{A}^1 \xrightarrow{j^*} X\times(\mathbb{A}^1\setminus\{0\}) \xrightarrow{\{-1/t\}} X\times(\mathbb{A}^1\setminus\{0\}) \xrightarrow{\partial_\infty} X,$$

$$H\colon X\times\mathbb{A}^1 \xrightarrow{p_2^*} X\times(\mathbb{A}^1\times\mathbb{A}^1\setminus\Delta) \xrightarrow{\{s-t\}} X\times(\mathbb{A}^1\times\mathbb{A}^1\setminus\Delta) \xrightarrow{p_{1*}} X\times\mathbb{A}^1.$$

Here t is the coordinate of $\mathbb{A}^1 = \operatorname{Spec}\mathbb{Z}[t]$ and s, t are the coordinates of $\mathbb{A}^1 \times \mathbb{A}^1 = \operatorname{Spec}\mathbb{Z}[s] \times \operatorname{Spec}\mathbb{Z}[t]$. Moreover $\Delta = \{s - t = 0\}$ is the diagonal, j is the standard inclusion, p_1 and p_2 are given by the standard projections and ∂_∞ is induced by $X = X \times \infty \subset X \times (\mathbb{P}^1 \setminus \{0\})$ with open complement $X \times (\mathbb{A}^1 \setminus \{0\})$.

We have to verify for (r,H) the defining properties of h-data. (9.0.1) and (9.0.3) are immediate and (9.0.2) follows from Lemma 4.5. To check (9.0.4) consider the decomposition

$$(p_1)_*\colon X\times(\mathbb{A}^1\times\mathbb{A}^1\setminus\Delta)\overset{q_*}{\bullet\!\!\longrightarrow} X\times\mathbb{A}^1\times\mathbb{P}^1\overset{\bar p_{1*}}{\bullet\!\!\longrightarrow} X\times\mathbb{A}^1$$

where q is the inclusion and $\bar p_1$ is the projection. $\bar p_1$ is proper and therefore

$$\delta(H)=(\bar p_1)_*\circ\delta(q_*)\circ\{s-t\}\circ p_2^*.$$

Moreover

$$\delta(q_*)=(i_\Delta)_*\circ\partial_\Delta+(i_\infty)_*\circ\partial_\infty$$

where $i_\Delta\colon X\times\Delta\to X\times\mathbb{A}^1\times\mathbb{P}^1$, $i_\infty\colon X\times\mathbb{A}^1\times\infty\to X\times\mathbb{A}^1\times\mathbb{P}^1$ are the inclusions and ∂_Δ, ∂_∞ are the boundary maps for $X\times\Delta\to X\times\mathbb{A}^1\times\mathbb{A}^1$, $X\times\mathbb{A}^1\times\infty\to X\times(\mathbb{A}^1\times\mathbb{P}^1\setminus\Delta)$, respectively.

Since $s-t$ is a parameter for Δ one finds

$$(\bar p_1)_*\circ(i_\Delta)_*\circ\partial_\Delta\circ\{s-t\}\circ p_2^*=\mathrm{id}$$

by Lemma 4.5.

Let $W=\mathbb{A}^1\times\mathbb{P}^1\setminus(\Delta\cup\mathbb{A}^1\times 0)$. Moreover let $\tilde p_2$ be the restriction of p_2 to $U=X\times(W\setminus\mathbb{A}^1\times\infty)$ and let $\tilde\partial_\infty$ be the boundary map corresponding to the inclusion $X\times\mathbb{A}^1\times\infty\to X\times W$. Then

$$\partial_\infty\circ\{s-t\}\circ p_2^*=\tilde\partial_\infty\circ\{s-t\}\circ\tilde p_2^*.$$

Since $(s-t)/(-t)$ is a unit on W with constant value 1 on $X\times\mathbb{A}^1\times\infty$ one has

$$\tilde\partial_\infty\circ\{s-t\}\circ\tilde p_2^*=\tilde\partial_\infty\circ\{-t\}\circ\tilde p_2^*.$$

The compatibility of the boundary maps with flat pull-back now gives

$$(\bar p_1)_*\circ(i_\infty)_*\circ\partial_\infty\circ\{s-t\}\circ p_2^*=-\pi^*\circ r.$$

Putting things together yields (9.0.4).

(9.2) The case $V=X\times\mathbb{A}^n$. Let $\pi_n\colon X\times\mathbb{A}^n\to X$ be the projection and put $\pi_X^n=\pi_n^*\colon X\bullet\!\!\rightsquigarrow X\times\mathbb{A}^n$. By induction on n we define h-data (r_X^n,H_X^n) for π_X^n. Let $Y=X\times\mathbb{A}^1$ so that $Y\times\mathbb{A}^{n-1}=X\times\mathbb{A}^n$. Note that $\pi_X^n=\pi_Y^{n-1}\circ\pi^*$ where $\pi\colon Y\to X$ is the projection. Put

$$\begin{aligned} r_X^n&=r\circ r_Y^{n-1},\\ H_X^n&=H_Y^{n-1}+\pi_Y^{n-1}\circ H\circ r_Y^{n-1}.\end{aligned}$$

Here (r,H) are the h-data for π^* from (9.1); moreover, $r_Y^0=\pi_Y^0=\mathrm{id}_Y^*$ and $H_Y^0=0$.

The properties (9.0.1) and (9.0.2) can be easily verified. For (9.0.3) note that

$$\begin{aligned} H_X^n \circ \pi_X^n &= H_X^n \circ \pi_Y^{n-1} \circ \pi^* \\ &= (H_Y^{n-1} \circ \pi_Y^{n-1}) \circ \pi^* + \pi_Y^{n-1} \circ H \circ (r_Y^{n-1} \circ \pi_Y^n) \circ \pi^* \\ &= 0 + \pi_Y^{n-1} \circ (H \circ \pi^*) = 0. \end{aligned}$$

Finally (9.0.4) follows from

$$\begin{aligned} \delta(H_X^n) &= \delta(H_Y^{n-1}) + \pi_Y^{n-1} \circ \delta(H) \circ r_Y^{n-1} \\ &= 1 - \pi_Y^{n-1} \circ r_Y^{n-1} + \pi_Y^{n-1} \circ (1 - \pi^* \circ r) \circ r_Y^{n-1} \\ &= 1 - \pi_X^n \circ r_Y^n. \end{aligned}$$

(9.3) Glueing. Let $\pi\colon V \to X$ be an affine bundle, let $Y \subset X$ be closed, let $U = X \setminus Y$ and put $V' = V \mid Y$, $V'' = V \mid U$. For given h-data (r', H') for $(\pi \mid V')^*\colon Y \multimap V'$ and (r'', H'') for $(\pi \mid V'')^*\colon U \multimap V''$ we define h-data (r, H) for $\pi^*\colon X \multimap V$ by the formulae:

$$r = \begin{pmatrix} r' & -r' \circ \partial \circ H'' \\ 0 & r'' \end{pmatrix}, \qquad H = \begin{pmatrix} H' & -H' \circ \partial \circ H'' \\ 0 & H'' \end{pmatrix}.$$

Here the matrix notation refers to the natural decompositions

$$\begin{aligned} C_*(X; M) &= C_*(Y; M) \oplus C_*(U; M), \\ C_*(V; M) &= C_*(V'; M) \oplus C_*(V''; M). \end{aligned}$$

Moreover $\partial\colon V'' \multimap V'$ is the boundary map corresponding to $V' \subset V$. The verification of (9.0.1)–(9.0.4) is straightforward and omitted.

(9.4) The general case. Given a coordination τ one uses iteratively the recipe of (9.3) to construct h-data $\big(r(\tau), H(\tau)\big)$ for π^*.

It turns out that the glueing process of (9.3) is "associative" in the sense that $\big(r(\tau), H(\tau)\big)$ does not depend on the ordering in which the different pieces are glued together. However, this is not at all important for us; one should just decide oneself for some fixed standard ordering.

(9.5) Functoriality. The construction of $\big(r(\tau), H(\tau)\big)$ is compatible with manipulations on the base given by the four types of maps f_*, g^*, $\{a\}$ and ∂. We omit a formulation, since this will be used only in the trivial case of open immersions g.

(9.6) Proposition.

(1) *Let $\pi\colon V \to X$ be an affine bundle of dimension n with coordination τ and let $\rho\colon X \to X'$ be a morphism. Then*

$$\begin{aligned} r(\tau)\big(C_{p,l}(\rho \circ \pi)\big) &\subset C_{p,l-n}(\rho), \\ H(\tau)\big(C_{p,l}(\rho \circ \pi)\big) &\subset C_{p,l+1}(\rho). \end{aligned}$$

(2) *Let*

$$\begin{array}{ccc} V & \xrightarrow{\pi} & X \\ \downarrow{\scriptstyle\eta} & & \downarrow{\scriptstyle\rho} \\ V' & \xrightarrow{\pi'} & X' \end{array}$$

be a pull-back diagram with π' an affine bundle of dimension n and let τ' be a coordination for π'. Then

$$r(\rho^*\tau')\big(C_{p,l}(\eta)\big) \subset C_{p-n,l-n}(\rho),$$
$$H(\rho^*\tau')\big(C_{p,l}(\eta)\big) \subset C_{p+1,l+1}(\rho).$$

Proof: This is straightforward (but nevertheless tedious) by following the constructions. □

In order to define h-data as above one needs less than the choice of a coordination. For example, in (9.2) one refers alone to trivializations of the one-dimensional bundles $X \times \mathbb{A}^{m+1} \to X \times \mathbb{A}^m$.

We have not tried to describe the precise amount of information of a coordination needed in order to perform the above construction.

10. Deformation to the Normal Cone

This section describes three general constructions associated to closed imbeddings: the normal cone, the deformation space and the double deformation space.

For the general role of the deformation space in intersection theory, we refer to (Fulton 1984). The double deformation space will be our tool to verify associativity of the intersection operations.

We first fix notations and describe some significant properties. Explicit descriptions are given in (10.3)–(10.5) below.

Let $Z \to Y \to X$ be closed imbeddings.

The normal cone to Y in X is denoted by $N = N_Y X = N(X,Y)$. There is the projection $N_Y X \to Y$ and the inclusion $Y \to N_Y X$. If $Y \to X$ is a regular imbedding, then $N_Y X$ is a vector bundle over Y with the inclusion as zero section.

The deformation space $D = D(X,Y)$ is a scheme over $X \times \mathbb{A}^1$. It is flat over $\mathbb{A}^1$. Over $\mathbb{A}^1 \setminus \{0\} \subset \mathbb{A}^1$ one has

$$D \mid (\mathbb{A}^1 \setminus \{0\}) = X \times (\mathbb{A}^1 \setminus \{0\}).$$

Furthermore the projection $D \,|\, \{0\} \to X \times \{0\}$ factors through $Y \to X \times \{0\}$ and one has

$$D \mid \{0\} = N_Y X$$

as schemes over Y. (Our D is in Fulton 1984, Chap. 5 denoted by M^0; moreover we have taken 0 instead of ∞ as the basepoint of the special fiber.)

The double deformation space $\overline{D} = \overline{D}(X,Y,Z)$ is a scheme over $X \times \mathbb{A}^2$. It is flat over $\mathbb{A}^2$ and one has the following canonical identifications of schemes over $\mathbb{A}^2$, assuming in (10.0.3)–(10.0.5) that $Z \to Y \to X$ are regular imbeddings.

$$\overline{D} \mid \mathbb{A}^1 \times (\mathbb{A}^1 \setminus \{0\}) = D(X,Y) \times (\mathbb{A}^1 \setminus \{0\}), \tag{10.0.1}$$

$$\overline{D} \mid (\mathbb{A}^1 \setminus \{0\}) \times \mathbb{A}^1 = (\mathbb{A}^1 \setminus \{0\}) \times D(X,Z), \tag{10.0.2}$$

$$\overline{D} \mid \mathbb{A}^1 \times \{0\} = D(N_Z X, N_Z Y), \tag{10.0.3}$$

$$\overline{D} \mid \{0\} \times \mathbb{A}^1 = D(N_Y X, N_Y X \mid Z). \tag{10.0.4}$$

Moreover the projection $\overline{D} \mid \{(0,0)\} \to X \times \{(0,0)\}$ factors through $Z \to X$ and one has

$$\overline{D} \mid \{(0,0)\} = N(N_Z X, N_Z Y) = N(N_Y X, N_Y X \mid Z) \tag{10.0.5}$$

as schemes over Z.

There is a more symmetric but less general version of the double deformation space. Namely, let Y, Y' be closed subschemes of X and let Z be the intersection of Y and Y', i.e., $Z = Y \times_X Y'$. Then there is a double deformation space $\widetilde{D} = \widetilde{D}(X;Y,Y') \to X \times \mathbb{A}^2$ relating (in the transversal case) all five inclusions induced from $Z \subset Y, Y' \subset X$. The deformation space $\widetilde{D}$ is flat over $\mathbb{A}^2$ and symmetric with respect to a simultaneous interchange of Y, Y' and of the factors of $\mathbb{A}^2 = \mathbb{A}^1 \times \mathbb{A}^1$. Suppose that Y and Y' meet transversally. Then

$$\widetilde{D} \mid (\mathbb{A}^1 \setminus \{0\}) \times \mathbb{A}^1 = D(X,Y) \times \mathbb{A}^1,$$
$$\widetilde{D} \mid \{0\} \times \mathbb{A}^1 = D(N_Y X, N_Y X \mid Z).$$

Moreover one has $\widetilde{D} \mid L = D(X,Z)$ for any line $L \subset \mathbb{A}^2$ through the origin as long as L is different from the two axes.

In the case $Z = Y \times_X Y'$, the space $\overline{D}(X,Y,Z)$ is the pull-back of the space $\widetilde{D}(X;Y,Y')$ along the affine blow up $\mathbb{A}^2 \to \mathbb{A}^2$, $(t,s) \to (ts,s)$. We don't need $\widetilde{D}$, but we have included below its definition, since it might be a bit simpler to understand than $\overline{D}$.

We have to recall facts from local algebra. Remark 10.1 is a special version of the local criterion of flatness (Matsumura 1980, (20.G), p. 152). Remark 10.2 may be deduced by considering locally regular sequences for J containing regular sequences for I (see Serre 1957). For a compact account of other facts needed in the following we refer to (Fulton 1984, App. A, App. B).

(10.1) Remark. A morphism $U \to V \times \mathbb{A}^1$ is flat if and only if the morphisms $U \times_{\mathbb{A}^1} (\mathbb{A}^1 \setminus \{0\}) \to V \times (\mathbb{A}^1 \setminus \{0\})$, $U \times_{\mathbb{A}^1} \{0\} \to V \times \{0\}$ and $U \to \mathbb{A}^1$ are flat. □

(10.2) Remark. If $Z \to Y$ and $Y \to X$ are regular imbeddings, then $Z \to X$ is a regular imbedding. If X is affine, and if $X = \operatorname{Spec} A$ and $I \subset J \subset A$ are the ideals corresponding to Y and Z, respectively, then

$$I^n J^m \cap J^{m+n+1} = I^n J^{m+1}, \tag{1}$$

$$I^n J^m \cap I^{n+1} = I^{n+1} J^{m-1}. \tag{2}$$

Here we understand n, $m \in \mathbb{Z}$ with $I^n = J^n = A$ for $n \leq 0$. □

We next give the definitions of N, D, $\overline{D}$ and $\widetilde{D}$ for affine X. From the naturality of the affine constructions it will be obvious that they extend to global ones.

We keep the notations of Remark 10.2. Moreover we use the coordinates $\mathbb{A}^1 = \operatorname{Spec} k[t]$ and $\mathbb{A}^2 = \operatorname{Spec} k[t, s]$. The indices n, m always run in $\mathbb{Z}$.

(10.3) The normal cone. $N = N_Y X$ is defined as the spectrum of the ring

$$O_N = \coprod_n I^n / I^{n+1}.$$

O_N is a ring over $O_Y = A/I$ and projection to the degree zero summand gives a homomorphism $O_N \to O_Y$.

(10.4) The deformation space. $D = D(X, Y)$ is defined as the spectrum of the subring

$$O_D = \sum_n I^n \cdot t^{-n} \subset A[t, t^{-1}].$$

O_D is a finitely generated ring over $A[t]$ (with generators $x_i t^{-1}$ if x_i are generators of I). After inverting t one has

$$O_D[t^{-1}] = A[t, t^{-1}].$$

Since t is not a zero divisor, it follows that O_D is flat over $k[t]$. Moreover

$$O_D \,/\, t \cdot O_D = \coprod_n I^n / I^{n+1} = O_N.$$

For later purposes we are very precise about this identification: for $x \in I^n$ the residue of $x \cdot t^{-n} \bmod I^{n+1} \cdot t^{-n}$ corresponds to $(-1)^n x \bmod I^{n+1}$. (This sign convention will avoid some other signs later on.)

(10.5) The double deformation space. $\overline{D} = \overline{D}(X, Y, Z)$ is the spectrum of the subring

$$O_{\overline{D}} = \sum_{n,m} I^n J^{m-n} \cdot t^{-n} s^{-m} \subset A[t, s, t^{-1}, s^{-1}].$$

$O_{\overline{D}}$ is finitely generated over $A[s, t]$. After inverting s or t one has (with $D' = D(X, Z)$)

$$O_{\overline{D}}[s^{-1}] = \sum_{n,m} I^n \cdot t^{-n} s^{-m} = O_D[s, s^{-1}],$$

$$O_{\overline{D}}[t^{-1}] = \sum_{n,m} J^m \cdot t^{-n} s^{-m} = O_{D'}[t, t^{-1}].$$

This shows (10.0.1) and (10.0.2). For (10.0.3) note first

$$O_{\overline{D}} \,/\, s \cdot O_{\overline{D}} = \coprod_{n,m} [I^n J^{m-n} / I^n J^{m-n+1}] \cdot t^{-n} s^{-m}.$$

In order to make clear the ring structures (in particular as ring over $k[t,s]$) we keep here the terms $t^{-n}s^{-m}$, having now merely the meaning of symbols.

Moreover $N_Z X$ and $N_Z Y$ are the spectra of

$$R = \coprod_m [J^m / J^{m+1}] \cdot s^{-m},$$

$$R' = \coprod_m [(J^m + I)/(J^{m+1} + I)] \cdot s^{-m},$$

The projection $R \to R'$ yields an inclusion $N_Z Y \to N_Z X$.

Let $\widetilde{I} = \ker(R \to R')$. By Remark 10.2.2 one has $J^m \cap I \subset I \cdot J^{m-1}$ and therefore

$$\widetilde{I} = \coprod_m [(I \cdot J^{m-1} + J^{m+1})/J^{m+1}] \cdot s^{-m}$$

and

$$\widetilde{I}^n = \coprod_m [(I^n \cdot J^{m-n} + J^{m+1})/J^{m+1}] \cdot s^{-m}.$$

Hence $D(N_Z X, N_Z Y)$ is the spectrum of

$$\coprod_n \widetilde{I}^n = \coprod_{n,m} [(I^n \cdot J^{m-n} + J^{m+1})/J^{m+1}] \cdot t^{-n} s^{-m}.$$

(10.0.3) follows now from Remark 10.2.1.

For (10.0.4) note first

$$O_{\overline{D}} \,/\, t \cdot O_{\overline{D}} = \coprod_{n,m} [I^n J^{m-n} / I^{n+1} J^{m-n-1}] \cdot t^{-n} s^{-m}.$$

For the ring of $N_Y X$ we write now

$$O_N = \coprod_n [I^n / I^{n+1}] \cdot u^{-n}.$$

Let $\widetilde{J} \subset O_N$ be the ideal corresponding to the closed subscheme $N_Y X \,|\, Z$. Its powers are

$$\widetilde{J}^m = \coprod_n [(J^m I^n + I^{n+1})/I^{n+1}] \cdot u^{-n}.$$

Hence $D(N_Y Z, N_Y X \mid Z)$ is the spectrum of

$$R'' = \coprod_{n,m} [(J^m I^n + I^{n+1})/I^{n+1}] \cdot u^{-n} s^{-m}.$$

Define

$$\varphi\colon O_{\overline{D}} \,/\, t \cdot O_{\overline{D}} \to R'',$$
$$(x \bmod I^{n+1} J^{m-n-1}) \cdot t^{-n} s^{-m} \to (x \bmod I^{n+1}) \cdot u^{-n} s^{-m+n}.$$

It is easy to see that φ is a surjective ring homomorphism over $k[s]$. Moreover φ is injective by Remark 10.2.1. The map φ gives the identification of (10.0.4). Now (10.0.5) is obvious. The flatness over $\mathbb{A}^2$ (not needed in the following) may be deduced from Remark 10.1.

(10.6) THE DOUBLE DEFORMATION SPACE $\widetilde{D}$. We just give the definition. Let $I' \subset A$ be the ideal corresponding to $Y' \to X$. One puts

$$O_{\widetilde{D}} = \sum_{n,m} I^n I'^m \cdot v^{-n} s^{-m} \subset A[v, s, v^{-1}, s^{-1}].$$

One may handle with $\widetilde{D}$ similar as with $\overline{D}$ in (10.5). In the transversal case one has

$$O_{\overline{D}} = O_{\widetilde{D}} \otimes_{k[v,s]} k[t, s]$$

where $k[v, s] \subset k[t, s]$ via $v \to ts$, $s \to s$ and with $J = I + I'$.

11. THE BASIC CONSTRUCTION

For a closed immersion $i: Y \to X$ we define

$$J(i) = J(X, Y): X \bullet\!\!\to N_Y X$$

as the composite of

$$X \overset{\pi^*}{\bullet\!\!\longrightarrow} X \times (\mathbb{A}^1 \setminus \{0\}) \overset{\{t\}}{\bullet\!\!\longrightarrow} X \times (\mathbb{A}^1 \setminus \{0\}) \overset{\partial}{\bullet\!\!\longrightarrow} N_Y X.$$

Here $\pi: X \times (\mathbb{A}^1 \setminus \{0\}) \to X$ is the projection, $\mathbb{A}^1 = \operatorname{Spec} k[t]$ and ∂ is the boundary map for $N_Y X \to D(X, Y)$. One has $\delta\big(J(X, Y)\big) = 0$ so that $J(X, Y)$ is a homomorphism of complexes $C_*(X; M) \to C_*(N_Y X; M)$.

If $M = K_*$, then the restriction of $J(X, Y)$ to the classical cycle groups coincides with the specialization homomorphisms σ of (Fulton 1984, Chap. 5.2); this may be deduced from the description of σ in (Fulton 1984, Prop. 5.2) via Cartier divisors. As for classical cycles, one may think of $J(X, Y)$ as the pull-back along tubular neighborhoods followed by a linearization process. In the following we have collected the remarks on $J(X, Y)$ which are needed in further sections. We have not tried to give a detailed geometrical description.

The construction of $J(X, Y)$ is local in the sense that

$$J(U, Y \cap U) \circ j^* = \tilde{\jmath}^* \circ J(X, Y)$$

where $j: U \to X$ is an open immersion and $\tilde{\jmath}: N_Y X \mid (Y \cap U) \to N_Y X$ is the induced inclusion.

(11.1) LEMMA. *Let $\sigma: Y \to N_Y X$ be the inclusion. Then*

$$J(X, Y) \circ i_* = \sigma_*.$$

Proof: The statement follows from Lemma 4.5 and the fact that the closure of $Y \times (\mathbb{A}^1 \setminus \{0\})$ in $D(X, Y)$ is $Y \times \mathbb{A}^1$. □

Let X be normal, $y \in X^{(1)}$ and $Y = \overline{\{y\}}$. Moreover let F and E be the function fields of X and $N_Y X$, respectively. We want to compute the codimension 0 component

$$J^0 \colon M(F) \to M(E)$$

of $J(X,Y)$. The problem is purely local in y. Let v be the valuation on F corresponding to y and let $\kappa = \kappa(y) = \kappa(v)$. Moreover let $\mathfrak{m}$ be the ideal of y, let $\pi \in \mathfrak{m}$ be a prime and let $\bar{\pi} \in \mathfrak{m}/\mathfrak{m}^2$ be its image. The normal cone $N_Y X$ is the spectrum of

$$\kappa[\bar{\pi}] = \coprod_n \mathfrak{m}^n/\mathfrak{m}^{n+1}$$

and one has $E = \kappa(\bar{\pi})$.

The following lemma shows that there is a factorization

$$J^0 \colon M(F) \xrightarrow{p} M(v) \to M(E)$$

where p is from Remark 1.6.

(11.2) Lemma.

$$J^0 = r_{E|\kappa} \circ s_v^\pi + \{\bar{\pi}\} \cdot r_{E|\kappa} \circ \partial_v .$$

Proof: We may suppose $X = \operatorname{Spec} A$ and that the ideal I corresponding to y is generated by π. Then $D(X,Y)$ is the spectrum of

$$A[t, \pi t^{-1}] \subset A[t, t^{-1}].$$

By definition we have

$$J^0 = \partial_w \circ \{t\} \circ r_{F(t)|F}$$

where w is the valuation on $F(t)$ corresponding to the principal ideal $t \cdot A[t, \pi t^{-1}]$. Note that $E = \kappa(w)$ and that $\bar{\pi}$ is the residue of the w-unit $-\pi t^{-1}$ (by the sign convention in 10.4). The claim follows now from

$$\begin{aligned} \partial_w \circ \{t\} \circ r_{F(t)|F}(\rho) &= \partial_w \circ \{-\pi\} \circ r_{F(t)|F}(\rho) - \partial_w \circ \{-\pi t^{-1}\} \circ r_{F(t)|F}(\rho) \\ &= \partial_w \circ r_{F(t)|F}(\{-\pi\} \cdot \rho) + \{\bar{\pi}\} \circ \partial_w \circ r_{F(t)|F}(\rho) \end{aligned}$$

and the fact that w restricts on F to v. □

The preceding remarks yield a complete description of $J(X,Y)$ for smooth curves X.

The rest of the section contains a series of technical lemmata.

(11.3) Lemma. *Let $Y \to X$ be a closed immersion, let $g \colon V \to X$ be flat (of constant relative dimension) and let*

$$N(g) \colon N(V, Y \times_X V) = N(X,Y) \times_X V \to N(X,Y)$$

be the projection. Then

$$J(V, Y \times_X V) \circ g^* = N(g)^* \circ J(X,Y).$$

Proof: This follows from the flatness of $D(V, Y \times_X V) \to D(X,Y)$. □

(11.4) LEMMA. *Let $U \to V$ be a closed immersion and let $p\colon V \to W$ be flat. Suppose that the composite*

$$q\colon N_U V \to U \to V \to W$$

is flat of the same relative dimension as p. Then

$$J(V,U) \circ p^* = q^*\colon W \bullet\!\!\to N_U V.$$

Proof: Let $\pi\colon W \times (\mathbb{A}^1 \setminus \{0\}) \to W$ be the projection and let f be the composite

$$f\colon D(V,U) \longrightarrow V \times \mathbb{A}^1 \xrightarrow{p\times \mathrm{id}} W \times \mathbb{A}^1.$$

Then, by definition,

$$J(V,U) \circ p^* = \partial \circ \{t\} \circ \big(f \mid V \times (\mathbb{A}^1 \setminus \{0\})\big)^* \circ \pi^*.$$

Now f is flat by Remark 10.1 and $f \mid N_U V = q$. Hence

$$J(V,U) \circ p^* = q^* \circ \partial' \circ \{t\} \circ \pi^*$$

where ∂' is the boundary map corresponding to $W \times \{0\} \hookrightarrow W \times \mathbb{A}^1$. But $\partial' \circ \{t\} \circ \pi^* =$ id by Lemma 4.5. $\square$

(11.5) LEMMA. *Let $U \to V$ be a regular imbedding, let $p\colon V \to W$ be smooth of constant relative dimension and suppose $p \circ i$ is a regular imbedding. Then the projection*

$$q\colon N_U V \to N_U W$$

is an epimorphism of vector bundles and

$$J(V,U) \circ p^* = q^* \circ J(W,U).$$

Proof: Use the flatness of $D(V,U) \to D(W,U)$. $\square$

(11.6) LEMMA. *Let $\rho\colon T \to T'$ be a morphism, let $T_1', T_2' \subset T'$ be closed subschemes and let $T_i = T \times_{T'} T_i'$ for $i = 1, 2$.*

Put $T_3 = T \setminus (T_1 \cup T_2)$, $T_0 = T_1 \cap T_2$, $\widetilde{T}_1 = T_1 \setminus T_0$, $\widetilde{T}_2 = T_2 \setminus T_0$ and let ∂_1^3, ∂_0^1, ∂_2^3, ∂_0^2 be the boundary maps for the closed immersions

$$\widetilde{T}_1 \to T \setminus T_2, \quad T_0 \to T_1, \quad \widetilde{T}_2 \to T \setminus T_1, \quad T_0 \to T_2,$$

respectively. Then

$$0 \simeq \partial_0^1 \circ \partial_1^3 + \partial_0^2 \circ \partial_2^3\colon T_3 \bullet\!\!\to T_0$$

under a filtration preserving homotopy of degree $(-1,-1)$.

Proof: Corresponding to the set theoretic decomposition of T we have

$$C_*(T;M) = C_*(T_0;M) \oplus C_*(\widetilde{T}_1;M) \oplus C_*(\widetilde{T}_2;M) \oplus C_*(T_3;M).$$

Let

$$\partial_0^3 \colon T_3 \bullet\!\!\rightarrow T_0$$

be the corresponding component of d_T. Then $d_T \circ d_T = 0$ gives

$$\delta(\partial_0^3) + \partial_0^1 \circ \partial_1^3 + \partial_0^2 \circ \partial_2^3 = 0.$$

Hence $-\partial_0^3$ is a homotopy as required. $\square$

Let $T = \overline{D} = \overline{D}(X,Y,Z)$, $T_1 = \overline{D}\,|\,(\{0\} \times \mathbb{A}^1)$, $T_2 = \overline{D}\,|\,(\mathbb{A}^1 \times \{0\})$. We keep the notations of Lemma 11.6. Then $T_3 = X \times (\mathbb{A}^1 \setminus \{0\}) \times (\mathbb{A}^1 \setminus 0)$ and $T_0 = \overline{D}\,|\,\{(0,0)\}$. Let $\pi \colon T_3 \to X$ be the projection and let t, s be the coordinates of $\mathbb{A}^2$ (as in (10.5), so that $T_1 = \{t=0\}$, $T_2 = \{s=0\}$).

(11.7) Lemma. *Let $Z \to Y \to X$ be regular imbeddings. Then*

$$\partial_0^1 \circ \partial_1^3 \circ \{t,s\} \circ \pi^* = J(N_Y X, N_Y X \mid Z) \circ J(X,Y),$$
$$\partial_0^2 \circ \partial_2^3 \circ \{s,t\} \circ \pi^* = J(N_Z X, N_Z Y) \circ J(X,Z).$$

Proof: Let

$$\pi_i \colon X \times (\mathbb{A}^1 \setminus \{0\}) \times (\mathbb{A}^1 \setminus \{0\}) \to X \times (\mathbb{A}^1 \setminus \{0\}), \quad \pi^i \colon X \times (\mathbb{A}^1 \setminus \{0\}) \to X$$

be the projections with

$$\pi_1(x,t,s) = (x,s), \quad \pi^1(x,s) = x,$$
$$\pi_2(x,t,s) = (x,t), \quad \pi^2(x,t) = x.$$

One finds (using in particular Lemma 11.3):

$$\begin{aligned}
\partial_0^1 \circ \partial_1^3 \circ \{t,s\} \circ \pi^* &= \partial_0^1 \circ \{s\} \circ \partial_1^3 \circ \{t\} \circ \pi_1^* \circ \pi^{1*} \\
&= \partial_0^1 \circ \{s\} \circ J\big(X \times (\mathbb{A}^1 \setminus 0); Y \times (\mathbb{A}^1 \setminus \{0\})\big) \circ \pi^{1*} \\
&= \partial_0^1 \circ \{s\} \circ \big(N_Y X \times (\mathbb{A}^1 \setminus \{0\}) \to N_Y X\big)^* \circ J(X,Y) \\
&= J(N_Y X, N_Y X \mid Z) \circ J(X,Y),
\end{aligned}$$

$$\begin{aligned}
\partial_0^2 \circ \partial_2^3 \circ \{s,t\} \circ \pi^* &= \partial_0^2 \circ \{t\} \circ \partial_1^2 \circ \{s\} \circ \pi_2^* \circ \pi^{2*} \\
&= \partial_0^2 \circ \{t\} \circ J\big(X \times (\mathbb{A}^1 \setminus \{0\}), Z \times (\mathbb{A}^1 \setminus \{0\})\big) \circ \pi^{2*} \\
&= \partial_0^2 \circ \{t\} \circ \big(N_Z X \times (\mathbb{A}^1 \setminus \{0\}) \to N_Z X\big)^* \circ J(X,Z) \\
&= J(N_Z X, N_Z Y) \circ J(X,Z).
\end{aligned}$$
$\square$

12. THE PULL-BACK MAP

In this section we define the pull-back maps for morphisms to smooth varieties. Some properties are formulated, in particular the functoriality of the spectral sequences. We conclude with applications and discussions concerning birational questions. The proofs of Theorems 12.1 and 12.7 are given in the next section.

In the following all schemes $X, Y, X', \ldots$ are flat over B of some constant relative dimension denoted by $\dim_B X, \ldots$. All products $Y \times X$, $Y' \times X'$, $\ldots$ are taken over B and cycle modules will be induced via projection to the second factor (projection to the first factor does not exist for us). We use the notations $T_S X = N_X(X \times_S X)$ and $TX = T_B X$. We are primarily interested in the case $B = \operatorname{Spec} k$, but we don't have to pay much for considering arbitrary B. M is a cycle module over X.

Let X be smooth over B. Then TX is a vector bundle on X. For a morphism $f\colon Y \to X$ let

$$f\colon Y \xrightarrow{i} Y \times X \xrightarrow{p} X$$

be the factorization with $i(y) = \bigl(y, f(y)\bigr)$ and $p(y,x) = x$. Then i is a regular imbedding and $N_Y(Y \times X) = f^* TX$.

We choose a coordination τ on TX and define

$$I(f) = I(f;\tau) = r(f^*\tau) \circ J(Y \times X, Y) \circ p^*\colon X \bullet\!\!\to Y.$$

Note that the construction is local in the sense that for an open immersion $j\colon U \to X$ one has

$$I(\tilde{f}; j^*\tau) \circ j^* = \tilde{\jmath}^* \circ I(f;\tau)$$

where $\tilde{f}\colon f^{-1}(U) \to U$ is the restriction of f and $\tilde{\jmath}\colon f^{-1}(U) \to Y$ is the inclusion.

One has $\delta\bigl(I(f)\bigr) = 0$ and

$$I(f)\bigl(C_p(X;M)\bigr) \subset C_{p+r}(Y;M)$$

where $r = \dim_B Y - \dim_B X$. If B is equidimensional, then

$$I(f)\bigl(C^p(X;M)\bigr) \subset C^p(Y;M).$$

We define

$$f^\bullet\colon A_p(X;M) \to A_{p+r}(Y;M)$$

and

$$f^\bullet\colon A^p(X;M) \to A^p(Y;M)$$

as the induced maps on (co-)homology. $f^\bullet$ does not depend on the choice of τ. One has the following properties.

(12.1) THEOREM. *For $g\colon Z \to Y$ and $f\colon Y \to X$ with X and Y smooth over B one has $(f \circ g)^\bullet = g^\bullet \circ f^\bullet$.*

For the proof see the next section.

(12.2) Proposition. *If f is flat, then $I(f) = f^*$.*

Proof: It suffices to show

$$J(Y \times X, Y) \circ p^* = \pi^* \circ f^*$$

where π: $f^*TX \to Y$ is the projection. For this apply Lemma 11.4 with $U = Y$, $V = Y \times X$ and $W = X$. □

(12.3) Proposition. *If i: $Y \to X$ is a regular imbedding and X is smooth over B, then $I(i)$ is homotopic to $r \circ J(i)$ where r is any retraction to $Y \hookrightarrow N_Y X$.*

Proof: Apply Lemma 11.5 with $U = Y$, $V = Y \times X$ and $W = X$. □

The following corollary applied to the blow up at x_0 implies (together with Theorem 12.1) property (E) of Section 2.

(12.4) Corollary. *Let X be smooth over $B = \operatorname{Spec} k$, let $x \in X^{(p)}$ and let*

$$i_x\colon \overline{\{x\}} \to X$$

be the inclusion. Moreover let $\pi_1, \ldots, \pi_p$ be any regular sequence at x and let $v_1, \ldots, v_p$ be the induced sequence of valuations with the fraction fields $k(X)$, $\kappa(v_1), \ldots, \kappa(v_{p-1})$ and with the (residue classes of) $\pi_1, \ldots, \pi_p$ as primes. Then

$$i_x^\bullet\colon A^0(X; M) \to A^0(\overline{\{x\}}; M)$$

is the restriction of

$$s_{v_p}^{\pi_p} \circ \cdots \circ s_{v_1}^{\pi_1}\colon M\bigl(k(X)\bigr) \to M\bigl(\kappa(x)\bigr).$$

Proof: Let $X = X_0 \supset X_1 \supset \cdots \supset X_p$ be the sequence of smooth schemes locally around x with X_i defined by $\langle \pi_1, \ldots, \pi_i \rangle$. Using Theorem 12.1 one reduces to $p = 1$. This case follows from Proposition 12.3 and Lemma 11.2. □

(12.5) Proposition. (Projection formula.) *Consider a pull-back square*

$$\begin{array}{ccc} \overline{Y} & \xrightarrow{\bar f} & \overline{X} \\ \Big\downarrow{\scriptstyle \tilde h} & & \Big\downarrow{\scriptstyle h} \\ Y & \xrightarrow{f} & X \end{array}$$

with h smooth and proper and with X smooth over B. Then

$$\tilde h_* \circ \bar f^\bullet = f^\bullet \circ h_*.$$

Proof: One considers the diagram

$$\begin{array}{ccccc}
\overline{Y} & \longrightarrow & \overline{Y}\times\overline{X} & \longrightarrow & \overline{X} \\
\| & & \downarrow{\scriptstyle \hat{h}} & & \| \\
\overline{Y} & \longrightarrow & Y\times\overline{X} & \longrightarrow & \overline{X} \\
\downarrow{\scriptstyle \tilde{h}} & & \downarrow & & \downarrow{\scriptstyle h} \\
Y & \longrightarrow & Y\times X & \longrightarrow & X.
\end{array}$$

Here the bottom diagram is the pull-back along h and $\hat{h}=\tilde{h}\times\mathrm{id}_{\overline{X}}$. We have three maps $\overline{X} \bullet\!\!\!\to Y$: the first is constructed along $D(\overline{Y}\times\overline{X},\overline{Y})$ and is given by pull-back to $\overline{Y}\times\overline{X}$ and specialization to $\overline{Y}$ followed by push-forward; the second goes along $D(Y\times\overline{X},\overline{Y})$ and is given by pull-back to $Y\times\overline{X}$ and specialization to $\overline{Y}$ followed by push-forward; the third goes along $D(Y\times X,Y)$ and is given by push-forward, pull-back to $Y\times X$ and specialization to Y. The first two may be related using Lemma 11.5 (with $U=\overline{Y}$, $V=\overline{Y}\times\overline{X}$, and $W=Y\times\overline{X}$), the last two by the compatibility of the constructions with proper push-forward. $\square$

Consider the triangle (7.4.0) and assume that ρ is smooth and η is flat. Define

$$[h/B]\colon A^p[\rho;M]\to A^p[\eta;M]$$

by $[h/B]_F=(h_F)^\bullet$. Here we understand $B=\operatorname{Spec}F$ in the definition of $(h_F)^\bullet$.

(12.6) Proposition. *$[h/B]$ is a homomorphism of cycle modules over B.*

Proof: We apply Proposition 7.4. Since the projection $\pi\colon N(Y\times X,Y)\to Y$ is a vector bundle we know that

$$[\pi^*]\colon A^q[\eta]\to A^q[\eta\circ\pi]$$

is an isomorphism of cycle modules. Moreover

$$[\pi^*]\circ[h/B]=[\partial]\circ[\{t\}]\circ[p^*]$$

where $p\colon Y\times X\times(\mathbb{A}^1\setminus\{0\})\to X$ is the projection and ∂ is the boundary for $N(Y\times X,Y)\to D(Y\times X,Y)$. $\square$

(12.7) Theorem. *Consider the square* (8.4.0) *and its decomposition* (8.5.0). *Suppose that B is equidimensional, η is flat, ρ is smooth and X' (hence also X) is smooth over B. Then the spectral sequences*

$$E_2^{p,q}(\rho)=A^p(X';A^q[\rho;M])\Longrightarrow A^{p+q}(X;M),$$
$$E_2^{p,q}(\eta)=A^p(Y';A^q[\eta;M])\Longrightarrow A^{p+q}(Y;M)$$

commute with the maps

$$A^p(X';A^q[\rho;M])\xrightarrow{(f')^\bullet}A^p(Y';A^q[\rho;M])\xrightarrow{[\hat{f}/Y']_\#}A^p(Y';A^q[\eta;M]),$$
$$f^\bullet\colon A^p(X;M)\to A^p(Y;M).$$

For the proof see the next section. Switching to dimension indices this theorem holds without the equidimensionality assumption on B.

For the rest of the section we assume $B = \operatorname{Spec} k$.

(12.8) Lemma. *Let X be smooth, let Y be integral, let $f\colon Y \to X$ be a dominant morphism and let $\varphi\colon k(X) \to k(Y)$ be the induced homomorphism of the function fields. Then*

$$I(f) \mid M\big(k(X)\big) = \varphi_*\colon M\big(k(X)\big) \to M\big(k(Y)\big).$$

Proof: After replacing Y by an open subset we may assume that f is flat. The claim follows from Proposition 12.2. □

(12.9) Lemma. *Assume in* (12.8) *additionally that f is proper and that φ is an isomorphism. Then $f_* \circ I(f) = \mathrm{id}$.*

Proof: Let

$$\bar{f}\colon D(Y \times X, Y) \to D(X \times X, X)$$

be the proper map induced from f. There is the commutative diagram

$$\begin{array}{ccccccc}
X & \bullet\!\!\longrightarrow & Y \times X \times (\mathbb{A}^1 \setminus \{0\}) & \xrightarrow{\partial\circ\{t\}} & f^*TX & \xrightarrow{r(f^*\tau)} & Y \\
\downarrow{=} & & \downarrow{\tilde{f}_*} & & \downarrow{\hat{f}_*} & & \downarrow{f_*} \\
X & \bullet\!\!\longrightarrow & X \times X \times (\mathbb{A}^1 \setminus \{0\}) & \xrightarrow{\partial\circ\{t\}} & TX & \xrightarrow{r(\tau)} & X
\end{array}$$

where $\tilde{f}$, $\hat{f}$ are the restrictions of $\bar{f}$. The diagram shows $f_* \circ I(f) = I(\mathrm{id})$. But $I(\mathrm{id}) = \mathrm{id}$ by Proposition 12.2. □

Lemma 12.9 shows in particular that for any blow up $Y \to X$ the complex $C_*(X; M)$ is a direct summand of $C_*(Y; M)$. This splitting via $I(f)$ depends alone on the choice of a coordination of TX near the singular locus and is unique up to homotopy.

(12.10) Corollary. *Let X be a proper smooth variety over k and let M be a cycle module over k. Then the group $A^0(X; M)$ is a birational invariant of X.*

Proof: If X_1, X_2 are proper and birational isomorphic there exist a proper Y and birational morphisms $Y \to X_i$ (take for Y the closure in $X_1 \times X_2$ of a common open subset of the X_i). Then as subgroups of $M(\xi_{X_i}) = M(\xi_Y)$ one has the trivial inclusions $A^0(Y; M) \subset A^0(X_i; M)$; Lemma 12.8 shows $A^0(X_i; M) \subset A^0(Y; M)$. □

For an illustration let X be a smooth and proper variety over k with function field F. Then for any geometric valuation v on F (of rank 1) there is a birational morphism $f\colon Y \to X$ such that v has center y in $Y^{(1)}$ with $\kappa(v) = \kappa(y)$. The map $I(f)$ yields a formula

(12.11) $$\partial_v = \sum_{x \in X^{(1)}} \alpha_v^x \circ \partial_x$$

where

$$\alpha_v^x\colon M\big(\kappa(x)\big) \to M\big(\kappa(v)\big)$$

equals the component $I(f)_y^x$.

This formula is a sort of higher dimensional analogue of the sum formula for one-dimensional function fields. It has the following properties:

— it is local, that is x runs only through $X_{(z)}^{(1)}$ where z is the center of v in X (in other words: $\alpha_v^x = 0$ for $x \notin X_{(z)}^{(1)}$).

— it is not unique, but depends only on the choice of a coordination of the tangent bundle of X restricted to $X_{(z)}$.

— it is universal in the sense that the α_v^x can be written as sums of compositions of the data of cycle modules, independent of M. This is quite obvious from the construction of $I(f)$. One can make this more precise by interpreting the α_v^x as morphisms in the category $\widetilde{\mathcal{F}}$ of Remark 1.10. In this way the category $\widetilde{\mathcal{F}}$ appears as the natural place for the coefficients α_y^x of formulas like (12.11).

Exercise: Describe the α_v^x for $\dim X = 2$ and v the valuation corresponding to the exceptional fiber of the blow up in a closed point (see Remark 2.8).

Birational invariants like $A^0(X; M)$ have been considered in various contexts like étale cohomology and K-theory, see (Colliot-Thélène 1992) for a survey. The advantage of the method of proof of Corollary 12.10 lies in its general and essentially simple nature (after having established the functors in question as cycle modules); moreover the formula (12.11) makes things perhaps more enlightening. A similar method works probably for functors related with the Witt ring of quadratic forms.

To mention a particular example, let $\pi\colon Z \to \operatorname{Spec} k$ be proper and let M (resp. N) be the $\mathbb{Z}$-graded cycle module over k given as the cokernel (resp. image) of

$$[\pi_*]\colon A_0[\pi; K_*] \to A_0[\mathrm{id}_{\operatorname{Spec} k}; K_*] = K_*.$$

By Corollary 12.10 the group $A^0(X; M, 1)$ (which is a subquotient of $k(X)^*$) is a birational invariant for proper smooth X over k. The proof of this fact was the main aim of (Rost 1990). There it was achieved by a different method using the triviality of $A^1(\widetilde{X}; N, 1)$ for smooth local $\widetilde{X}$ (proved in this paper in more generality in Section 6).

13. Intersection Theory for Fibrations

The purpose of this section is to prove Theorems 12.1 and 12.7. We define pull-back maps on complex level for regular imbeddings and for morphisms to smooth varieties in fibered situations. Moreover we establish functoriality of the constructions. Most of the work has been done already in Sections 7–11.

Consider a commutative square

$$(\Delta)\qquad \begin{array}{ccc} Y & \xrightarrow{f} & X \\ \big\downarrow{\scriptstyle \eta} & & \big\downarrow{\scriptstyle \rho} \\ Y' & \xrightarrow{f'} & X'. \end{array}$$

The square Δ is called a regular imbedding, if f and f' are regular imbeddings of some constant codimensions and if the induced map

$$p\colon N_Y X \to \eta^* N_{Y'} X'$$

is an epimorphism of vector bundles over Y. The kernel bundle of p is denoted by N_Δ. We consider p also as vector bundle and identify it with $q^* N_\Delta$ where $q\colon \eta^* N_{Y'} X' \to Y$ is the projection.

Let Δ be a regular imbedding and let $\tilde\tau$ and $\tilde\tau'$ be coordinations of $N_\Delta \to Y$ and of $N_{Y'}X' \to Y'$, respectively. We define

$$\overline{J}(\Delta)\colon X \bullet\!\!\to Y$$

by

$$\overline{J}(\Delta) = \overline{J}(\Delta, \tilde\tau, \tilde\tau') = r(\eta^*\tilde\tau') \circ r(q^*\tilde\tau) \circ J(X, Y).$$

The following is clear from Sections 8–9. One has $\delta\big(\overline{J}(\Delta)\big) = 0$ and

$$\overline{J}(\Delta)\big(C_{p,l}(\rho)\big) \subset C_{p+s,l+t}(\eta)$$

with $s = -\operatorname{codim}(f')$ and $t = -\operatorname{codim}(f)$. Moreover the homotopy class of $\overline{J}(\Delta)$ (with respect to the degree (s,t)) does not depend on the choice of $\tilde\tau$ and $\tilde\tau'$.

In the definition of $\overline{J}(\Delta)$ we wanted to be as canonical as possible. If one is interested only in the homotopy class, one may put $\overline{J}(\Delta) = r \circ J(X, Y)$ for any filtration preserving retraction r to $N_Y X \to X$ ("filtration preserving" means always with respect to the natural degrees).

Next consider a diagram

$$(13.1.0)\qquad \begin{array}{ccccc} Z & \xrightarrow{f_1} & Y & \xrightarrow{f_2} & X \\ \big\downarrow{\scriptstyle \mu} & & \big\downarrow{\scriptstyle \eta} & & \big\downarrow{\scriptstyle \rho} \\ Z' & \xrightarrow{f_1'} & Y' & \xrightarrow{f_2'} & X'. \end{array}$$

We denote the left square by Δ_1, the right square by Δ_2 and the composed square by Δ_3.

(13.1) THEOREM. *If Δ_1 and Δ_2 are regular imbeddings, then Δ_3 is a regular imbedding and*
$$\bar J(\Delta_1)\circ\bar J(\Delta_2)\simeq\bar J(\Delta_3)$$
under a filtration preserving homotopy.

Proof: The first statement is straightforward. In the following deduction of the second statement the letters r_i stand for some filtration preserving retractions. We make use of Lemma 11.3 for the homotopy (1), of Lemmata 11.6 and 11.7 for the homotopy (2) and of Lemma 11.4 for the equality (3).

$$\begin{aligned}
\bar J(\Delta_1)\circ\bar J(\Delta_2) &\simeq r_1\circ J(Y,Z)\circ r_2\circ J(X,Y)\\
&\simeq r_3\circ\big(N(N_YX,N_YX\mid Z)\to N(Y,Z)\big)^*\circ J(Y,Z)\circ r_2\circ J(X,Y)\\
&\overset{(1)}{\simeq} r_3\circ J(N_YX,N_YX\mid Z)\circ(N(Y,X)\to Y)^*\circ r_2\circ J(X,Y)\\
&\simeq r_3\circ J(N_YX,N_YX\mid Z)\circ J(X,Y)\\
&\overset{(2)}{\simeq} r_3\circ J(N_ZX,N_ZY)\circ J(X,Z)\\
&\simeq r_3\circ J(N_ZX,N_ZY)\circ(N_ZX\to Z)^*\circ r_4\circ J(X,Z)\\
&\overset{(3)}{=} r_3\circ\big(N(N_ZX,N_ZY)\to Z\big)^*\circ r_4\circ J(X,Z)\\
&\simeq r_5\circ J(X,Z)\\
&\simeq \bar J(\Delta_3). \qquad\square
\end{aligned}$$

The square Δ is called admissible if η is flat, ρ is smooth and X' is smooth over B. Consider the diagram

$$\begin{array}{ccccc}
Y & \xrightarrow{i} & Y\times X & \xrightarrow{\rho} & X\\
\big\downarrow{\scriptstyle\eta} & & \big\downarrow & & \big\downarrow{\scriptstyle\rho}\\
Y' & \xrightarrow{i'} & Y'\times X' & \xrightarrow{\rho'} & X'
\end{array}$$

and denote the left square by Δ_i and the right square by Δ_p. If Δ is admissible, then Δ_i is a regular imbedding and Δ_p is a flat square (see (8.5.0)). Moreover the normal bundles of i and i' are given by f^*TX and f'^*TX', respectively, and N_{Δ_i} is given by $f^*T_{X'}X$.

Let Δ be admissible and let τ and τ' be coordinations of $T_{X'}X$ and TX', respectively. We define
$$\bar I(\Delta)\colon X \bullet\!\!\to Y$$
by
$$\bar I(\Delta)=\bar I(\Delta,\tau,\tau')=\bar J(\Delta_i,f^*\tau,f'^*\tau')\circ p^*.$$
One has $\delta\big(\bar I(\Delta)\big)=0$ and
$$\bar I(\Delta)\big(C_{p,l}(\rho)\big)\subset C_{p+s,l+t}(\eta)$$
with $s=\dim_B Y'-\dim_B X'$ and $t=\dim_B Y-\dim_B X$. Moreover the homotopy class of $\bar I(\Delta)$ (with respect to the degree (s,t)) does not depend on the choice of τ and τ'.

(13.2) Theorem. *If in* (13.1.0) *the squares* Δ_1 *and* Δ_2 *are admissible, then the square* Δ_3 *is admissible and*

$$\bar{I}(\Delta_3) \cong \bar{I}(\Delta_1) \circ \bar{I}(\Delta_2)$$

under a filtration preserving homotopy.

Proof: The first statement is trivial. For the second we consider the diagrams

$$\begin{array}{ccccc} Z & \xrightarrow{i_1} & Z \times Y & \xrightarrow{i_4} & (Z \times Y) \times X \\ & & \downarrow{\scriptstyle p_1} & & \downarrow{\scriptstyle p_4} \\ & & Y & \xrightarrow{i_2} & Y \times X \\ & & & & \downarrow{\scriptstyle p_2} \\ & & & & X, \end{array}$$

$$Z \xrightarrow{i_3} Z \times X \xrightarrow{i_5} (Z \times Y) \times X \xrightarrow{p_5} X.$$

The regular imbeddings i_j lie over accordingly defined regular imbeddings i'_j; the corresponding squares are denoted by Σ_j.

The Σ_j are regular imbeddings. An application of Lemma 11.3 and Theorem 13.1 shows that (by noting $i_4 \circ i_1 = i_5 \circ i_3$ and $p_2 \circ p_4 = p_5$)

$$\bar{I}(\Delta_1) \circ \bar{I}(\Delta_2) \simeq \bar{J}(\Sigma_3) \circ \bar{J}(\Sigma_5) \circ p_5^*.$$

By definition we have

$$\bar{I}(\Delta_3) \simeq \bar{J}(\Sigma_3) \circ (p_5 \circ i_5)^*.$$

Finally Lemma 11.4 shows

$$\bar{J}(\Sigma_5) \circ p_5^* = (p_5 \circ i_5)^*. \qquad \square$$

Theorem 13.2 implies Theorem 12.1 by passing to homology. For a proof of Theorem 12.7 we consider six squares with the top arrows

$$\begin{array}{ccccccc} Y & \xrightarrow{i_1} & Y \times_{X'} X & \xrightarrow{i_2} & Y \times X & & \\ & & \downarrow{\scriptstyle p_1} & & \downarrow{\scriptstyle p_2} & & \\ & & Y' \times_{X'} X & \xrightarrow{i_3} & Y' \times X & \xrightarrow{p_3} & X \end{array}$$

lying over the bottom arrows

$$\begin{array}{ccccccc} Y' & = & Y' & \xrightarrow{i'} & Y' \times X' & & \\ & & \| & & \| & & \\ & & Y' & \xrightarrow{i'} & Y' \times X' & \xrightarrow{p'} & X'. \end{array}$$

Let Σ_j be the square corresponding to i_j for $j = 1, 2, 3$. The Σ_j are regular imbeddings. The map $f^\bullet$ is induced from $\bar{I}(\Delta)$. By the definition of $\bar{I}(\Delta)$ and Theorem 13.1 we have

$$\bar{I}(\Delta) \simeq \bar{J}(\Sigma_1) \circ \bar{J}(\Sigma_2) \circ p_2^* \circ p_3^*.$$

Lemma 11.3 shows that

$$\bar{J}(\Sigma_2) \circ p_2^* \simeq p_1^* \circ \bar{J}(\Sigma_3).$$

Therefore $f^\bullet$ is the composition of the maps induced by $\bar{J}(\Sigma_1) \circ p_1^*$ and by $\bar{J}(\Sigma_3) \circ p_3^*$.

Next note that $p_1 \circ i_1 = \hat{f}$. An application of Proposition 8.5 shows that $\bar{J}(\Sigma_1) \circ p_1^*$ induces on the E^2-terms the map $[\hat{f}/Y']_\#$. Finally note that $p' \circ i' = f'$ and that the squares under i_3 and p_3 are pull-back squares. An application of Proposition 8.5 shows that $\bar{J}(\Sigma_3) \circ p_3^*$ induces on the E^2-terms the map $(f')^\bullet$. $\square$

14. Products

In this section M is a cycle module over B and N is a cycle module over k. We assume that either $N = K_*$ or that $M = N$ is a cycle module with ring structure over $B = \operatorname{Spec} k$. So in any case we are given a pairing $N \times M \to M$ of cycle modules over B.

The restriction to these special cases are made for simplification. For example, in forming intersections of cycles with coefficients in a cycle module M with ring structure, one needs to know that its pairing factors through a cycle module over $B \times B$. However, the existence of a corresponding appropriate notion of tensor product of cycle modules is not clear to me (and a settling of this question would lead to far here anyway). The problem could be avoided in the following by assuming the necessary factorizations, but this is somewhat tiring.

(14.1) Cross products. Let Y be a scheme over k and let Z be a scheme over B (all of finite type over k). We define the cross product

$$\times : C_p(Y; N) \times C_q(Z; M) \to C_{p+q}(Y \times Z; M)$$

as follows. For $y \in Y$ let $Z_y = \operatorname{Spec} \kappa(y) \times Z$, let $\pi_y \colon Z_y \to Z$ be the projection and let $i_y \colon Z_y \to Y \times Z$ be the inclusion. For $z \in Z$ we understand similarly $Y_z, \pi_z \colon Y_z \to Y$ and $i_z \colon Y_z \to Y \times Z$. We give the following two equivalent definitions:

$$\rho \times \mu = \sum_{y \in Y_{(p)}} (i_y)_* \big(\rho_y \cdot \pi_y^*(\mu)\big),$$
$$\rho \times \mu = \sum_{z \in Z_{(q)}} (i_z)_* \big(\pi_z^*(\rho) \cdot \mu_z\big).$$

Here $\rho_y \in N(y)$ is the y-component of ρ and the product is understood after pointwise restriction of ρ_y. The map

$$(i_y)_* \colon C_q(Z_y; M) \to C_{p+q}(Y \times Z; M)$$

is the inclusion corresponding to $Z_{y(q)} \subset (Y \times Z)_{(p+q)}$. Similarly we understand $\mu_z \in M(z)$ and

$$(i_z)_*\colon C_p(Y_z; M) \to C_{p+q}(Y \times Z; M).$$

To check equality of the two definitions consider the u-components for $u \in Y \times Z$. Let y, z be the images of u under the projections $Y \times Z \to Y$, Z and let $R = \kappa(y) \otimes_k \kappa(z)$. Then the u-components are either trivial or u is a minimal prime of R. In the latter case the u-components are given by

$$(\rho \times \mu)_u = r_{\kappa(u)|\kappa(y)}(\rho_y) \cdot r_{\kappa(u)|\kappa(z)}(\mu_z) \cdot l_R(R_{(u)}).$$

(14.2) Associativity. Additionally let X be of finite type over k and let $\eta \in C_r(X; N)$. Then

$$\eta \times (\rho \times \mu) = (\eta \times \rho) \times \mu.$$

For a proof consider the u-components for $u \in X \times Y \times Z$. Let x, y, z be the images of u in X, Y, Z, respectively, and let $R = \kappa(x) \otimes_k \kappa(y) \otimes_k \kappa(z)$. Then the u-components are either trivial or u is a minimal prime of R. In the latter case it follows from standard rules for length that the u-components are given by

$$(\eta \times \rho \times \mu)_u = r_{\kappa(u)|\kappa(x)}(\eta_x) \cdot r_{\kappa(u)|\kappa(y)}(\rho_y) \cdot r_{\kappa(u)|\kappa(z)}(\mu_z) \cdot l_R(R_{(u)}).$$

(14.3) Commutativity. Suppose $M = N$ is a cycle module with ring structure over $B = \operatorname{Spec} k$. Let $\tau\colon Y \times Z \to Z \times Y$ be the interchange of factors. For $\rho \in C_p(Y; M, n)$ and $\mu \in C_q(Z; M, m)$ one has

$$\tau_*(\rho \times \eta) = (-1)^{nm} \eta \times \rho \in C_{p+q}(Z \times Y; M, n+m).$$

This is immediate from the definitions.

(14.4) Chain rule. For $\rho \in C_p(Y; N, n)$ and $\mu \in C_q(Z; M, m)$ one has

$$d(\rho \times \mu) = d(\rho) \times \mu + (-1)^n \rho \times d(\mu).$$

For a proof we may assume $\rho \in M(y)$, $\mu \in M(z)$ for some $y \in Y_{(p)}$ and $z \in Z_{(q)}$. Consider for $u \in Y \times Z$ the u-components of the three terms. If one of them is nontrivial, we must have $\dim(u, Y \times Z) = p + q - 1$ and the images y', z' of u must be in the closures of y, z, respectively. Dimension reasons show $y' = y$ or $z' = z$. Now the claim follows from one of the two definitions of the cross product and Proposition 4.6.2.

(14.5) Compatibility. The cross product is compatible with the four basic types of maps f_*, g^*, $\{a\}$ and ∂ acting on one of the two factors. This follows from the compatibility with flat pull-back and Definition 1.3. We omit a detailed formulation.

We conclude with a consideration of the intersection pairing for cycles on a smooth variety. Let X be smooth over k and let τ be a coordination of TX. We define

$$I_X\colon C^*(X; N) \times C^*(X; M) \to C^*(X; M),$$

$$I_X(\rho, \mu) = \big(r(\tau) \circ J(X \times X, X)\big)(\rho \times \mu).$$

By (14.4) this is a pairing of complexes. Let

$$\smile\colon A^*(X; N) \times A^*(X; M) \to A^*(X; M)$$

be the induced pairing.

The next theorem follows from the preceding remarks and in particular from Theorem 12.1. It holds accordingly on chain level up to homotopy.

(14.6) Theorem. *If $M = N$ is a cycle module with ring structure over $B = \operatorname{Spec} k$, the pairing $\smile$ turns $A^*(X;M)$ into an anti-commutative associative ring. If $N = K_*$, the pairing $\smile$ turns $A^*(X;M)$ into a module over $A^*(X;K_*)$.* □

We have defined in particular a ring structure on the classical Chow groups

$$\mathrm{CH}^*(X) = \coprod_p A^p(X;K_*,p)$$

of a smooth variety. This ring structure coincides with the classical one. This may be deduced from the remark at the beginning of Section 11 and (Fulton 1984, Chaps. 5, 6, 8).

References

Bass, H., Tate, J. (1972): The Milnor ring of a global field, in: Algebraic K-Theory II—"Classical" Algebraic K-Theory and Connections with Arithmetic, Proceedings, Battelle Institute 1972, *Lecture Notes in Math.* **342** (1972) 349–446

Bloch, S. (1986): Algebraic cycles and higher algebraic K-theory, *Adv. Math.* **61** (1986) 267–304

Bloch, S., Ogus, A. (1974): Gersten's conjecture and the homology of schemes, *Ann. scient. Èc. Norm. Sup.* **7** (1974) 181–202

Colliot-Thélène, J-L. (1992): Birational invariants, purity and the Gersten conjecture, K-Theory and algebraic geometry: connections with quadratic forms and division algebras (Santa Barbara, CA, 1992), *Proc. Sympos. Pure Math.*, **58**, Part 1, 1–64, Amer. Math. Soc., Providence, RI, 1995

Fulton, W. (1984): *Intersection Theory*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, 1984

Grayson, D. (1976): Higher Algebraic K-Theory II, in: Algebraic K-Theory, Evanston 1976, *Lecture Notes in Math.* **551** (1976) 217–240

Grayson, D. (1978): Products in K-Theory and Intersecting Algebraic Cycles, *Invent. Math.* **47** (1978) 71–83

Hartshorne, R. (1977): *Algebraic Geometry*, Graduate Texts in Mathematics, Springer, 1977

Hilton, P. J., Stammbach, U. (1971): *A Course in Homological Algebra*, Graduate Texts in Mathematics, Springer, 1971

Karpenko, N. A., Merkurjev, A. S. (1991): Chowgroups of projective quadrics, *Leningrad Math. J.* **2** No.3 (1991)

Kato, K. (1980): A generalization of local class field theory by using K-groups II, *J. Fac. of Sc.*, Univ. of Tokyo 1A **27** (1980) 603–683

Kato, K. (1986): Milnor K-theory and the Chowgroup of zero cycles, in: Applications of Algebraic K-Theory to Algebraic Geometry and Number Theory, Part I, Proceedings, *Contemp. Math.* **55** (1986) 241–253

Matsumura, H. (1980): *Commutative Algebra*, Second Edition, Math. Lecture Note Series, Benjamin/Cummings, Reading, 1980

Merkurjev, A. S., Suslin, A. A. (1982): K-cohomology of Severi-Brauer varieties and the norm residue homomorphism, *Izv. Akad. Nauk SSSR* **46** (1982) 1011–1046 [engl. transl. in: *Math. USSR Izv.* **21** (1983) 307–340]

Merkurjev, A. S., Suslin, A. A. (1986): On the norm residue homomorphism of degree three, preprint, LOMI, 1986 [also: *Math. USSR Izv.* **36** (1991) 541–565]

Merkurjev, A. S., Suslin, A. A. (1987): On the K_3 for a field, preprint, LOMI (1987) [also: *Math. USSR Izv.* **36** (1991) 541–565]

Milnor, J. (1970): Algebraic K-theory and quadratic forms, *Invent. Math.* **9** (1970) 318–344

Milnor, J. (1971): Introduction to Algebraic K-Theory, *Annals of Math. Studies* **72**, Princeton Univ. Press, 1971

Quillen, D. (1973): Higher algebraic K-Theory, in: Algebraic K-Theory I—Higher K-Theories, Proceedings, Battelle Institute 1972, *Lecture Notes in Math.* **341** (1973) 77–139

Rost, M. (1990): Durch Normengruppen definierte birationale Invarianten, *C. R. Acad. Sci. Paris* **310** Série I (1990) 189–192

Serre, J-P. (1957): Algèbre locale · Multiplicités, 1957/58, *Lecture Notes in Math.* **11** (1975)

Serre, J-P. (1968): *Corps Locaux*, Hermann, Paris, 1968

Serre, J-P. (1992): Résumé de cours 1991-1992, *Annuaire du Collège de France* (1992) 105–113 [reproduced in: Serre 1994]

Serre, J-P. (1994): Cohomologie Galoisienne, cinquième édition, *Lecture Notes in Math.* **5** (1994)

Shatz, S. (1972): Profinite groups, arithmetic and geometry, *Annals of Math. Studies* **67**, Princeton Univ. Press, 1972

Somekawa, M. (1990): On Milnor K-groups attached to Semi-Abelian Varieties, *K-Theory* **4** no.2 (1990) 105–119

Srinivas, V. (1991): *Algebraic K-Theory*, Progress in Math. **90** (1991)

Suslin, A. A., Voevodsky, V. (1996): Singular homology of abstract algebraic varieties, *Invent. Math.* **123** (1996) 61–94

Tate, J. (1976): Relations between K_2 and Galois Cohomology, *Invent. Math.* **36** (1976) 257–274

Fakultät für Mathematik
Universität Regensburg
D-93040 Regensburg
Germany
markus.rost@rphs1.physik.uni-regensburg.de

Applications of Weight-Two Motivic Cohomology

Bruno Kahn

Received: July 26, 1996

Communicated by Peter Schneider

Abstract. Using Lichtenbaum's complex $\Gamma(2)$, we reprove and extend a little bit some known results relating the kernel of $H^3(F,\mathbf{Q}/\mathbf{Z}(2)) \to H^3(F(X),\mathbf{Q}/\mathbf{Z}(2))$ to the torsion of CH^2X for rational varieties X.

1991 Mathematics Subject Classification: Primary 14C35, secondary 19E20.

Contents

Introduction

Let F be a field and X be a smooth, geometrically integral variety over F. In [6, prop. 3.6], Colliot-Thélène and Raskind produced an exact sequence:

$$(1)\quad H^1_{\mathrm{Zar}}(X,\mathcal{K}_2) \to H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)^{G_F} \\ \to H^1(F,K_2(\overline{F}(X))/H^0_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)) \to \mathrm{Ker}(CH^2X \to CH^2\overline{X}) \\ \to H^1(F,H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)) \to H^2(F,K_2(\overline{F}(X))/H^0_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)).$$

Here, $\overline{X}$ denotes the variety X viewed over the separable closure $\overline{F}$ of F, $\mathcal{K}_2$ is the Zariski sheaf associated to the presheaf $U \mapsto K_2(U)$ and G_F is the absolute Galois group of F. On the other hand, in [17, th. 3.1], we produced an isomorphism

$$(2)\quad H^1(F,K_2(\overline{F}(X))/K_2(\overline{F})) \simeq \mathrm{Ker}(H^3(F,\mathbf{Q}/\mathbf{Z}(2)) \to H^3(F(X),\mathbf{Q}/\mathbf{Z}(2))).$$

In (2), the coefficients $\mathbf{Q}/\mathbf{Z}(2)$ are

$$\varinjlim \mu_n^{\otimes 2} \text{ if } \operatorname{char} F = 0 \text{ and } \varinjlim_{(n,\operatorname{char} F)=1} \mu_n^{\otimes 2} \oplus \varinjlim_r W_r\Omega^2_{\log}[-2] \text{ if } \operatorname{char} F > 0,$$

where $W_r\Omega^2_{\log}$ is the weight-two logarithmic part of the de Rham-Witt complex over the big étale site of $\operatorname{Spec} F$ [13] (see comments at the end of the introduction).

When X is a *complete rational variety, i.e.* the extension $\overline{F}(X)/\overline{F}$ is purely transcendental, the group $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)$ coincides with $K_2(\overline{F})$. One may therefore replace the group $H^1(\overline{F}, K_2(F(X))/H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$ in (1) by $\operatorname{Ker}(H^3(F, \mathbf{Q}/\mathbf{Z}(2)) \to H^3(F(X), \mathbf{Q}/\mathbf{Z}(2)))$ in this case. The resulting exact sequence has been used in [29] and [30].

Moreover, the left map in (1) is injective when X is a complete rational variety ([6, prop. 4.3] in characteristic 0, [24, prop. 1.5] in general). Putting all this together, one therefore gets an exact sequence:

$$\begin{aligned} 0 \to H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) &\to H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} \\ &\to \operatorname{Ker}(H^3(F, \mathbf{Q}/\mathbf{Z}(2)) \to H^3(F(X), \mathbf{Q}/\mathbf{Z}(2))) \\ &\qquad \to \operatorname{Ker}(CH^2 X \to CH^2 \overline{X}) \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \end{aligned}$$

for any complete rational variety X.

In this paper, we use the Lichtenbaum complex $\Gamma(2)$ of [22], [23] to recover this exact sequence directly, and extend it to the right. Our main result is:

Theorem 1. *Let X be a smooth variety over F.*
a) Assume that $K_2(\overline{F}) \xrightarrow{\sim} H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)$. Let us denote by

$$\eta : H^3(F, \mathbf{Q}/\mathbf{Z}(2)) \to H^0_{\mathrm{Zar}}(X, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2)))$$
$$\xi : CH^2 X \to (CH^2 \overline{X})^{G_F}$$
$$\mathrm{cl}^2_X : CH^2 X \otimes \mathbf{Q}/\mathbf{Z} \to H^4(X, \mathbf{Q}/\mathbf{Z}(2))$$

the natural maps and the divisible cycle class map. Then there is an exact sequence

$$(3) \quad 0 \to H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \to H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} \to \operatorname{Ker} \eta \to \operatorname{Ker} \xi \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)).$$

b) Assume moreover that $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2)))$ is p-primary torsion, where p is the characteristic exponent of F and $\mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))$ is the Zariski sheaf associated to the presheaf $U \mapsto H^3_{\text{ét}}(U, \mathbf{Q}/\mathbf{Z}(2))$ (if $\operatorname{char} F = 0$, this means $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))) = 0$). Then the exact sequence (3) *extends to a complex*

$$(4) \qquad \operatorname{Ker} \xi \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \to H^4(F, \mathbf{Q}/\mathbf{Z}(2)) \to \operatorname{Coker} \mathrm{cl}^2_X .$$

Let A (resp. B) denote the homology of (4) *at $H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$ (resp. at $H^4(F, \mathbf{Q}/\mathbf{Z}(2))$). Then there is another complex*

$$(5) \quad 0 \to \operatorname{Coker} \eta \otimes \mathbf{Z}[1/p] \to \operatorname{Coker} \xi \otimes \mathbf{Z}[1/p] \to H^2(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \otimes \mathbf{Z}[1/p]$$

whose homology at $\operatorname{Coker} \eta \otimes \mathbf{Z}[1/p]$ (resp. at $\operatorname{Coker} \xi \otimes \mathbf{Z}[1/p]$) is $A \otimes \mathbf{Z}[1/p]$ (resp. $B \otimes \mathbf{Z}[1/p]$).
If $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))) = 0$, we can remove $\otimes \mathbf{Z}[1/p]$ everywhere.

REMARK. The assumptions are satisfied if X is a complete rational variety, but also if it is a torsor under a semi-simple, simply connected algebraic group [7]. If $\operatorname{char} k = p > 0$, in the second case the group $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))$ is in general nonzero, as higher logarithmic Hodge-Witt cohomology is not homotopy invariant; hence the complicated statement of theorem 1. However, we do have $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2)) = 0$ in the first case (compare corollaries 5.3 and 6.2 c)).

COROLLARY. *Let X be as in theorem 1 b).*
1) Suppose $\operatorname{cd} F \leq 3$. *Then there is an exact sequence*

$$0 \to H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \to H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} \to \operatorname{Ker} \eta \to \operatorname{Ker} \xi \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \\ \to \operatorname{Coker} \eta \to \operatorname{Coker} \xi \to H^2(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$$

after tensorisation by $\mathbf{Z}[1/p]$. *The part of this sequence up to* $H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$ *exists and is exact without tensoring by* $\mathbf{Z}[1/p]$.
2) Suppose $\operatorname{cd} F \leq 2$. *Then there is an isomorphism*

$$H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \xrightarrow{\sim} H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F}$$

and an exact sequence

$$0 \to \operatorname{Ker} \xi \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \\ \to H^0_{\mathrm{Zar}}(X, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))) \to \operatorname{Coker} \xi \to H^2(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$$

after tensorisation by $\mathbf{Z}[1/p]$. *The injection* $\operatorname{Ker} \xi \hookrightarrow H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))$ *holds without tensoring by* $\mathbf{Z}[1/p]$.
If $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))) = 0$, *the results hold without tensoring by* $\mathbf{Z}[1/p]$.

To try and get a relationship between theorem 1 and the last term in (1), we observe that a closer examination of the spectral sequence used in [17, proof of th. 3.1] yields an exact sequence:

$$(6) \quad H^3(F, \mathbf{Q}/\mathbf{Z}(2)) \to \operatorname{Ker}(H^3(F(X), \mathbf{Q}/\mathbf{Z}(2)) \to H^3(\overline{F}(X), \mathbf{Q}/\mathbf{Z}(2))) \\ \to H^2(F, K_2(F(X))/K_2(F)) \to H^4(F, \mathbf{Q}/\mathbf{Z}(2)) \to H^4(F(X), \mathbf{Q}/\mathbf{Z}(2)).$$

How to derive theorem 1 from sequence (6) does not seem obvious, however.

This paper is organized as follows. In section 1, we compute the étale hypercohomology of X with coefficients in $\Gamma(2)$: this is done in theorem 1.1, which is of independent interest. In sections 2 and 3, we introduce two relative complexes $\Gamma(F(X)/X, 2)$ (over $X_{\text{ét}}$) and $\Gamma(X/F, 2)$ (over $(\operatorname{Spec} F)_{\text{ét}}$). Considering the Hochschild-Serre spectral sequence for the hypercohomology of $\Gamma(F(X)/X, 2)$, we get back the Colliot-Thélène-Raskind exact sequence (1) in a straightforward manner (see proposition 2.2). To prove theorem 1, we similarly examine the Hochschild-Serre spectral sequence for the hypercohomology of X with coefficients $\Gamma(X/F, 2)$ (see section 3). In sections 4, 5 and 6, we respectively prove a purity theorem, compute the motivic cohomology of a projective bundle and prove a Bloch-Ogus type theorem. Finally, in section 7, we look at projective homogeneous varieties.

The proof of the isomorphism (2) in [17] consisted of considering the Hochschild-Serre spectral sequence for the hypercohomology of F with coefficients in a relative

Lichtenbaum complex $\Gamma(F(X)/F, 2)$, relative to the extension $\overline{F}/F$. What we do here can be considered as a refinement of this method, by factoring the morphism $\operatorname{Spec} F(X) \to \operatorname{Spec} F$ into

$$\operatorname{Spec} F(X) \to X \to \operatorname{Spec} F.$$

REMARKS ON CHARACTERISTIC p. We have to be a little careful if $\operatorname{char} F > 0$ when defining the coefficients $\mathbf{Q}/\mathbf{Z}(2)$. In characteristic 0, they are defined as $\varinjlim \mu_n^{\otimes 2}$. If $\operatorname{char} F = p > 0$, we set $\mathbf{Z}/p^r(2) = W_r\Omega^2_{\log}[-2]$, where $W_r\Omega^2_{\log}$ is the sheaf of logarithmic de Rham-Witt differentials over the big étale site of $\operatorname{Spec} F$, defined as the subsheaf of the de Rham-Witt sheaf $W_r\Omega^2$ generated locally for the étale topology by sections of the form $d\log \underline{x}_1 \wedge d\log \underline{x}_2$ [13, I.5.7]. So $\mathbf{Z}/p^r(2)$ is a complex of étale sheaves concentrated in degree 2. The Verlagerung maps $V : W_n\Omega^2 \to W_{n+1}\Omega^2$ preserve logarithmic differentials, hence can be used to define $\mathbf{Q}_p/\mathbf{Z}_p(2)$ as $\varinjlim_r \mathbf{Z}/p^r(2)$. Corollaires I.3.5 and I.5.7.5 of [13] yield exact sequences of étale sheaves

$$0 \to \mathbf{Z}/p^r(2) \xrightarrow{V^s} \mathbf{Z}/p^{r+s}(2) \to \mathbf{Z}/p^s(2) \to 0 \tag{7}$$

hence exact sequences

$$0 \to \mathbf{Z}/p^r(2) \to \mathbf{Q}_p/\mathbf{Z}_p(2) \xrightarrow{p^r} \mathbf{Q}_p/\mathbf{Z}_p(2) \to 0. \tag{8}$$

We now define $\mathbf{Q}/\mathbf{Z}(2)$ as $\varinjlim_{(n,\operatorname{char} F)=1} \mu_n^{\otimes 2} \oplus \mathbf{Q}_p/\mathbf{Z}_2(2)$. We sometimes abbreviate $\mathbf{Q}/\mathbf{Z}(2)$ by '2'.

NOTATION. We denote by $\Gamma_{\mathrm{Zar}}(2)$ (*resp.* $\Gamma_{\text{ét}}(2)$) the complex of sheaves over the big Zariski (*resp.* étale) site of $\operatorname{Spec} F$ associated to the presheaf $U \mapsto \Gamma(U, 2)$ of [22]. When necessary, we denote by $\Gamma_{\mathrm{Zar}}(X, 2)$ (*resp.* $\Gamma_{\text{ét}}(X, 2)$) the restriction of $\Gamma_{\mathrm{Zar}}(2)$ (*resp.* $\Gamma_{\text{ét}}(2)$) to the small Zariski (*resp.* étale) site of a scheme X. We drop indices when the context makes it clear what site we are in.

1. MOTIVIC COHOMOLOGY OF SMOOTH VARIETIES

Let X be a smooth, connected variety over a field F. We compute the étale hypercohomology groups $\mathbb{H}^*_{\text{ét}}(X, \Gamma(2)) = \mathbb{H}^*_{\text{ét}}(X, \Gamma_{\text{ét}}(2))$:

1.1. THEOREM. *$\mathbb{H}^i_{\text{ét}}(X, \Gamma(2))$ is*

(i) 0 *for* $i \leq 0$.
(ii) $K_3(F(X))_{\mathrm{ind}}$ *for* $i = 1$.
(iii) $H^0_{\mathrm{Zar}}(X, \mathcal{K}_2)$ *for* $i = 2$.
(iv) $H^1_{\mathrm{Zar}}(X, \mathcal{K}_2)$ *for* $i = 3$
(v) $\operatorname{Coker} \mathrm{cl}^2_X$ *for* $i = 5$
(vi) $H^{i-1}_{\text{ét}}(X, \mathbf{Q}/\mathbf{Z}(2))$ *for* $i \geq 6$

where cl^2_X is defined in theorem 1. Moreover, for $i = 4$ there is a short exact sequence:

$$0 \to CH^2 X \to \mathbb{H}^4_{\text{ét}}(X, \Gamma(2)) \to H^0_{\mathrm{Zar}}(X, \mathcal{H}^3(\mathbf{Q}/\mathbf{Z}(2))) \to 0. \tag{9}$$

As an immediate application, we get:

1.2. COROLLARY. *In characteristic* 0*, weight-two étale motivic cohomology is homotopy invariant. In characteristic* >0*, this is still true up to (cohomological) degree* 3.

To prove theorem 1.1, we shall use the Leray spectral sequence

$$E_2^{p,q} = H^p_{\mathrm{Zar}}(X, R^q\alpha_*\Gamma(2)) \Longrightarrow \mathbb{H}^{p+q}_{\text{ét}}(X, \Gamma(2)) \tag{10}$$

associated to the change-of-sites map $\alpha : X_{\text{ét}} \to X_{\mathrm{Zar}}$. For the convenience of the reader, we prove a well-known general lemma:

1.3. Lemma. *Let $\eta \xrightarrow{j} X$ be the generic point of the irreducible normal scheme X, and let A be an étale sheaf over η. Then the cohomology groups $H^q_{\text{ét}}(X, j_*A)$ are torsion for all $q > 0$.*

Proof. Let $\eta = \operatorname{Spec} K$. Consider the Leray spectral sequence for j

$$E_2^{p,q} = H^p_{\text{ét}}(X, R^q j_*A) \Longrightarrow H^{p+q}_{\text{ét}}(K, A).$$

Since the abutment is Galois cohomology, it is torsion for $p + q > 0$ and we have to prove that $R^q j_*A$ is torsion for all $q > 0$. But since X is normal, it is geometrically unibranch and the stalks of $R^q j_*A$ are Galois cohomology of the strict Henselizations of K relatively to the points of X, hence the claim. □

1.4. Lemma. *The Zariski sheaves $R^q\alpha_*\Gamma(2)$ are as follows:*

(i) 0 *for* $q \le 0$.
(ii) *The constant sheaf* $K_3(F(X))_{\mathrm{ind}}$ *for* $q = 1$.
(iii) $\mathcal{K}_2$ *for* $q = 2$.
(iv) 0 *for* $q = 3$.
(v) $\mathcal{H}^{q-1}(\mathbf{Q}/\mathbf{Z}(2))$ *for* $q \ge 4$.

Proof. (i) is obvious, (iii) is proved in [23, th. 2.10]) and (ii) (*resp.* (iv)) is proved in [23, prop. 2.11] (*resp.* in [23, prop. 2.12]) but only up to 2-torsion. This partially comes from the insistence to deal with $gr^2_\gamma K_3$ rather than with $K_{3,\mathrm{ind}}$. We give proofs of (ii), (iv) and (v).

Denote by $\mathbf{K}_{3,\mathrm{ind}}$ (*resp.* $\mathbf{H}^1(\Gamma(2))$) the étale sheaf associated to the presheaf $R \mapsto K_3(R)_{\mathrm{ind}}$ (*resp.* $R \mapsto H^1(\Gamma(R,2))$) for étale $\operatorname{Spec} R \to X$. Let $x \in X$. We claim that there is a chain of isomorphisms

$$\begin{aligned} H^1_{\text{ét}}(\mathcal{O}_{X,x}, \Gamma(2)) \xrightarrow{\sim} H^0_{\text{ét}}(\mathcal{O}_{X,x}, \mathbf{H}^1(\Gamma(2))) \xleftarrow{\sim} H^0_{\text{ét}}(\mathcal{O}_{X,x}, \mathbf{K}_{3,\mathrm{ind}}) \\ \xrightarrow{\sim} H^0_{\text{ét}}(K, \mathbf{K}_{3,\mathrm{ind}}) \xleftarrow{\sim} K_3(K)_{\mathrm{ind}}. \end{aligned} \tag{11}$$

The first isomorphism (from the left) simply comes from the fact that $\mathbf{H}^i(\Gamma(2)) = 0$ for $i \le 0$. The last one is proven in [26, prop. 11.4] (see also [21, th. 4.13]). By [16, theorem], if A is a local ring of a smooth variety, then $K_3(A)_{\mathrm{ind}} \to K_3(K)_{\mathrm{ind}}$ is bijective, where K is the field of fractions of A. Letting $j : \operatorname{Spec} K \hookrightarrow X$ be the inclusion of the generic point, this shows that the map $\mathbf{K}_{3,\mathrm{ind}} \to j_*j^*\mathbf{K}_{3,\mathrm{ind}}$ is an isomorphism, hence the third isomorphism in (11). Finally, by [22, prop. 1.8], for any local ring A whose residue field contains more than 2 elements, there is a surjection

$$K_3(A)_{\mathrm{ind}} \twoheadrightarrow H^1(\Gamma(A,2))$$

which is bijective if A is a field. Therefore, the commutative diagram

$$\begin{tikzcd} K_3(\mathcal{O}^{sh}_{X,x})_{\mathrm{ind}} \arrow[r, two heads] \arrow[d, "\wr"'] & H^1(\Gamma(\mathcal{O}^{sh}_{X,x}, 2)) \arrow[d] \\ K_3(K^{sh}_x)_{\mathrm{ind}} \arrow[r, "\sim"] & H^1(\Gamma(K^{sh}_x, 2)) \end{tikzcd}$$

where $\mathcal{O}^{sh}_{X,x}$ is the strict Henselisationes of $\mathcal{O}_{X,x}$ and K^{sh}_x is its field of fractions, shows that $K_3(\mathcal{O}^{sh}_{X,x})_{\rm ind} \to H^1(\Gamma(\mathcal{O}^{sh}_{X,x}, 2))$ is an isomorphism (we used [16] again for the left vertical isomorphism). This proves the second isomorphism in (11), which proves lemma 1.4 (ii).

We note that (iv) follows from (iii), the Merkurjev-Suslin theorem for the local rings of X [22, th. 9.1], the fact that $R^3\alpha_*\Gamma(2)$ is torsion [22, th. 9.2] and the triangles

$$
\begin{array}{ccccccc}
 & \Gamma(2) & \xrightarrow{n} & \Gamma(2) & \qquad & \Gamma(2) \xrightarrow{p^r} \Gamma(2) & \\
(12) & \nwarrow & & \swarrow & & \nwarrow \qquad \swarrow & \\
 & & \mu_n^{\otimes 2} & & & \mathbf{Z}/p^r(2) &
\end{array}
$$

in the derived category (the second triangle in the case $\operatorname{char} F = p > 0$). The first triangle is proven exact in [22] and [23] only for n odd, relying on the computation of torsion and cotorsion in $\mathbf{K}_{3,\rm ind}$ [22, lemma 8.2]. However, the proof goes through just as well for n even by using the isomorphism from [16] already mentioned. The second triangle is proven exact in [23, lemma 2.7] only for $r = 1$ and $p > 2$ (this fact was overlooked in [17]). However, the proof of [23, lemma 2.7] carries over in the same way, using (ii) and the Bloch-Gabber-Kato isomorphism $K_2(E)/p^r \xrightarrow{\sim} W_r\Omega^2_{E,\log}$ for any field E of characteristic p [2, cor. 2.8].

Finally, let us prove (v). By the triangle (12), we have a long exact sequence of Zariski sheaves

$$\cdots \to R^{i-1}\alpha_*\Gamma(2)\otimes \mathbf{Q} \to R^{i-1}\alpha_*\mathbf{Q}/\mathbf{Z}(2) \to R^i\alpha_*\Gamma(2) \to R^i\alpha_*\Gamma(2)\otimes\mathbf{Q} \to \dots$$

so that it is enough to see that $R^i\alpha_*\Gamma(2)$ is torsion for $i \geq 3$. For $i = 3$, this is (iv). For $i > 3$, we have a long exact sequence of sheaves

$$\cdots \to R^{i-1}\alpha_*\mathbf{K}_{3,\rm ind} \to R^i\alpha_*\Gamma(2) \to R^{i-2}\alpha_*\mathbf{K}_2 \to \dots$$

so it is enough to see that $R^i\alpha_*\mathbf{K}_{3,\rm ind}$ and $R^i\alpha_*\mathbf{K}_2$ are torsion for $i > 0$. In view of the isomorphism (see above)

$$\mathbf{K}_{3,\rm ind} \xrightarrow{\sim} j_*j^*\mathbf{K}_{3,\rm ind}$$

the first one follows from lemma 1.3. We are left with proving that $R^i\alpha_*\mathbf{K}_2$ is torsion for $i > 0$. As in [23, proof of lemma 2.2], we have a "Gersten resolution"

$$0 \to \mathbf{K}_2 \to j_*K_{2,K} \to \coprod_{x\in X^{(1)}} i_x^*\mathbb{G}_m \to \coprod_{x\in X^{(2)}} i_x^*\mathbf{Z} \to 0.$$

This complex of étale sheaves is not exact, but up to torsion it is. Therefore, up to torsion, there is a spectral sequence of Zariski sheaves

$$E_1^{p,q} = R^q\alpha_*C^p \Longrightarrow R^{p+q}\alpha_*\mathbf{K}_2$$

where C^p is the p-th term of the above "resolution" of $\mathbf{K}_2$. Since C^0 is of the form $j_*\mathcal{F}$, the same argument as above shows that $E_1^{0,q}$ is torsion for $q > 0$. The stalks of $E_1^{1,q}$ and $E_1^{2,q}$ are sums of Galois cohomology groups, so are torsion for $q > 0$. This shows that $E_2^{p,q}$ is torsion for $p+q>0$, except perhaps when $q=0$. But, for $x\in X$, the stalks of $E_2^{1,0}$ and $E_2^{2,0}$ at x are the cohomology groups of the complex

$$H^0(K,K_2(\overline{K}))\to\coprod_{y\in Y^{(1)}}F(y)^*\to\coprod_{y\in Y^{(2)}}\mathbf{Z}\to 0 \tag{13}$$

where $Y=\operatorname{Spec}\mathcal{O}_{X,x}$. Comparing with the exact sequence (Gersten's conjecture)

$$K_2(K)\to\coprod_{y\in Y^{(1)}}F(y)^*\to\coprod_{y\in Y^{(2)}}\mathbf{Z}\to 0$$

and using the fact that the map $K_2(K)\to H^0(K,K_2(\overline{K}))$ has torsion kernel and cokernel, we get that (13) has torsion cohomology groups, which concludes the proof of lemma 1.4 (v). □

Proof of theorem 1.1. As indicated above, we use the spectral sequence (10). (i) is obvious in view of lemma 1.4 (i) and so is (ii) in view of the isomorphism

$$\mathbb{H}^1_{\text{ét}}(X,\Gamma(2))\xrightarrow{\sim}H^0_{\text{Zar}}(X,R^1\alpha_*\Gamma(2))$$

and lemma 1.4 (ii). To get further, we observe that $E_2^{p,1}=0$ for $p>0$ since $R^1\alpha_*\Gamma(2)$ is constant, and $E_2^{p,3}=0$ for all p in view of lemma 1.4 (iv). This and lemma 1.4 (iii) immediately imply (iii) and (iv). Still by lemma 1.4 (iii) and Gersten's conjecture, $E_2^{p,2}=0$ for $p>2$ and $E_2^{2,2}\simeq CH^2X$; this and lemma 1.4 (v) (for $q=4$) gives the exact sequence (9). We now note that the above information and lemma 1.4 (v) imply that $\mathbb{H}^i_{\text{ét}}(X,\Gamma(2))$ is torsion for $i\geq 5$. (v) and (vi) now follow from (9) and the long exact sequence

$$\cdots\to\mathbb{H}^{i-1}_{\text{ét}}(X,\Gamma(2))\otimes\mathbf{Q}\to H^{i-1}_{\text{ét}}(X,\mathbf{Q}/\mathbf{Z}(2))\to\mathbb{H}^i_{\text{ét}}(X,\Gamma(2))\to\mathbb{H}^i_{\text{ét}}(X,\Gamma(2))\otimes\mathbf{Q}\to\cdots$$

□

1.5. Remark. The same computation gives the cohomology sheaves of $\Gamma_{\text{Zar}}(X,2)$:

$$\begin{aligned}\mathcal{H}^1(\Gamma_{\text{Zar}}(X,2))&=K_3(K)_{\text{ind}}\\ \mathcal{H}^2(\Gamma_{\text{Zar}}(X,2))&=\mathcal{K}_2\\ \mathcal{H}^i(\Gamma_{\text{Zar}}(X,2))&=0\,\text{for}\,i\neq 1,2.\end{aligned}$$

From this, we deduce a triangle, precising [23, prop. 3.1]:

$$\begin{array}{ccc}\Gamma_{\text{Zar}}(2) & \longrightarrow & R\alpha_*\Gamma_{\text{ét}}(2)\\ \nwarrow & & \swarrow\\ & \tau_{\geq 3}(R\alpha_*\mathbf{Q}/\mathbf{Z}(2))[-1] & \end{array}$$

In particular,

$$\Gamma_{\text{Zar}}(2)\otimes\mathbf{Q}\xrightarrow{\sim}R\alpha_*\Gamma_{\text{ét}}(2)\otimes\mathbf{Q}. \tag{14}$$

We also get the following analogue of theorem 1.1:

1.6. Theorem. $\mathbb{H}^i_{\text{Zar}}(X,\Gamma_{\text{Zar}}(2))=\begin{cases}K_3(K)_{\text{ind}} & \textit{if } i=1\\ H^{i-2}_{\text{Zar}}(X,\mathcal{K}_2) & \textit{if } 2\leq i\leq 4\\ 0 & \textit{otherwise.}\end{cases}$ □

2. Relative motivic cohomology, I

Let $j : \operatorname{Spec} F(X) \hookrightarrow X$ be the inclusion of the generic point and $\Gamma(F(X)/X, 2)$ be the homotopy fibre of the morphism

$$\Gamma_{\text{ét}}(X, 2) \to Rj_*\Gamma_{\text{ét}}(F(X), 2).$$

Denote the hypercohomology group $\mathbb{H}^i_{\text{ét}}(X, \Gamma(F(X)/X, 2))$ by $\mathbb{H}^i(F(X)/X, \Gamma(2))$, so that we have a long exact sequence

$$\to \mathbb{H}^i(F(X)/X, \Gamma(2)) \to \mathbb{H}^i_{\text{ét}}(X, \Gamma(2)) \to \mathbb{H}^i_{\text{ét}}(F(X), \Gamma(2)) \to \mathbb{H}^{i+1}(F(X)/X, \Gamma(2)) \to$$

This gives:

2.1. Lemma. *The groups $\mathbb{H}^i(F(X)/X, \Gamma(2))$ are 0 for $i \leq 2$; there are exact sequences:*

$$0 \to K_2(F(X))/H^0_{\text{Zar}}(X, \mathcal{K}_2) \to \mathbb{H}^3(F(X)/X, \Gamma(2)) \to H^1_{\text{Zar}}(X, \mathcal{K}_2) \to 0$$

$$\mathbb{H}^4(F(X)/X, \Gamma(2)) \xrightarrow{\sim} CH^2 X$$

$$(15) \quad 0 \to H^0_{\text{Zar}}(X, \mathcal{H}^3(2)) \to H^3_{\text{ét}}(F(X), 2) \to \mathbb{H}^5(F(X)/X, \Gamma(2)) \to \operatorname{Coker} \mathrm{cl}^2_X \to H^4_{\text{ét}}(F(X), 2).$$

Proof. The first claim is clear for $i \leq 0$; for $i = 1$ and 2 it follows from theorem 1.1 and the injectivity of $H^0_{\text{Zar}}(X, \mathcal{H}^2) \to K_2(F(X))$. For $i = 3$, it follows from theorem 1.1 again, plus the vanishing of $\mathbb{H}^3(F(X), \Gamma(2))$. For $i = 4, 5$, we have a cross of exact sequences:

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & \mathbb{H}^4(F(X)/X, \Gamma(2)) \arrow[d] & & \\
0 \arrow[r] & CH^2 X \arrow[r] & \mathbb{H}^4_{\text{ét}}(X, \Gamma(2)) \arrow[r] \arrow[d] & H^0_{\text{Zar}}(X, \mathcal{H}^3(2)) \arrow[r] & 0 \\
 & & H^3_{\text{ét}}(F(X), 2) \arrow[d] & & \\
 & & \mathbb{H}^5(F(X)/X, \Gamma(2)) \arrow[d] & & \\
 & & \operatorname{Coker} \mathrm{cl}^2_X \arrow[d] & & \\
 & & H^4_{\text{ét}}(F(X), 2) & &
\end{tikzcd}$$

The map $\mathbb{H}^4_{\text{ét}}(X, \Gamma(2)) \to H^3_{\text{ét}}(F(X), 2)$ factors through $H^0_{\text{Zar}}(X, \mathcal{H}^3(2)) \to H^3_{\text{ét}}(F(X), 2)$, which is injective. A diagram chase concludes the proof. $\square$

For simplicity, let us denote by $\overline{K}_2(F(X))$ the group $K_2(F(X))/H^0_{\rm Zar}(X,\mathcal{K}_2)$. Using the "Hochschild-Serre" (hypercohomology) spectral sequence

$$H^p_{\text{ét}}(F,\mathbb{H}^q(\overline{F}(X)/\overline{X},\Gamma(2)))\Rightarrow \mathbb{H}^{p+q}(F(X)/X,\Gamma(2))$$

and the vanishing of $\mathbb{H}^i(F(X)/X,\Gamma(2))$ for $i\le 2$, we get an isomorphism

$$\mathbb{H}^3(F(X)/X,\Gamma(2))\xrightarrow{\sim} H^0(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2)))$$

and an 5-terms exact sequence

$$\begin{aligned}&0\to H^1(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2)))\to \mathbb{H}^4(F(X)/X,\Gamma(2))\\ &\to H^0(F,\mathbb{H}^4(\overline{F}(X)/\overline{X},\Gamma(2)))\to H^2(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2)))\to \mathbb{H}^5(F(X)/X,\Gamma(2))\end{aligned}$$

hence, using lemma 2.1:

2.2. Proposition. *There are exact sequences:*

$$\begin{aligned}0\to \overline{K}_2(F(X))\to \overline{K}_2(\overline{F}(X))^{G_F}\to H^1_{\rm Zar}(X,\mathcal{K}_2)\to H^1_{\rm Zar}(\overline{X},\mathcal{K}_2)^{G_F}\\ \to H^1(F,\overline{K}_2(F(X)))\to H^1(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2))\\ \to H^1(F,H^1_{\rm Zar}(\overline{X},\mathcal{K}_2))\to H^2(F,\overline{K}_2(\overline{F}(X)))\end{aligned}$$

$$\begin{aligned}0\to H^1(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2))\to CH^2X\to (CH^2\overline{X})^{G_F}\\ \to H^2(F,\mathbb{H}^3(\overline{F}(X)/\overline{X},\Gamma(2))\to \mathbb{H}^5(F(X)/X,\Gamma(2)).\end{aligned}$$

□

The exact sequence (1) follows immediately. Moreover, we also get [6, lemma 4.1].

3. Relative motivic cohomology, II

We recall some notation:

- As above, $H^i(X,j)$ (*resp.* $\mathcal{H}^i(j)$) is shorthand for $H^i_{\text{ét}}(X,\mathbf{Q}/\mathbf{Z}(j))$ (*resp.* for $\mathcal{H}^i(\mathbf{Q}/\mathbf{Z}(j))$).
- η is the map $H^3(F,2)\to H^0(X,\mathcal{H}^3(2))$.
- ξ is the map $CH^2X\to (CH^2\overline{X})^{G_F}$.

We also denote by $\overline{H}^0(X,\mathcal{K}_2)$ the group $H^0(X,\mathcal{K}_2)/K_2(F)$.

Let $\pi: X\to \operatorname{Spec} F$ be the structural morphism and $\Gamma(X/F,2)$ be the homotopy fibre (in the derived category) of the morphism

$$\Gamma_{\text{ét}}(F,2)\to R\pi_*\Gamma_{\text{ét}}(X,2).$$

Denote the hypercohomology group $\mathbb{H}^i_{\text{ét}}(F,\Gamma(X/F,2))$ by $\mathbb{H}^i(X/F,\Gamma(2))$, so that we have a long exact sequence

(16)

$$\cdots\to \mathbb{H}^i(X/F,\Gamma(2))\to \mathbb{H}^i_{\text{ét}}(F,\Gamma(2))\to \mathbb{H}^i_{\text{ét}}(X,\Gamma(2))\to \mathbb{H}^{i+1}(X/F,\Gamma(2))\to\cdots$$

This gives:

3.1. Lemma. *The groups* $\mathbb{H}^i(X/F,\Gamma(2))$ *are:*

(i) 0 *for* $i\le 1$.
(ii) $K_3(F(X))_{\rm ind}/K_3(F)_{\rm ind}$ *for* $i=2$.
(iii) $H^0_{\rm Zar}(X,\mathcal{K}_2)/K_2(F)$ *for* $i=3$.

Moreover, there is a complex

$$(17)\quad 0 \to H^1_{\mathrm{Zar}}(X,\mathcal{K}_2) \to \mathbb{H}^4(X/F,\Gamma(2)) \to \operatorname{Ker}\eta$$
$$\to CH^2X \to \mathbb{H}^5(X/F,\Gamma(2)) \to H^4(F,\mathbf{Q}/\mathbf{Z}(2)) \to \operatorname{Coker}\mathrm{cl}^2_X\,.$$

This complex is exact, except perhaps at $\mathbb{H}^5(X/F,\Gamma(2))$, *where its homology is* $\operatorname{Coker}\eta$. *In particular, we have an isomorphism*

$$H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2) \xrightarrow{\sim} \mathbb{H}^4(\overline{X}/\overline{F},\Gamma(2))$$

and a short exact sequence

$$(18)\qquad 0 \to CH^2\overline{X} \to \mathbb{H}^5(\overline{X}/\overline{F},\Gamma(2)) \to H^0_{\mathrm{Zar}}(\overline{X},\mathcal{H}^3(2)) \to 0.$$

Proof. (i), (ii) and (iii) immediately follow from theorem 1.1 and the exact sequence (16). The complex (17) and the value of its homology follow from the cross of exact sequences ((9) and (16))

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & H^1_{\mathrm{Zar}}(X,\mathcal{K}_2) \arrow[d] & & \\
 & & \mathbb{H}^4(X/F,\Gamma(2)) \arrow[d] & & \\
 & & H^3_{\text{ét}}(F,2) \arrow[d] \arrow[dr, "\eta"] & & \\
0 \arrow[r] & CH^2X \arrow[r] & \mathbb{H}^4_{\text{ét}}(X,\Gamma(2)) \arrow[r] \arrow[d] & H^0_{\mathrm{Zar}}(X,\mathcal{H}^3(2)) \arrow[r] & 0 \\
 & & \mathbb{H}^5(X/F,\Gamma(2)) \arrow[d] & & \\
 & & H^4_{\text{ét}}(F,2) \arrow[d] & & \\
 & & \operatorname{Coker}\mathrm{cl}^2_X & &
\end{tikzcd}$$

and the “lemma of the 700th” [27]. □

We now consider the hypercohomology spectral sequence

$$(19)\qquad H^p(F,\mathbb{H}^q(\overline{X}/\overline{F},\Gamma(2))) \Longrightarrow \mathbb{H}^{p+q}(X/F,\Gamma(2)).$$

Note that $E_2^{p,2}=0$ for $p>0$, since the group $K_3(F(X))_{\mathrm{ind}}/K_3(F)_{\mathrm{ind}}$ is uniquely divisible by [26, prop. 11.6]. Hence we get an isomorphism

$$\overline{H}^0_{\mathrm{Zar}}(X,\mathcal{K}_2) \xrightarrow{\sim} \overline{H}^0_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)^{G_F}$$

and an exact sequence

$$\begin{aligned}0 \to H^1(F, \overline{H}^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) &\to \mathbb{H}^4(X/F, \Gamma(2)) \\ &\to H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} \to H^2(F, \overline{H}^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2))\end{aligned}$$

(noting that $H^4(\overline{X}/\overline{F}, \Gamma(2)) = H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)$ by lemma 3.1). The isomorphism is Suslin's [32, cor. 5.9], but we get it here by a formal argument, in the vein of [17, th. 3.1 (a)]. The cross of complexes (the above exact sequence and (17)):

$$\begin{tikzcd}
 & & 0 \arrow[d] & & \\
 & & H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \arrow[d] & & \\
0 \arrow[r] & H^1(F, \overline{H}^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \arrow[r] & \mathbb{H}^4(X/F, \Gamma(2)) \arrow[r] \arrow[d] & H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} \arrow[r] & H^2(F, \overline{H}^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) \\
 & & \operatorname{Ker} \eta \arrow[d] & & \\
 & & CH^2 X \arrow[d] & & \\
 & & \mathbb{H}^5(X/F, \Gamma(2)) \arrow[d] & & \\
 & & H^4_{\text{ét}}(F, 2) \arrow[d] & & \\
 & & \operatorname{Coker} \mathrm{cl}^2_X & &
\end{tikzcd}$$

contains, via the lemma of the 700th, all the information one can easily get in this generality.

Assume now that $\overline{H}^0_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2) = 0$. Then the exact row in the above diagram reduces to an isomorphism $\mathbb{H}^4(X/F, \Gamma(2)) \xrightarrow{\sim} H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F}$, hence we get a complex:

$$\begin{aligned}(20)\quad 0 \to H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \to H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)^{G_F} &\to \operatorname{Ker} \eta \to CH^2 X \\ &\to \mathbb{H}^5(X/F, \Gamma(2)) \to H^4_{\text{ét}}(F, 2) \to \operatorname{Coker} \mathrm{cl}^2_X\end{aligned}$$

with homology $\operatorname{Coker} \eta$ at $\mathbb{H}^5(X/F, \Gamma(2))$ and 0 elsewhere.

Moreover the spectral sequence (19) and lemma 3.1 give an exact sequence

$$\begin{aligned}0 \to H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) &\to \mathbb{H}^5(X/F, \Gamma(2)) \\ &\to (\mathbb{H}^5(\overline{X}/\overline{F}, \Gamma(2)))^{G_F} \to H^2(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)).\end{aligned}$$

Putting (20) and () together, we get a cross of complexes (the horizontal one exact, the vertical one exact except perhaps at the crossing point):

$$
\begin{tikzcd}[column sep=small]
 & & 0 \arrow[d] & & \\
 & & H^1_{\mathrm{Zar}}(X,\mathcal{K}_2) \arrow[d] & & \\
 & & H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)^{G_F} \arrow[d] & & \\
 & & \operatorname{Ker}\eta \arrow[d] & & \\
 & & CH^2X \arrow[d] \arrow[dr, "\xi'"] & & \\
0 \arrow[r] & H^1(F,H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)) \arrow[r] & \mathbb{H}^5(X/F,\Gamma(2)) \arrow[r] \arrow[d] & (\mathbb{H}^5(\overline{X}/\overline{F},\Gamma(2)))^{G_F} \arrow[r] & H^2(F,H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)) \\
 & & H^4_{\text{ét}}(F,2) \arrow[d] & & \\
 & & \operatorname{Coker}\mathrm{cl}^2_X\,. & &
\end{tikzcd}
$$

Note that $\operatorname{Ker}\xi = \operatorname{Ker}\xi'$ by (18). We get theorem 1 a) from this cross and the latter remark, by a diagram chase analogous to the lemma of the 700th. The same diagram chase gives us the complex (4), and shows that its cohomology coincides with that of a complex

$$0 \to \operatorname{Coker}\eta \to \operatorname{Coker}\xi' \to H^2(F, H^1_{\mathrm{Zar}}(\overline{X},\mathcal{K}_2)).$$

Notice the short exact sequence from (18)

$$0 \to \operatorname{Coker}\xi \to \operatorname{Coker}\xi' \to H^0_{\mathrm{Zar}}(\overline{X},\mathcal{H}^3(2))^{G_F}.$$

Using this exact sequence, we easily conclude the proof of theorem 1. □

4. Purity

In this section, we establish a purity theorem for Zariski and étale weight-two motivic cohomology, generalizing results of [23]. Recall that $\Gamma(1)$ is defined as $\mathbb{G}_m[-1]$ and $\Gamma(0)$ as $\mathbf{Z}[0]$ (in both the Zariski and étale topologies). We also need such complexes for $i < 0$:

4.1. Definition. *For $i < 0$, we define:*

$$\Gamma_{\mathrm{Zar}}(i) = 0;$$
$$\Gamma_{\text{ét}}(i) = \mathbf{Q}/\mathbf{Z}(i)[-1] \qquad \textit{(no p-primary part in characteristic p).}$$

The following theorem extends and precises [23, th. 4.5]; the method of proof is different.

4.2. Theorem. *Let X be a smooth variety over a field and let $Z \xrightarrow{i} X$ be a closed immersion, with Z smooth of codimension c.*

a) There is an isomorphism (in the derived category of complexes of sheaves over Z_{Zar})

$$\Gamma_{\mathrm{Zar}}(Z,2-c)[-2c] \xrightarrow{\sim} Ri^!_{\mathrm{Zar}}\Gamma_{\mathrm{Zar}}(X,2).$$

b) There is a map (in the derived category of complexes of sheaves over $Z_{\text{ét}}$)

$$\Gamma_{\text{ét}}(Z,2-c)[-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X,2)$$

whose homotopy cofibre is concentrated in degree $c+4$ and has p-primary torsion cohomology, where p is the characteristic exponent of F. In particular, if $\operatorname{char} F = 0$, *this map is an isomorphism.*

4.3. Lemma. *Let $Z \overset{i}{\hookrightarrow} X$ be a smooth subvariety of X of codimension c. Then:*
a) For any constant sheaf A over X_{Zar}, $R^p i^!_{\mathrm{Zar}} A = 0$ for all p.
b) For any n, $R^p i^!_{\mathrm{Zar}}\mathcal{K}_n = \begin{cases} 0 & \text{for } p \neq c \\ \mathcal{K}_{n-c} & \text{for } p = c, \end{cases}$ where $\mathcal{K}_{n-c} := 0$ if $n < c$.

Proof. a) is trivial and b) follows in a well-known way from Gersten's conjecture (*e.g.* [9, § 7]). □

Proof of theorem 4.2 a). Apply $Ri^!$ to the triangle

$$\begin{array}{ccc} (\mathcal{K}_3)_{\mathrm{ind}}[-1] & \longrightarrow & \Gamma_{\mathrm{Zar}}(2) \\ & \nwarrow \quad \swarrow & \\ & \mathcal{K}_2[-2] & \end{array}$$

and apply lemma 4.3, noting that the Zariski sheaf $(\mathcal{K}_3)_{\mathrm{ind}}$ is constant.

For the proof of theorem 4.2 b), we need some facts on étale cohomological purity. For all $m \geq 1$, there is a morphism

$$\mathbf{Z}/m(2-c)[-2c] \to Ri^!_{\text{ét}}\mathbf{Z}/m(2). \tag{21}$$

For m prime to the characteristic exponent of F, this morphism is the classical purity isomorphism of SGA4, e.g. [28, th. 6.1]. For $\operatorname{char} F = p > 0$ and m a power of p, it is comes from Gros' thesis [10, II.3.5]: its homotopy cofibre is concentrated in degree $c+3$. In the general case, we define the morphism component-wise, on the prime-to-p and p-primary parts.

The following rather trivial lemma is very useful:

4.4. Lemma. *a) Let $f : S \to T$ be a morphism of sites and $Rf_* : \mathcal{D}^+(S) \to \mathcal{D}^+(T)$ the functor induced from the bounded below derived category of Abelian S-sheaves to that of Abelian T-sheaves. Let C be a bounded below complex of Abelian groups, that we view as a complex of constant sheaves over S. Then there is a natural isomorphism of functors*

$$Rf_* \circ (C \overset{L}{\otimes} ?) \approx C \overset{L}{\otimes} (Rf_* ?)$$

and a natural morphism of functors

$$f^* \circ (C \overset{L}{\otimes} ?) \to C \overset{L}{\otimes} (f^* ?).$$

b) Denote by i_λ ($\lambda = \mathrm{Zar}$ or $\mathrm{\acute{e}t}$) the map corresponding to i from Z_λ to X_λ (small sites). Then, with C as in a), there is a natural isomorphism of functors

$$Ri^!_\lambda \circ (C \overset{L}{\otimes} ?) \approx C \overset{L}{\otimes} (Ri^!_\lambda ?).$$

Proof. a) For A, B two Abelian groups, let $A * B$ denote $Tor_1^{\mathbf{Z}}(A, B)$. We note that $*$ is left exact and its unique nonzero higher derived functor is $R^1 * = \otimes$. Hence there is a natural isomorphism

$$C \overset{L}{\otimes} D \approx C \overset{R}{*} D[1]$$

for all $C, D \in \mathcal{D}(Ab)$.

Therefore the natural isomorphism of the lemma is equivalent to a natural isomorphism of functors

$$Rf_* \circ (C \overset{R}{*} ?) \approx C \overset{R}{*} (Rf_* ?)$$

which in turn will follow from a natural isomorphism

$$f_*(A * \mathcal{F}) \approx A * f_*\mathcal{F} \tag{22}$$

for any Abelian group A and any sheaf $\mathcal{F}$ over S. Note that, since $*$ is left exact, the presheaf $U \mapsto A * \mathcal{G}(U)$ is a sheaf for any sheaf $\mathcal{G}$ over any site. Therefore, given $U \in S$, both sides of (22) evaluated on U are $A * \mathcal{F}(f^{-1}(U))$. Finally, the second natural transformation, say, follows from the first one by adjunction.

b) Follows from a), considering the triangle of functors (with $j : X - Z \hookrightarrow X$ the complementary open immersion)

$$i_* Ri^! \to Id_{X_\lambda} \to Rj_* j^* \to i_* Ri^![1] \tag{23}$$

and the fact that i_* is fully faithful. Here we dropped the index $_\lambda$ for notational simplicity. □

Note that the triangle (12) and its analogues for $i = 0, 1$ can be reformulated as quasi-isomorphisms

$$\Gamma_{\mathrm{\acute{e}t}}(i) \overset{L}{\otimes} \mathbf{Z}/m \xrightarrow{\sim} \mathbf{Z}/m(i) \quad (0 \le i \le 2) \tag{24}$$

over the big étale site of $\operatorname{Spec} F$. Note also the obvious quasi-isomorphisms

$$\alpha^* \Gamma_{\mathrm{Zar}}(i) \xrightarrow{\sim} \Gamma_{\mathrm{\acute{e}t}}(i) \quad (0 \le i \le 2). \tag{25}$$

Using (24) and lemma 4.4, they give by adjunction morphisms

$$\Gamma_{\mathrm{Zar}}(i) \overset{L}{\otimes} \mathbf{Z}/m \to R\alpha_* \mathbf{Z}/m(i) \quad (o \le i \le 2) \tag{26}$$

over the big Zariski site of $\operatorname{Spec} F$.

Let finally $\alpha_X : X_{\mathrm{\acute{e}t}} \to X_{\mathrm{Zar}}$ and $\alpha_Z : Z_{\mathrm{\acute{e}t}} \to Z_{\mathrm{Zar}}$ be the natural morphisms of (small) sites. Note the natural isomorphism of functors

$$Ri^!_{\mathrm{Zar}} R(\alpha_X)_* \xrightarrow{\sim} R(\alpha_Z)_* Ri^!_{\mathrm{\acute{e}t}} \tag{27}$$

over the small Zariski site of Z. (It can be obtained for example with the help of (23); compare [14, II.6.14].)

There is a diagram

$$\begin{array}{ccc}\Gamma_{\mathrm{Zar}}(Z,2-c)[-2c]\overset{L}{\otimes}\mathbf{Z}/m & \longrightarrow & Ri^!_{\mathrm{Zar}}\Gamma_{\mathrm{Zar}}(X,2)\overset{L}{\otimes}\mathbf{Z}/m\\ \downarrow & & \downarrow\\ R(\alpha_Z)_*\mathbf{Z}/m(2-c)[-2c] & \longrightarrow & Ri^!_{\mathrm{Zar}}R(\alpha_X)_*\mathbf{Z}/m(2)\end{array}\tag{28}$$

where the vertical maps are given by (26), the top horizontal map by theorem 4.2 a) and the bottom horizontal map is defined by applying $R(\alpha_Z)_*$ to (21) and using (27). The notation in the top right corner is unambiguous, thanks to lemma 4.4.

4.5. Lemma. *Diagram* (28) *commutes up to sign.*

Proof. As in the proof of lemma 4.3, this boils down to the fact that the Gersten complex for K-theory is compatible with the Gersten complex for étale cohomology via the Galois symbol (m prime to $\operatorname{char} F$) or the differential symbol (m a power of $\operatorname{char} F$). The first case is well-known; see [11, cor. 1.6 and proof of lemma 4.11] for the second one. □

Proof of theorem 4.2 b). We first construct the map. There is a tautological natural transformation (stemming from (27))

$$\alpha_Z^* Ri^!_{\mathrm{Zar}} \to Ri^!_{\text{ét}}\alpha_X^*\tag{29}$$

hence a morphism (in the derived category of étale sheaves over Z)

$$\alpha_Z^*\Gamma_{\mathrm{Zar}}(Z,2-c)[-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X,2)\tag{30}$$

where we used a) and (25). On the other hand, the triangle

$$\begin{array}{ccc}\Gamma(2) & \longrightarrow & \Gamma(2)\otimes\mathbf{Q}\\ & \nwarrow \quad \swarrow & \\ & \mathbf{Q}/\mathbf{Z}(2) & \end{array}\tag{31}$$

deduced from (12) yields a map

$$Ri^!_{\text{ét}}\mathbf{Q}/\mathbf{Z}(2)[-1] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X,2).\tag{32}$$

Passing to the colimit in (21), we get a morphism

$$\mathbf{Q}/\mathbf{Z}(2-c)[-2c] \to Ri^!_{\text{ét}}\mathbf{Q}/\mathbf{Z}(2)\tag{33}$$

whose homotopy cofibre is concentrated in degree $c+3$ and has p-primary torsion cohomology. Shifting and composing with (32), we get a morphism

$$\mathbf{Q}/\mathbf{Z}(2-c)[-1-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X,2).\tag{34}$$

For $c \le 2$, we use (30) to define the map of b), noting that it becomes then

$$\Gamma_{\text{ét}}(Z,2-c)[-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X,2)$$

via (25). For $c > 2$, we use (34) to define this map.

We now prove the property of the map of b) as claimed in the statement of theorem 4.2. It is enough to do this after tensoring (30) and (34) by $\mathbf{Q}$ and $\mathbf{Z}/m$ for all m (in the derived sense). Since $R(\alpha_Z)_*$ is fully faithful, we may even apply this functor to the situation.

Suppose first that $c \leq 2$. Using (27), (14) and a), we see that the morphism

$$R(\alpha_Z)_*\Gamma_{\text{ét}}(Z, 2-c)[-2c] \otimes \mathbf{Q} \to R(\alpha_Z)_* Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X, 2) \otimes \mathbf{Q}$$

is a quasi-isomorphism. On the other hand, there is a $\pm$-commutative diagram

$$\begin{tikzcd}
\Gamma_{\text{ét}}(Z, 2-c) \overset{L}{\otimes} \mathbf{Z}/m[-2c] \arrow[r] \arrow[d, "\wr"] & \alpha_Z^* Ri^!_{\text{Zar}}\Gamma_{\text{Zar}}(X, 2) \overset{L}{\otimes} \mathbf{Z}/m \arrow[r] \arrow[d] & Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X, 2) \overset{L}{\otimes} \mathbf{Z}/m \arrow[dl, "\approx"'] \\
\mathbf{Z}/m(2-c)[-2c] \arrow[r] & Ri^!_{\text{ét}}\mathbf{Z}/m(2). &
\end{tikzcd}$$

In this diagram, the left square is obtained via (25) and (27) by applying adjunction to (28) and using lemma 4.5; the right triangle is obtained via (25) and (29). The left vertical map and the southwest map come from the triangle (12).

The bottom horizontal map is none else than (21): its homotopy cofibre is p-primary torsion and concentrated in degree $c + 3$. The left vertical map and the south-west map are quasi-isomorphisms by (24), hence the top composite has the same cofibre as the bottom map. This proves theorem 4.2 b) in the case $c \leq 2$.

Suppose now that $c > 2$. We first have

$$R(\alpha_Z)_* Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X, 2) \otimes \mathbf{Q} \approx Ri^!_{\text{Zar}} R(\alpha_X)_*\Gamma_{\text{ét}}(X, 2) \otimes \mathbf{Q} \approx Ri^!_{\text{Zar}}\Gamma_{\text{Zar}}(X, 2) \otimes \mathbf{Q} = 0$$

by (14) and a). On the other hand, tensoring (34) by $\mathbf{Z}/m$ and using (31) yields

$$\mathbf{Z}/m(2-c)[-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X, 2) \overset{L}{\otimes} \mathbf{Z}/m.$$

Using (24), we get a composition

$$\mathbf{Z}/m(2-c)[-2c] \to Ri^!_{\text{ét}}\Gamma_{\text{ét}}(X, 2) \overset{L}{\otimes} \mathbf{Z}/m \xrightarrow{\sim} Ri^!_{\text{ét}}\mathbf{Z}/m(2)$$

which is clearly (33). This concludes the proof of theorem 4.2 b). □

5. Cohomology of projective bundles

Let $E \to X$ be a vector bundle of rank n, and $P \xrightarrow{\pi} X$ the associated projective bundle. Our aim in this section is to compute $R\pi_*\Gamma_{\text{Zar}}(P, 2)$ and $R\pi_*\Gamma_{\text{ét}}(P, 2)$.

In order to state the theorem, we remark that there are pairings ($i \leq 2$):

$$\Gamma_{\text{Zar}}(i-1) \overset{L}{\otimes} \Gamma_{\text{Zar}}(1) \to \Gamma_{\text{Zar}}(i) \tag{35}$$

over the big Zariski site of $\operatorname{Spec} F$, if F has more than two elements, and

$$\Gamma_{\text{ét}}(i-1) \overset{L}{\otimes} \Gamma_{\text{ét}}(1) \to \Gamma_{\text{ét}}(i) \tag{36}$$

over the big étale site of $\operatorname{Spec} F$.

For $i = 2$, (35) and (36) are the pairings of [22, prop. 2.5]; for $i = 1$ they are tautological. For $i < 0$ (and in the étale case), the triangle analogous to (31) for $\Gamma(1)$

shows that, for all i, the morphism $\mathbf{Q}/\mathbf{Z}(i-1)\overset{L}{\otimes}\mathbf{Q}/\mathbf{Z}(1)[-1]\to\mathbf{Q}/\mathbf{Z}(i-1)\overset{L}{\otimes}\Gamma_{\text{ét}}(1)$ is a quasi-isomorphism. Therefore it suffices to define morphisms

$$\mathbf{Q}/\mathbf{Z}(i-1)\overset{L}{\otimes}\mathbf{Q}/\mathbf{Z}(1)\to\mathbf{Q}/\mathbf{Z}(i)[1]$$

for all $i\in\mathbf{Z}$. This is nothing else than Tate twists of the natural isomorphisms (in $\mathcal{D}(Ab)$)

$$\mathbf{Q}_l/\mathbf{Z}_l\overset{L}{\otimes}\mathbf{Q}_l/\mathbf{Z}_l\approx\mathbf{Q}_l/\mathbf{Z}_l[1]$$

for $l\neq\operatorname{char}F$. Finally, for $i=0$, the pairing is defined similarly, using the natural map

$$\mathbf{Q}/\mathbf{Z}[-1]\to\mathbf{Z}[0]=\Gamma_{\text{ét}}(0).$$

Let L be a line bundle over an F-scheme S. Let $\lambda=\text{Zar}$ or ét. Via (36), its class $[L]\in H^1_\lambda(S,\mathbb{G}_m)=\mathbb{H}^2_\lambda(S,\Gamma(1))$ defines morphisms of complexes

$$\Gamma_\lambda(i-j)_{|S}[-2j]\xrightarrow{[L]^j}\Gamma_\lambda(i)_{|S}$$

where $_{|S}$ means "restriction to the big λ site of S". In particular, for $S=P$ and $L=\mathcal{O}(1)$, we get maps

$$\Gamma_\lambda(2-j)_{|P}[-2j]\xrightarrow{[\mathcal{O}(1)]^j}\Gamma_\lambda(2)_{|P}\quad(j\geq 0)$$

hence, by adjunction, a morphism

$$\coprod_{j=0}^{n}\Gamma_\lambda(2-j)_{|X}[-2j]\xrightarrow{\rho_\lambda}R\pi_*(\Gamma_\lambda(2)_{|P}).\tag{37}$$

We are now ready to state the result:

5.1. Theorem. *The morphism ρ_λ is a quasi-isomorphism for $\lambda=\text{Zar}$ or ét (for $\lambda=\text{Zar}$, assume F has more than two elements).*

Proof. We proceed as in the last section, first proving the Zariski case. Let A be a local ring of X, and K be its field of fractions. The restriction of E to $\operatorname{Spec}A$ is trivial, hence $P_{|\operatorname{Spec}A}\simeq\mathbf{P}^n_A$. Looking at the maps induced by ρ_{Zar} on cohomology sheaves and using theorem 1.6, we can identify them to:

$$\begin{aligned}K_3(K)_{\text{ind}}&\to K_3(K(T_1,\dots,T_n))_{\text{ind}}\\K_2(A)&\to H^0_{\text{Zar}}(\mathbf{P}^n_A,\mathcal{K}_2)\\A^*&\to H^1_{\text{Zar}}(\mathbf{P}^n_A,\mathcal{K}_2)\\\mathbf{Z}&\to H^2_{\text{Zar}}(\mathbf{P}^n_A,\mathcal{K}_2).\end{aligned}$$

We have to show that all these maps are isomorphisms. The first one is an isomorphism because $K_{3,\text{ind}}$ is invariant under rational extensions. The other ones follow from [9, lemma 8.11].

In the étale case, it is enough to check that ρ is a quasi-isomorphism after tensoring by $\mathbf{Q}$ and by $\mathbf{Z}/l$ for all prime l. In the case of $\mathbf{Q}$, we reduce to the Zariski case as above, by applying $R\alpha_*$ and using (14).

For $\mathbf{Z}/l$, we first need a lemma. Note that there are products:

$$\mathbf{Z}/l(i-1)\overset{L}{\otimes}\mathbf{Z}/l(1)\to\mathbf{Z}/l(i). \tag{38}$$

For $l\neq\operatorname{char}F$, they are nothing else than Tate twists of the natural product in $\mathcal{D}(Ab)$. For $l=\operatorname{char}F$ (and $i>0$), they come from the products

$$\Omega^{i-1}_{\log}\otimes\Omega^1_{\log}\to\Omega^i_{\log}.$$

5.2. Lemma. *For any prime l and any $i\le 2$, the diagram*

$$\begin{array}{ccc}
\Gamma_{\text{ét}}(i-1)\overset{L}{\otimes}\Gamma_{\text{ét}}(1)\overset{L}{\otimes}\mathbf{Z}/l & \longrightarrow & \Gamma_{\text{ét}}(i)\overset{L}{\otimes}\mathbf{Z}/l\\
\downarrow & & \downarrow\\
\mathbf{Z}/l(i-1)\overset{L}{\otimes}\mathbf{Z}/l(1) & \longrightarrow & \mathbf{Z}/l(i)
\end{array}$$

commutes, where the top horizontal map is (36)$\overset{L}{\otimes}\mathbf{Z}/l$*, the bottom horizontal map is* (38) *and the vertical maps are deduced from* (24)*.*

Proof. For $l\neq\operatorname{char}F$, this follows from [22]. For $l=\operatorname{char}F$, it follows from the definition of the logarithmic symbol, since (for $i=1$) the étale sheaf $\mathbf{K}_{3,\text{ind}}$ is uniquely l-divisible. □

If $l\neq\operatorname{char}F$, using lemma 5.2, $\rho\overset{L}{\otimes}\mathbf{Z}/l$ becomes the map γ of [15, th. 2.2.1], which is a quasi-isomorphism, Tate-twisted twice. If $p=\operatorname{char}F$, still using lemma 5.2, $\rho\overset{L}{\otimes}\mathbf{Z}/p$ becomes the map

$$\mathbf{Z}/p[-2]\oplus(\Omega^1_{\log})_{|X}[-1]\oplus(\Omega^2_{\log})_{|X}\to R\pi_*(\Omega^2_{\log})_{|P}$$

shifted, which is an isomorphism by [10, cor. I.2.1.12]. □

5.3. Corollary. $H^0_{\text{Zar}}(X,\mathcal{H}^3(2))\xrightarrow{\sim}H^0_{\text{Zar}}(P,\mathcal{H}^3(2))$.

Proof. (We don't really need $\Gamma(2)$ for this.) Consider the commutative diagram with exact rows

$$\begin{array}{ccccccccc}
0 & \longrightarrow & CH^2X & \longrightarrow & \mathbb{H}^4_{\text{ét}}(X,\Gamma(2)) & \longrightarrow & H^0_{\text{Zar}}(X,\mathcal{H}^3(2)) & \longrightarrow & 0\\
 & & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & CH^2P & \longrightarrow & \mathbb{H}^4_{\text{ét}}(P,\Gamma(2)) & \longrightarrow & H^0_{\text{Zar}}(P,\mathcal{H}^3(2)) & \longrightarrow & 0
\end{array}$$

where the rows come from (9). Using theorem 5.1 and the analogous result for Chow groups, the bottom left horizontal map can be rewritten

$$CH^0(X)\oplus CH^1(X)\oplus CH^2(X)\to\mathbb{H}^0_{\text{ét}}(X,\Gamma(0))\oplus\mathbb{H}^2_{\text{ét}}(X,\Gamma(1))\oplus\mathbb{H}^4_{\text{ét}}(X,\Gamma(2)).$$

The result now comes from the fact that $CH^i(X)\to\mathbb{H}^{2i}_{\text{ét}}(X,\Gamma(i))$ is an isomorphism for $i=0,1$. □

6. The coniveau spectral sequence and Gersten's conjecture

By the standard procedure, we can construct a coniveau spectral sequence ([3], [5])

$$E_1^{p,q} = \coprod_{x\in X^{(p)}} \mathbb{H}_x^{q+p}(X_{\text{ét}},\Gamma(2)) \Longrightarrow \mathbb{H}_{\text{ét}}^{p+q}(X,\Gamma(2))$$

where $\mathbb{H}_x^{q+p}(X_{\text{ét}},\Gamma(2)) = \varinjlim_{U\ni x} \mathbb{H}^{q+p}_{\{x\}\cap U}(U_{\text{ét}},\Gamma(2))$.

Applying theorem 4.2, we get, for $x\in X^{(p)}$:

$$\mathbb{H}_x^{q+p}(X_{\text{ét}},\Gamma(2)) = \begin{cases} \mathbb{H}^q(F(X),\Gamma(2)) & \text{for } p=0\\ H^{q-2}(F(x),\mathbb{G}_m) & \text{for } p=1 \text{ and } q\neq 4,5\\ H^{q-2}(F(x),\mathbf{Z}) & \text{for } p=2 \text{ and } q\neq 4,5\\ H^{q-p-1}(F(x),\mathbf{Q}/\mathbf{Z}(-p)) & \text{for } p>2 \text{ and } q\neq 4,5. \end{cases}$$

Moreover, we have exact sequences:

$$0\to H^{3-p}(F(x),\mathbf{Q}/\mathbf{Z}(2-p))\to \mathbb{H}_x^{p+4}(X_{\text{ét}},\Gamma(2))\to H^0(F(x),\mathcal{F})$$
$$\to H^{4-p}(F(x),\mathbf{Q}/\mathbf{Z}(2-p))\to \mathbb{H}_x^{p+5}(X_{\text{ét}},\Gamma(2))\to 0$$

where $\mathcal{F}$ is an l-primary torsion sheaf if $\operatorname{char} F = l>0$ (and is 0 if $\operatorname{char} F=0$). For $p>2$, the map $H^0(F(x),\mathcal{F})\to H^{4-p}(F(x),\mathbf{Q}/\mathbf{Z}(2-p))$ has to be 0, so the sequence splits into

$$0\to H^{3-p}(F(x),\mathbf{Q}/\mathbf{Z}(2-p))\to \mathbb{H}_x^{p+4}(X_{\text{ét}},\Gamma(2))\to H^0(F(x),\mathcal{F})\to 0$$
$$H^{4-p}(F(x),\mathbf{Q}/\mathbf{Z}(2-p))\xrightarrow{\sim} \mathbb{H}_x^{p+5}(X_{\text{ét}},\Gamma(2)).$$

This shows that $E_1^{p,5}=0$ for $p\geq 5$ and $E_1^{p,4}$ is l-primary torsion for $p\geq 4$. For $q\neq 4,5$, $E_1^{p,q}=0$ for $p\geq q$, except for $E_1^{2,2}=Z^2(X)$ (codimension 2 cycles). Note also that

$$E_1^{p,3}=0 \quad \text{for all } p.$$

Using theorem 5.1 for $P=\mathbf{P}^1_X$, the arguments of [8], [5] show that Gersten's conjecture holds for étale weight-two motivic cohomology. Therefore we get a Bloch-Ogus-type theorem:

6.1. Theorem. *The $E_2^{p,q}$ term of the coniveau spectral sequence for weight-two motivic cohomology coincides with $H^p(X_{\text{Zar}},R^q\alpha_*\Gamma(2)) =: H^p_{\text{Zar}}(X,\mathcal{H}^q(\Gamma(2)))$.* □

6.2. Corollary. *For any $i\geq 0$,*
a) The functor $X\mapsto \mathbb{H}^i_{\text{ét}}(X,\Gamma(2))$ satisfies "codimension 1 purity" for regular local rings of a smooth variety in the sense of [4, def. 2.1.4 (b)].
b) $H^0_{\text{Zar}}(X,\mathcal{H}^i(\Gamma(2)))$ is a birational invariant of smooth, proper varieties X/F.
c) For any proper morphism $P\xrightarrow{f} X$ of smooth, integral F-varieties such that the generic fibre of f is $F(X)$-rational,

$$H^0_{\text{Zar}}(X,\mathcal{H}^i(\Gamma(2)))\xrightarrow{\sim} H^0_{\text{Zar}}(P,\mathcal{H}^i(\Gamma(2))).$$

Proof. a) follows from theorem 6.1. b) follows from theorem 6.1 and [4, prop. 2.1.8]. Finally, c) follows from b) and corollary 5.3. (In [5, §8], we give a general proof of these properties for suitable "cohomology theories with supports".) □

REMARK. As for corollary 5.3, we could prove this without having recourse to $\Gamma(2)$, in view of lemma 1.4. More precisely, we could "merely" use Gersten's conjecture for K-theory (Quillen [31]), étale cohomology with coefficients in twisted roots of unity (Bloch-Ogus [3]) and logarithmic Hodge-Witt cohomology (Gros-Suwa [11]).

7. Projective homogeneous varieties

Let X be a projective homogeneous variety in the sense of [25] and [30]. In particular X is rational, so the assumptions of theorem 1 are satisfied, including $H^0_{\mathrm{Zar}}(\overline{X}, \mathcal{H}^3(2)) = 0$ by corollary 6.2 c). Moreover, we have $K_{j-i}(\overline{F}) \otimes CH^i\overline{X} \xrightarrow{\sim} H^i_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_j)$ for all $i \le j$ (loc. cit.). Finally, the G_F-modules $CH^i\overline{X}$ are permutation modules, hence torsion-free [30]. In particular:

$$H^1(F, H^1_{\mathrm{Zar}}(\overline{X}, \mathcal{K}_2)) = 0$$
$$\operatorname{Ker}\xi = (CH^2 X)_{\mathrm{torsion}}.$$

Let E be the étale F-algebra associated to X as in [25]. We get the following corollary of theorem 1, containing [25, Theorem] and [30, th. 1]:

7.1. COROLLARY. *If X is projective homogeneous, there is an exact sequence:*

$$0 \to H^1_{\mathrm{Zar}}(X, \mathcal{K}_2) \to E^* \xrightarrow{\rho} \operatorname{Ker}\eta \to (CH^2X)_{torsion} \to 0$$

and a complex

$$0 \to \operatorname{Coker}\eta \to \operatorname{Coker}\xi \to \operatorname{Br}(E)$$

which is exact, except perhaps at $\operatorname{Coker}\xi$*, where its homology is* $\operatorname{Ker}(H^4(F,2) \to \operatorname{Coker}\mathrm{cl}^2_X)$.

The map ρ in corollary 7.1 is described by Merkurjev [25]: there is an Azumaya E-algebra A associated to X, and ρ is cup-product by $[A]$ followed by transfer.

7.2. COROLLARY. $\operatorname{Coker}\eta$ *is finite.*

Indeed, $\operatorname{Coker}\xi$ is finite, as a torsion quotient of the finitely generated group $(CH^2\overline{X})^{G_F}$. □

In [19] we show that $\operatorname{Coker}\eta$ is isomorphic to $\operatorname{Ker}\mathrm{cl}^2_X$.

ACKNOWLEDGEMENTS. I wish to thank Jean-Louis Colliot-Thélène, Hélène Esnault, Marc Levine, Eckart Viehweg and the referee for useful comments.

References

[1] J.Kr. Arason *Cohomologische Invarianten quadratischer Formen*, J. Alg. 36 (1975), 448–491.
[2] S. Bloch; K. Kato *p-adic étale cohomology*, Publ. Math. IHES 63 (1986), 107–152.
[3] S. Bloch, A. Ogus *Gersten's conjecture and the homology of schemes*, Ann. Sci. Ec. Norm. Sup. 7 (1974), 181–202.
[4] J.-L. Colliot-Thélène *Birational invariants, purity and Gersten's conjecture*, Proc. Symposia Pure Math. 58.1 (1995), 1–64.
[5] J.-L. Colliot-Thélène, R. Hoobler, B. Kahn *The Bloch-Ogus-Gabber theorem*, preprint, 1996.

[6] J.-L. Colliot-Thélène, W. Raskind *$\mathcal{K}_2$-cohomology and the second Chow group*, Math. Ann. 270 (1985), 165–199.

[7] H. Esnault, B. Kahn, M. Levine, E. Viehweg *The Arason invariant and mod 2 algebraic cycles*, preprint, 1996.

[8] O. Gabber *Gersten's conjecture for some complexes of vanishing cycles*, Manuscripta Math. 85 (1994), 323–343.

[9] H. Gillet *Riemann-Roch theorems for higher algebraic K-theory*, Adv. in Math. 40 (1981), 203–289.

[10] M. Gros Classes de Chern et classes de cycles en cohomologie de Hodge-Witt logarithmique, Mém. Soc. Math. France 21 (1985).

[11] M. Gros, N. Suwa *La conjecture de Gersten pour la cohomologie de Hodge-Witt logarithmique*, Duke Math. J. 57 (1988), 615–628.

[12] A. Grothendieck, J. Dieudonné Eléments de géométrie algébrique, ch. IV: Etude locale des schémas et des morphismes de schémas (première partie), IHES Publ. Math. 20 (1964).

[13] L. Illusie *Complexe de de Rham-Witt et cohomologie cristalline*, Ann. Sci. E.N.S. 12 (1979), 501–661.

[14] B. Iversen Cohomology of sheaves, Springer, 1986.

[15] J.P. Jouanolou *Cohomologie de quelques schémas classiques et théorie cohomologique des classes de Chern*, Exposé 7 in SGA5, Lect. Note in Math. 589, Springer, 1977, 282–350.

[16] B. Kahn *K_3 d'un schéma régulier*, C. R. Acad. Sci. Paris 315 (1992), 433–436.

[17] B. Kahn *Descente galoisienne et K_2 des corps de nombres*, K-theory 7 (1993), 55–100.

[18] B. Kahn, M. Rost, R. Sujatha *Unramified cohomology of quadrics*, preprint, 1996.

[19] B. Kahn *Cohomologie non ramifiée des variétés homogènes*, in preparation.

[20] N. Karpenko *Algebro-geometric invariants of quadratic forms* (in Russian), Algebra i Analiz 2 (1990), 141–162. English translation: Leningrad (Saint-Petersburg) Math. J. 2 (1991), 119–138.

[21] M. Levine *The indecomposable K_3 of fields*, Ann. Sci. Ec. Norm. Sup. 22 (1989), 255–344.

[22] S. Lichtenbaum *The construction of weight-two arithmetic cohomology*, Invent. Math. 88 (1987), 183–215.

[23] S. Lichtenbaum *New results on weight-two motivic cohomology*, The Grothendieck Festschrift, vol. 3, Progress in Math. 88, Birkhäuser, Boston, 1990, 35–55.

[24] A.S. Merkurjev *Certain $\mathcal{K}$-cohomology groups of Severi-Brauer varieties*, K-theory and algebraic geometry: connections with quadratic forms and division algebras, AMS Summer Research Institute, Santa Barbara (1992), Proc. Symp. Pure Math. 58 (2), AMS, Providence 1995, 319–331.

[25] A.S. Merkurjev *The group $H^1(X, \mathcal{K}_2)$ for projective homogeneous varieties*, Algebra i Analiz. English translation: Leningrad (Saint-Petersburg) Math. J. 7 (1995), 136–164.

[26] A.S. Merkurjev, A.A. Suslin *The group K_3 for a field* (in Russian), Izv. Akad. Nauk. SSSR 54 (1990), 339–356. English translation: Math. USSR Izv. 36 (1990), 541–565.

[27] A.S. Merkurjev, J.-P. Tignol *Galois cohomology of biquadratic extensions*, Comm. Math. Helv. 68 (1993), 138–169.

[28] J.S. Milne Etale cohomology, Princeton University Press, 1980.

[29] E. Peyre *Products of Severi-Brauer varieties and Galois cohomology*, K-theory and algebraic geometry: connections with quadratic forms and division algebras, AMS Summer Research Institute, Santa Barbara (1992), Proc. Symp. Pure Math. 58 (2), AMS, Providence 1995, 369–401.

[30] E. Peyre *Corps de fonctions de variétés homogènes et cohomologie galoisienne*, C. R. Acad. Sci. Paris 321 (1995), 891–896.

[31] D. Quillen *Higher algebraic K-theory, I*, Lect. Notes in Math. 341, Springer, 1973, 83–147.

[32] A. Suslin *Torsion in K_2 of fields*, K-theory 1 (1987), 5–29.

Bruno Kahn
Institut de Mathématiques de Jussieu
Equipe Théories Géométriques
Université Paris 7, Case 7012
75251 Paris Cedex 05
France
kahn@math.jussieu.fr

Calabi-Yau Threefolds of Quasi-Product Type

Keiji Oguiso

Received: March 20, 1996

Communicated by Thomas Peternell

Abstract. According to the numerical Iitaka dimension $\nu(X, D)$ and $c_2(X) \cdot D$, fibered Calabi-Yau threefolds $\Phi_{|D|} : X \to W$ $(\dim W > 0)$ are coarsely classified into six different classes. Among these six classes, there are two peculiar classes called of type II_0 and of type III_0 which are characterized respectively by $\nu(X, D) = 2$ and $c_2(X) \cdot D = 0$ and by $\nu(X, D) = 3$ and $c_2(X) \cdot D = 0$. Fibered Calabi-Yau threefolds of type III_0 are intensively studied by Shepherd-Barron, Wilson and the author and now there are a satisfactory structure theorem and the complete classification. The purpose of this paper is to guarantee a complete structure theorem of fibered Calabi-Yau threefolds of type II_0 to finish the classification of these two peculiar classes. In the course of proof, the log minimal model program for threefolds established by Shokurov and Kawamata will play an important role. We shall also introduce a notion of quasi-product threefolds and show their structure theorem. This is a generalization of the notion of hyperelliptic surfaces to threefolds and will have other applicability, too.

1991 Mathematics Subject Classification: Primary: 14J, secondary 14D.

Introduction

Let us start this introduction by recalling a global picture of fibered Calabi-Yau threefolds known at the present and then state the Main Theorem precisely.

Throughout this paper, by a Calabi-Yau threefold, we mean a normal projective complex threefold X with only $\mathbb{Q}-$factorial terminal singularities (so that isolated) and with $\mathcal{O}_X(K_X) \simeq \mathcal{O}_X$ and $\pi_1^{alg}(X) = \{1\}$. The last condition is equivalent to $\pi_1^{alg}(X - \mathrm{Sing}\, X) = \{1\}$, because the local fundamental group of three dimensional terminal Gorenstein singularities is trivial ([Kw3]). This also implies $h^1(\mathcal{O}_X) = 0$ ([O1]). We define

$$c_2(X) \cdot D := c_2(X') \cdot \nu^*(D)$$

for any resolution $\nu : X' \to X$ of $\mathrm{Sing}\,(X)$.

It is known by Miyaoka that $c_2(X) \cdot D$ is non-negative if D is nef ([Mi]).

A surjective morphism $\Phi : X \to W$ is called a fibered Calabi-Yau threefold if X is a Calabi-Yau threefold, W is a normal projective variety (of positive dimension) and Φ has connected fibers. Note that Φ is nothing but $\Phi_{|D|}$ if D is the pull back of (any) very ample divisor H on W.

Fibered Calabi-Yau threefolds $\Phi_{|D|} : X \to W$ are divided into six classes by the numerical invariants $\nu(X, D)$ and $c_2 \cdot D$:

Type I_0 : $\nu(X,D) = 1$ and $c_2 \cdot D = 0$; *Type* I_+ : $\nu(X,D) = 1$ and $c_2 \cdot D > 0$;
Type II_0 : $\nu(X,D) = 2$ and $c_2 \cdot D = 0$; *Type* II_+ : $\nu(X,D) = 2$ and $c_2 \cdot D > 0$;
Type III_0 : $\nu(X,D) = 3$ and $c_2 \cdot D = 0$; *Type* III_+ : $\nu(X,D) = 3$ and $c_2 \cdot D > 0$.

The following (more or less tautological) coarse classification is proved in [O1].

THEOREM 1 ([O1]). *Each class of fibered Calabi-Yau threefolds* $\Phi(= \Phi_{|D|}) : X \to W$ *defined above is characterized as follows.*

Type I_0*: General fibers are smooth Abelian surfaces and* $W = \mathbb{P}^1$*,*
Type I_+*: General fibers are smooth K3 surfaces and* $W = \mathbb{P}^1$*,*
Type II_0*: General fibers are smooth elliptic curves and* W *is a normal projective rational surface with only quotient singularities and with* $K_W \equiv 0$*,*
Type II_+*: General fibers are smooth elliptic curves and* W *is a normal projective rational surface with only quotient singularities and with* $K_W + \Delta \equiv 0$ *for some non-zero effective* $\mathbb{Q}$*-divisor* Δ *such that* (W, Δ) *is klt,*
Type III_0*:* Φ *is a birational morphism and* W *is a normal projective threefold with only canonical singularities and with* $\mathcal{O}_W(K_W) \simeq \mathcal{O}_W$ *and* $c_2(W)(:= \Phi_* c_2(X)) = 0$ *as a linear form on* $Pic(W)$*,*
Type III_+*:* Φ *is a birational morphism and* W *is a normal projective threefold with only canonical singularities and with* $\mathcal{O}_W(K_W) \simeq \mathcal{O}_W$ *and* $c_2(W) \neq 0$*.*

Moreover, if $\Phi : X \to W$ *is a fibered Calabi-Yau threefold of type* II_0 *and* H *is a general very ample divisor on* W*, then the induced elliptic surface* $\Phi^{-1}(H) \to H$ *has no singular fibers while* $\Phi^{-1}(H) \to H$ *has at least one singular fiber composed of rational curves if* $\Phi : X \to W$ *is of type* II_+*.*

Theorem 1 shows that fibered Calabi-Yau threefolds of type III_0 or of type II_0 have rather special nature.

The following two theorems give a complete picture of fibered Calabi-Yau threefolds of type III_0.

THEOREM 2 ([SW]). *Let* $\Phi : X \to \overline{X}$ *be a fibered Calabi-Yau threefold of type* III_0*. Then, there exist an Abelian threefold* A *and a finite Gorenstein automorphism group* G *of* A *such that*

(1) $A^{[G]}$ *is a non-empty finite set, and*
(2) $\overline{X} = A/G$.

Theorem 3 ([O3]). *Two fiber spaces $\Phi_3 : X_3 \to \overline{X_3}$ and $\Phi_7 : X_7 \to \overline{X_7}$ defined in the following (1) and (2) are fibered Calabi-Yau threefolds of type* III_0.

(1) *Let E_{ζ_3} be the elliptic curve with period $\zeta_3 := \exp(2\pi i/3)$. Setting $\overline{X_3} := E^3_{\zeta_3}/\langle \mathrm{diag}(\zeta_3,\zeta_3,\zeta_3)\rangle$, we define $\Phi_3 : X_3 \to \overline{X_3}$ to be a unique crepant (toric) resolution of $\overline{X_3}$.*

(2) *Let A_7 be the Jacobian threefold of the Klein quintic curve $C := (x_0x_1^3 + x_1x_2^3 + x_2x_0^3 = 0) \subset \mathbb{P}^2_{[x_0:x_1:x_2]}$ and g_7 the automorphism of A_7 induced by the automorphism of C given by $[x_0 : x_1 : x_2] \mapsto [x_0^1 : x_1^2 : x_2^4]$. Setting $\overline{X_7} := A_7/\langle g_7\rangle$, we define $\Phi_7 : X_7 \to \overline{X_7}$ to be a unique crepant (toric) resolution of $\overline{X_7}$.*

Conversely, any fibered Calabi-Yau threefold of type III_0 *is isomorphic to either $\Phi_3 : X_3 \to \overline{X_3}$ or $\Phi_7 : X_7 \to \overline{X_7}$ as fiber spaces.*

In particular, there are exactly two fibered Calabi-Yau threefolds of type III_0 *and both of them are smooth and rigid.*

Now it is interesting to study another peculiar class of fibered Calabi-Yau threefolds called of type II_0.

Base surfaces W of fibered Calabi-Yau threefolds $\Phi : X \to W$ of type II_0 are classified into two classes by the global canonical covering $\pi : T \to W$, for which we have either

(1) T is a smooth Abelian surface, or
(2) T is a (projective) K3 surface with only Du Val singularities.

In case (1) (resp. (2)), a fibered Calabi-Yau threefold $\Phi : X \to W$ of type II_0 is called of type II_0A (resp. of type II_0K).

The following theorem gives a complete classification of fibered Calabi-Yau threefolds of type II_0A.

Theorem 4 ([O2]).

(1) *Let $\Phi_3 : X_3 \to E^3_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3,\zeta_3)$ be as in Theorem 3 and $p : X_3 \to E^2_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3)$ the natural map given by the composite of Φ_3 and the natural projection $p_{12} : E^3_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3,\zeta_3) \to E^2_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3)$. Then, any composite of flops $f : X_3 \cdots \to X_3'$ along curves in $p^{-1}(\mathrm{Sing}(E^2_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3)))$ gives a fibered Calabi-Yau threefolds $p\circ f^{-1} : X_3' \to E^2_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3)$ of type II_0A. In this case, $E^2_{\zeta_3}$ is nothing but the global canonical cover of the base surface $E^2_{\zeta_3}/\mathrm{diag}(\zeta_3,\zeta_3)$.*

(2) *Conversely, every fibered Calabi-Yau threefolds of type II_0A is obtained by the above process up to isomorphisms as fiber spaces. In particular, every fibered Calabi-Yau threefolds of type II_0A is smooth and rigid. Moreover, there are exactly 14 different fibered Calabi-Yau threefolds of type II_0A up to isomorphism as fiber spaces.*

The purpose of this paper is to show the following structure theorem of fibered Calabi-Yau threefolds of type II_0K. This theorem tells us how to construct all the fibered Calabi-Yau threefolds of type II_0K.

Main Theorem. *Let us prepare*

(i) *a smooth elliptic curve E with a fixed origin 0,*
(ii) *a projective K3 surface S with only Du Val singularities and its minimal resolution $\mu : S' \to S$, and*
(iii) *two groups*
$G \in \{\{1\}, \mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_4, \mathbb{Z}_5, \mathbb{Z}_6, \mathbb{Z}_7, \mathbb{Z}_8, (\mathbb{Z}_2)^2, (\mathbb{Z}_3)^2, (\mathbb{Z}_4)^2, \mathbb{Z}_2 \times \mathbb{Z}_4, \mathbb{Z}_2 \times \mathbb{Z}_6\}$,
and
$\langle g \rangle \simeq \mathbb{Z}_I \in \{\mathbb{Z}_2, \mathbb{Z}_3, \mathbb{Z}_4, \mathbb{Z}_6\}$,

such that $\tilde{G} := G \rtimes \langle g \rangle$ (semi-direct product) acts faithfully on both E and S (and then on S' and $E \times S'$) in such a way that

(iv) *$G \ni a : E \times S' \to E \times S', (x, y) \mapsto (x + a_E, a_{S'}(y))$ with $a_E \in (E)_{\mathrm{ord}\,(a)}$ and $a_{S'}^* \omega_{S'} = \omega_{S'}$, where $\omega_{S'}$ is a nowhere vanishing regular 2 form on S',*
(v) *$g : E \times S' \to E \times S', (x, y) \mapsto (\zeta_I^{-1} x, g_{S'}(y))$ with $g_{S'}^* \omega_{S'} = \zeta_I \omega_{S'}$, and*
(vi) *$(S')^{[\tilde{G}]} \subset \mathrm{Exc}(\mu)$ except for finitely many points in $(S')^{[\tilde{G}]}$, that is, $(S)^{[\tilde{G}]}$ is a finite set.*

Note that $\tilde{G}$ is a finite Gorenstein automorphism group of $E \times S'$. Let

$$\nu : Y(E, S, \tilde{G}) \to (E \times S')/\tilde{G}$$

be a crepant resolution (whose existence is now guaranteed by Roan [Ro]) and

$$p : Y(E, S, \tilde{G}) \to S/\tilde{G}$$

the natural projection given by the composite of $\nu : Y(E, S, \tilde{G}) \to (E \times S')/\tilde{G}$, $p_2 : (E \times S')/\tilde{G} \to S'/\tilde{G}$, and $\mu/\tilde{G} : S'/\tilde{G} \to S/\tilde{G}$.

Then,

(1) *any composite of flop $f : Y(E, S, \tilde{G}) \cdots \to Y'$ along curves in $p^{-1}(\mathrm{Sing}\,(S/\tilde{G}))$ gives a fibered Calabi-Yau threefold $p \circ f^{-1} : Y' \to S/\tilde{G}$ of type $\mathrm{II}_0 K$ provided that $\pi_1^{alg}(Y) = \{1\}$. In this case S/G gives the global canonical cover of the base space $S/\tilde{G}$.*
(2) *Conversely, every fibered Calabi-Yau threefold of type $\mathrm{II}_0 K$ is obtained by the above process for some triplet $(E, S, \tilde{G})$ satisfying the conditions (i)-(vi) up to isomorphisms as fiber spaces. In particular, every fibered Calabi-Yau threefold of type $\mathrm{II}_0 K$ is smooth.*

This together with Theorems 2, 3 and 4 will complete the structure theorem of the two peculiar classes of fibered Calabi-Yau threefolds called of types II_0 and III_0.

Remark. *Investigating the actions of G and $\langle g \rangle$ on E, we easily see that*

(1) *$\tilde{G}$ is uniquely determined by G and $\langle g \rangle$ as an abstract group, and*
(2) *among 52 possibilities of $(G, \langle g \rangle)$ in the Main Theorem, the following 18 combinations do not occur:*
$(\mathbb{Z}_4, \mathbb{Z}_3)$, $(\mathbb{Z}_5, \mathbb{Z}_3)$, $(\mathbb{Z}_6, \mathbb{Z}_3)$, $(\mathbb{Z}_8, \mathbb{Z}_3)$, $(\mathbb{Z}_2 \times \mathbb{Z}_6, \mathbb{Z}_3)$, $(\mathbb{Z}_2 \times \mathbb{Z}_8, \mathbb{Z}_3)$,
$(\mathbb{Z}_3, \mathbb{Z}_4)$, $(\mathbb{Z}_4, \mathbb{Z}_4)$, $(\mathbb{Z}_6, \mathbb{Z}_4)$, $(\mathbb{Z}_7, \mathbb{Z}_4)$, $(\mathbb{Z}_2 \times \mathbb{Z}_8, \mathbb{Z}_4)$,
$(\mathbb{Z}_2, \mathbb{Z}_6)$, $(\mathbb{Z}_4, \mathbb{Z}_6)$, $(\mathbb{Z}_5, \mathbb{Z}_6)$, $(\mathbb{Z}_6, \mathbb{Z}_6)$, $(\mathbb{Z}_8, \mathbb{Z}_6)$, $(\mathbb{Z}_2 \times \mathbb{Z}_6, \mathbb{Z}_6)$, $(\mathbb{Z}_2 \times \mathbb{Z}_8, \mathbb{Z}_6)$.

Remark. *There are examples of non-rigid fibered Calabi-Yau threefolds of type* II_0K *and the number of fibered Calabi-Yau threefolds of type* II_0K *is not finite any more ([O1]).*

Remark. *It is interesting to compare Theorems 2, 3, 4 and main theorem with the so called Bogomolov decomposition theorem (see for example [Bo]). These look very similar, while our proof is free from the Bogomolov decomposition theorem.*

The Main Theorem and Theorem 4 immediately imply

Corollary. *Let* $\Phi : X \to W$ *is a fibered Calabi-Yau threefold of type* II_0. *Then the global canonical index of* W *is either* 2, 3, 4 *or* 6.

Corollary. *Let* $\Phi : X \to W$ *be a fibered Calabi-Yau threefold of type* II_0K *(resp. of type* II_0A*). Then, there is a composite of flops* $Y \to W$ *of* $\Phi : X \to W$ *over* W *such that* Y *has at least two different fiber space structures,* $Y \to W$ *of type* II_0K *(resp. of type* II_0A*) and* $Y \to \mathbb{P}^1$ *of type* I_+ *(resp. of type* I_0*).*

Very little is known for a fibered Calabi-Yau threefold of type I_0, that is, a Calabi-Yau threefold with an Abelian fibration. However, our main theorem and Theorem 4 show

Corollary. *Let* X *be a Calabi-Yau threefold with at least two different Abelian fibrations. Then,* X *is a Calabi-Yau threefold described as in either the Main Theorem (2) or Theorem 4(2). In particular,* X *is smooth and birational to either a quotient of an Abelian threefolds or that of the product of a K3 surface and an elliptic curve.*

In fact, if $\Phi_{|D_i|} : X \to \mathbb{P}^1$ $(i = 1, 2)$ are two different Abelian fibrations on X, then $\Phi_{|m(D_1+D_2)|} : X \to W$ is of type II_0 for some m.

The outline of this paper is as follows.

In section 1, we introduce the notion of quasi-product threefolds ((1.1)) and show their structure theorem ((1.3)). This plays an important role for our proof of the Main Theorem.

Sections 2 - 4 are devoted to prove the Main Theorem. Since Main Theorem (1) is quite clear, we prove only Main Theorem (2).

Let $\Phi_T : X_T := X \times_W T \to T$ be the base change of a fibered Calabi-Yau threefold $\Phi : X \to W$ of type II_0K to the global canonical cover $\pi : T \to W$. Since Φ always has a two dimensional fibers ([O1]), X_T has very bad singularities and Φ_T itself is a very complicated map in general.

In section 2, we apply the log minimal model program established by Shokurov and Kawamata or Kollár et al. [Sh] and [Kw4] (also [Ko3]) to find a good birational (canonical) model $f : Z \to T$ of $\Phi_T : X_T \to T$ over T such that

(1) $Gal(T/W) := \langle g \rangle$ acts regularly on $f : Z \to T$ and
(2) $\Phi : X \to W$ is birational to the quotient $(f : Z \to T)/\langle g \rangle$.

Moreover applying the result in section 1, we show that there are a smooth elliptic curve E, a normal projective surface S which is either an Abelian surface or a K3 surface with only Du Val singularities, and a finite automorphism group G of the fiber space $p_2 : E \times S \to S$ such that $(f : Z \to T) = (p_2 : E \times S \to S)/G$.

In section 3, we show that the action of $\langle g\rangle$ on $f : Z \to T$ lifts to that on its covering $p_2 : E \times S \to S$ in an equivariant way. This is a rather special phenomenon, because a composite of Galois extensions is not Galois in general.

Till section 3, the main part of our proof of the Main Theorem is completed. It remains only to show the impossibility for S to be a smooth Abelian surface. This problem is treated in section 4. This requires our assumption $\pi_1^{alg}(X) = \{1\}$ and forces rather minute analysis of automorphism groups of an Abelian surface.

Acknowledgement

This work was completed during the author's stay at Johns Hopkins University in October 1995. The author would like to express his best thanks to Professors Y. Kawamata and V. Shokurov for their invitation. The author would like to express his thanks to JSPS for financial support during his stay.

Notation and Convention

Throughout this paper, we work over the complex number field $\mathbb{C}$.

We will employ standard notion and notation in minimal model program ([KMM] or [Ko3]) freely.

By a minimal threefold, we mean a normal projective threefold V with only $\mathbb{Q}-$factorial terminal singularities and with nef canonical (Weil) divisor K_V.

A surjective morphism $\Phi : V \to W$ is said to be relatively minimal if V has only $\mathbb{Q}-$factorial terminal singularities and the canonical divisor K_V is relatively nef with respect to Φ.

We often use the notion of klt (Kawamata log terminal) given in [Ko3]. This is same as the notion of log terminal in [KMM].

By a fiber space on a normal projective variety V, we mean a surjective morphism $\Phi : V \to W$ to a normal projective variety W with connected fibers. Note that Φ is not equi-dimensional in general. By $\Phi^{-1}(w)$ $(w \in W)$, we denote the scheme theoretic fiber over w. We denote its reduction by $\Phi^{-1}(w)_{\mathrm{red}}$. This is in some sense a set theoretical fiber.

Two fiber spaces $\Phi : V \to W$ and $\Phi' : V' \to W'$ are said to be isomorphic if there are isomorphisms $F : V \to V'$ and $f : W \to W'$ such that $\Phi' \circ F = f \circ \Phi$.

For two morphisms $\Phi : V \to W$ and $\pi : T \to W$, we sometimes denote natural morphisms $V \times_W T \to T$ and $V \times_W T \to V$ by $\Phi_T : V_T \to T$ and $\pi_V : V_T(= T_V) \to V$ respectively.

The primitive $n-$th root of unity $exp(2\pi i/n)$ is denoted by ζ_n.

We denote the cyclic group of order n by $\mathbb{Z}_n$.

The elliptic curve with period $\tau \in \mathbb{H}$ is written as E_τ.

The $n-$torsion group of an Abelian variety A with origin 0 is denoted by $(A)_n$. By global coordinates around a point P of an $n-$dimensional Abelian variety A, we mean those of its universal cover $\mathbb{C}^n$ or, equivalently, those of the tangent space $T_{A,P}$.

For a faithful group action of G on a variety V, we set

$$V^{[G]} := \{x \in V \mid \exists g \in G - \{1\}, g(x) = x\},$$

while,

$$H^G := \{v \in H \mid \forall g \in G, g^*(v) = v\}$$

for any cohomology group H of V.

Similarly, for an automorphism g of a variety V, we set

$$V^g := \{x \in V \mid g(x) = x\}.$$

An equivariant action of a finite group G on a fibration $\Phi : V \to W$ induces a new fibration $\Phi(\operatorname{mod} G) : X/G \to W/G$. We sometimes abbreviate this fibration by $(\Phi : V \to W)/G$.

We say that G acts on $\Phi : V \to W$ over W if the action of G is equivariant and is trivial on W.

An automorphism group G of a variety V with $\mathcal{O}_V(K_V) \simeq \mathcal{O}_V$ is called Gorenstein if the action of G on $H^0(V, \mathcal{O}_V(K_V))$ is trivial, that is, all elements g of G satisfy $g^*\omega_V = \omega_V$ for a generator ω_V of $H^0(V, \mathcal{O}_V(K_V))$.

For the automorphism group $\operatorname{Aut}(V)$ of a variety V and a subset B in V, we often consider the subgroup $\{g \in \operatorname{Aut}(V) \mid g(B) = B\}$. We denote this group by $\operatorname{Aut}(X, B)$. For example, if A is an Abelian variety with origin 0, then $\operatorname{Aut}(A, \{0\})$ is nothing but the so called Lie automorphism group of A.

§1. Quasi-product threefolds

In this preliminary section, we shall introduce the notion of quasi-product threefolds and prove their structure theorem (Theorem (1.3)). This is a rather wide generalisation of the notion of hyperelliptic surfaces to threefolds.

Definition (1.1). *A normal projective threefold V with only rational singularities is called a quasi-product threefold with distinguished morphisms a and f if*

(1) *V has a fiber space structure $a : V \to A$ over a smooth elliptic curve A,*
(2) *V has a fiber space structure $f : V \to T$ over a normal projective surface T with only rational singularities and with $H^1(\mathcal{O}_T) = 0$ such that $f^{-1}(t)_{\mathrm{red}}$ is a smooth elliptic curve for any $t \in T$, and that $f^{-1}(t)$ itself is smooth except at most finitely many points $t \in T$.*

Example (1.2). *Let S be a normal projective surface with only rational singularities and E a smooth elliptic curve. Assume that a finite group of translations G of E acts faithfully on S in such a way that $S^{[G]}$ is finite and $H^1(\mathcal{O}_S)^G = 0$. Then the quotient threefold $(E \times S)/G$ is a quasi-product threefold with distinguished morphisms $p_1 : (E \times S)/G \to E/G$ and $p_2 : (E \times S)/G \to S/G$.*

Conversely, we shall show

Theorem (1.3). *Let V be a quasi-product threefold with two distinguished morphisms $a : V \to A$ and $f : V \to T$. Let S be a general fiber of a.*

Then, there exist an elliptic curve E and a finite subgroup $G \subset E$, that is, a finite group of translations of E (and then is isomorphic to either $\mathbb{Z}_m$ or $\mathbb{Z}_n \times \mathbb{Z}_m$ with $(n|m)$) such that

(1) *there is an injective homomorphism $\iota : G \to \operatorname{Aut}(S)$,*
(2) *$V = (E \times S)/G$ under the (free) action of G on $E \times S$ defined by*

$$G \ni g : E \times S \ni (u, v) \mapsto (u + g, \iota(g)v) \in E \times S,$$

(3) *two distinguished morphisms* $a : V \to A$ *and* $f : V \to T$ *are given by the natural projections*

$$p_1 : (E \times S)/G \to E/G$$

and

$$p_2 : (E \times S)/G \to S/\iota(G)$$

respectively.

As a result, S *can be replaced by any fiber of* a*. We set* $G_S := \iota(G)(\simeq G)$*. Moreover, if* $\mathcal{O}_V(K_V) \simeq \mathcal{O}_V$*, then,*

(4) *any fiber* S *of* a *is either a K3 surface with only Du Val singularities or a smooth Abelian surface,*
(5) G_S *is a finite Gorenstein automorphism of* S*,*
(6) *if* S *is a K3 surface with only Du Val singularities, then* $S^{[G_S]}$ *is a non-empty finite set and* $G_S(\simeq G)$ *is isomorphic to either one of the following groups;* $\{1\}$, $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_4$, $\mathbb{Z}_5$, $\mathbb{Z}_6$, $\mathbb{Z}_7$, $\mathbb{Z}_8$, $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_4$, $\mathbb{Z}_2 \times \mathbb{Z}_6$, $\mathbb{Z}_3 \times \mathbb{Z}_3$, *or* $\mathbb{Z}_4 \times \mathbb{Z}_4$,
(7) *if* S *is a smooth Abelian surface, then* $S^{[G_S]}$ *is a non-empty finite set and* $G_S(\simeq G)$ *is isomorphic to either one of the following groups;* $\{1\}$, $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_4$, $\mathbb{Z}_6$, *or* $\mathbb{Z}_2 \times \mathbb{Z}_2$, $\mathbb{Z}_2 \times \mathbb{Z}_4$, $\mathbb{Z}_3 \times \mathbb{Z}_3$.
In addition, if $G_S \simeq \mathbb{Z}_m$*, then* $G_S \subset \mathrm{Aut}\,(S, \{0\})$ *for an appropriate origin* 0 *of* S*, while, if* $G_S \simeq \mathbb{Z}_n \times \mathbb{Z}_m$ $(n|m)$*, then* $\mathbb{Z}_n \subset (S)_n$ *and* $\mathbb{Z}_m \subset \mathrm{Aut}\,(S, \{0\})$ *for an appropriate origin* 0 *of* S*. Moreover,* $\mathrm{Sing}\,(S/G_S)$ *is described as follows for each* G_S *([Kt]).*
$(G_S, \mathrm{Sing}\,(S/G_S)) = (\mathbb{Z}_2, 16A_1), (\mathbb{Z}_2 \times \mathbb{Z}_2, 16A_1), (\mathbb{Z}_3, 9A_2), (\mathbb{Z}_3 \times \mathbb{Z}_3, 9A_2)$ $(\mathbb{Z}_4, 4A_3 + 6A_1), (\mathbb{Z}_2 \times \mathbb{Z}_4, 4A_3 + 6A_1), (\mathbb{Z}_6, A_5 + 4A_2 + 5A_1)$.

Remark. *Let* $\nu : S' \to S$ *be the minimal resolution of* S*. Then* G *induces an equivariant free action on* $id \times \nu : E \times S' \to E \times S$*. The induced morphism* $(E \times S')/G \to (E \times S)/G$ *gives a resolution of* $(E \times S)/G$*.*

Remark. *Our proof given here basically follows the argument of Bombieri and Mumford for hyperelliptic surfaces([BM]). However, since we work at threefolds, we should keep the following two essential differences in mind:*

(1) f *may not be flat over* T*,*
(2) *three dimensional relatively minimal models are not unique among their birational models (even if they exist) so that rational actions on a relatively minimal model are not necessarily regular in general.*

Proof. Set $B := \{t \in T |$ either $f^{-1}(t)$ is not reduced or T is singular at $t\}$, and denote $C_t := f^{-1}(t)(t \in T)$ and $S_x := a^{-1}(x)$ $(x \in A)$. By our assumption, B is a finite set. Let us fix a general point $0 \in A$ and regard this point as an origin of A. Set $S := S_0$. Then S is a normal surface with only rational singularities. Put $n := (C_t \cdot S)$. This is independent of $t \in T - B$ (because $T - B$ is smooth and $f|_{f^{-1}(T-B)}$ is a smooth morphism over $T - B$.)

Claim (1.4). *$a_t := a|_{C_t} : C_t \to A$ is surjective for each $t \in T - B$. In particular, a_t is an isogeny of elliptic curves of degree $n := (C_t \cdot S)$ for each $t \in T - B$ (and then $n > 0$).*

Proof of Claim (1.4). Assume the contrary that $a(C_t)$ is a point on A for some $t \in T - B$. Then, $a(C_{t'})$ must be a point for every $t' \in T - B$ because f is flat over $T - B$. Thus, a induces a morphism $\overline{a} : T - B \to A$. This gives a rational map $\overline{a} : T \cdot\cdot\to A$ with $a = \overline{a} \circ f$. Let $T' \to T$ be a resolution of both singularities of T and indeterminacy of $\overline{a}$. Since T has only rational singularities, we have $h^1(\mathcal{O}_{T'}) = h^1(\mathcal{O}_T) = 0$. Thus, $\overline{a} \circ \nu(T')$ is a point. Hence $\overline{a}$ is a morphism and $\overline{a}(T)$ is a point. Then, $a(V)$ would be a point because $a = \overline{a} \circ f$. But this contradicts the surjectivity of a. q.e.d. for (1.4).

Let t be an arbitrary point on $T - B$. Then, by (1.4), A acts on C_t via the composite of the group homomorphism $A \simeq Pic^0(A) \to Pic^0(C_t)$ given by a_t^* and the natural action of $Pic^0(C_t)$ on C_t. More concretely, this action is written as

$$A \ni x : C_t \ni P \mapsto P + x_1 + ... + x_n - 0_1 - ... - 0_n \in C_t,$$

where $\{x_1, ..., x_n\} := a_t^{-1}(x) = C_t \cap S_x$ and $\{0_1, ..., 0_n\} := a_t^{-1}(0) = C_t \cap S$. Note that f has a local section over $T - B$. Thus, gluing these together, we get a regular action of A on $\cup_{t \in T-B} C_t = f^{-1}(T - B)$ over $T - B$. This gives a rational action on V over T. But, since the possible indeterminacy $f^{-1}(B)$ of this action on V consists of elliptic curves (then no rational curves) and since V has only rational singularities, this action of A on V must be regular. Let us denote this action by $\sigma : A \times V \to V$. By construction, σ stabilizes each fiber of f. Set $\tau := \sigma|_{A \times S} : A \times S \to V$. Since a_t is an isogeny, we have

$$a_t(P + x_1 + ... + x_n - 0_1 - ... - 0_n) = a_t(P) + nx$$

for $t \in T - B$ and $x \in A$. So, once we define a new action of A on A by

$$A \ni x : A \to A; y \mapsto y + nx,$$

that is, by $n \times$ (*translation*), then A induces an equivariant action on the fibration $V - f^{-1}(B) \to A$. By the same reason as before, this action of A is extended to an equivariant regular action on the whole space $a : V \to A$.

By definition, we have $x(S)(= x(S_0)) = S_{nx}$ $(x \in A)$. In particular, $\tau : A \times S \to V$ is surjective. Moreover, the action of the $n-$torsion group $(A)_n$ of A on V stabilizes $S = S_0$. This induces a group homomorphism $\iota : (A)_n \to \mathrm{Aut}\,(S)$.

The following claim ([BM]) is now proved formally.

Claim (1.5). *Let (x, v) and (x', v') be points on $A \times S$. Then, the following (1) and (2) are equivalent to one another.*

(1) $\tau(x, v) = \tau(x', v')$,

(2) *(x, v) and (x', v') are in the same orbit of the action*

$$(A)_n \ni k : A \times S \to A \times S; (x, v) \mapsto (x - k, \iota(k)v).$$

Proof of Claim (1.5). Since $\tau(x-k,\iota(k)v)=\sigma(x-k,\sigma(k,v))=\sigma(x-k+k,v)=\tau(x,v)$, (2) implies (1). We prove the converse. Since $\tau(x,v)\in S_{nx}$ and $\tau(x',v')\in S_{nx'}$, it follows that $nx=nx'$, or equivalently, $k:=x-x'\in(A)_n$. We may show that $\iota(k)(v)=v'$. Using $\tau(x,v)=\tau(x',v')$, that is, $\sigma(x,v)=\sigma(x',v')$, we calculate

$$v'=\sigma(-x',\sigma(x',v'))=\sigma(-x',\sigma(x,v))=\sigma(x-x',v).$$

This is nothing but the desired equality, $\iota(k)(v)=v'$. q.e.d. for (1.5).

By (1.5), we get $V=(A\times S)/(A)_n$. Moreover, just by construction, we see that $f:(A\times S)/(A)_n\to T$ factors through the natural projection $p_2:(A\times S)/(A)_n\to S/(A)_n$. In fact, f factors through p_2 at least over $T-B$. But, since B is finite and $S/(A)_n$ is normal, this is so over the whole T. Let $\mu:S/(A)_n\to T$ be the induced morphism. Since both f and p_2 have only one dimensional connected fibers, μ must be a finite birational morphism. Thus, by the Zariski main theorem, μ is isomorphism and then $f=p_2$ under the identification $T=S/(A)_n$. Similarly, $a:(A\times S)/(A)_n\to A$ factors through $p_1:(A\times S)/(A)_n\to A/(A)_n=A$. Now the equality $a=p_2$ is shown by the same argument as before.

It only remains to make ι injective to complete the first half part of (1.3). But this is done as follows. Let $G=(A)_n/Ker\,\iota$. Then, $(A\times S)/(A)_n=(A/(Ker\,\iota)\times S)/G$ and $A/(A)_n=(A/Ker\,\iota)/G$, in which G acts on translation group of an elliptic curve $A/Ker\,\iota$. Now replacing A, $(A)_n$ and ι by $E=A/(Ker\,\iota)$, G, and the injection $\iota\circ(-1):G\to\mathrm{Aut}\,(S)$, we are done. Here we will compose (-1) only to change the sign $-$ in (1.5) into $+$ as in (1.3).

From now on, we shall prove the latter half part of (1.3). It is obvious that S is either a K3 surface with only Du Val singularities or a smooth Abelian surface. Moreover, since G acts on E as a translation group and $\mathcal{O}_V(K_V)\simeq\mathcal{O}_V$, it follows that G_S must be a Gorenstein automorphism group of S. In the rest we denote G_S simply by G if no confusion seems to arise.

Assume first that S is a K3 surface with only Du Val singularities. Let $S'\to S$ be the minimal resolution of S. Then G gives a commutative Gorenstein action on S'. Now the result follows from the Nikulin's classification ([Ni]). Note that two groups $(\mathbb{Z}_2)^3$ and $(\mathbb{Z}_2)^4$ in his list are excluded because G is isomorphic to either $\mathbb{Z}_n$ or $\mathbb{Z}_n\times\mathbb{Z}_m$ $(n|m)$.

Finally, assuming that S is a smooth Abelian surface, we show that G satisfies the condition in (1.3)(7). Since G is a finite Gorenstein automorphism group of S with $T=S/G$ and since $h^1(T,\mathcal{O}_T)=0$, it follows that $S^{[G]}$ is a non-empty finite set. Choose an appropriate origin 0 of S and identify S with its translation automorphism group. Set $Aut^0(S):=\{\sigma\in\mathrm{Aut}\,(S)|\sigma^*\omega_S=\omega_S\}$, $Aut^0(S,\{0\}):=\{\sigma\in Aut^0(S)|\sigma(0)=0\}$, where ω_S is a non-zero global regular two form on S. Then, $Aut^0(S)=S\rtimes Aut^0(S,\{0\})$ and $G\subset Aut^0(S)$. Identifying $Aut^0(S,\{0\})=Aut^0(S)/S$, we denote the natural projection by $p:Aut^0(S)\to Aut^0(S,\{0\})$. If we choose global coordinates around 0, we can explicitly write down the action of $g\in Aut^0(S)$ in its affine form

$$g(x)=M_gx+t_g, M_g\in SL(2,\mathbb{C}), t_g\in S.$$

Then p is nothing but the map taking the matrix part, that is, $g \mapsto M_g$. It follows from this expression that

(1) as an abstract group, $p(G)$ is independent of the choice of an origin of S,
(2) a finite Gorenstein automorphism $g \in Aut^0(S)$ has a fixed point if and only if g is not a translation.

On the other hand, Katsura's classification ([Kt]) of possible finite subgroups of $Aut^0(S, \{0\})$ shows that the commutative group $p(G)$ is isomorphic to either $\mathbb{Z}_2$, $\mathbb{Z}_3$, $\mathbb{Z}_4$ or $\mathbb{Z}_6$.

Thus we can choose $g \in G$ and $0 \in S$ such that $p(g)$ generates $p(G)$ and $g(0) = 0$. From now on, we regard this point 0 as the origin of S.

Claim (1.6).

(1) *$H := Ker(p)$ consists of translations in G, that is, $H \subset S$,*
(2) *$\langle g \rangle \simeq p(G)$.*
(3) *G is isomorphic to $H \times \langle g \rangle$.*
(4) *H is a subgroup of S^g (under the inclusion $H \subset S$).*

Proof of (1.6). The assertion (1) follows from $M_h = id$ for $h \in H$. By definition, $p|_{\langle g \rangle} : \langle g \rangle \to p(G)$ is surjective group homomorphism. Let h be an element of $Ker(p|_{\langle g \rangle})$. Then, $h(0) = 0$ and $h \in H$. Combining this with (1), we get $h = id$. Thus, $p|_{\langle g \rangle}$ is isomorphism. This shows that G is a semi-direct product of H and $\langle g \rangle$. Since G is commutative, this must be the direct product. The last statement now directly follows from the relation $gh = hg$ $(h \in H)$. q.e.d. of (1.6).

Claim (1.7). *According to* $\operatorname{ord}(g) = 2, 3, 4, 6$, *$S^g$ is isomorphic to $(\mathbb{Z}_2)^4$, $(\mathbb{Z}_3)^2$, $(\mathbb{Z}_2)^2$ and $\{0\}$.*

Proof of (1.7). If $\operatorname{ord}(g) = 2$, then $S^g = (S)_2$. Since $(S)_2 \simeq (\mathbb{Z}_2)^4$, we are done.

Assume that $\operatorname{ord}(g) = 3$. Then, using appropriate global coordinates (x, y) around 0, we can write $g = \operatorname{diag}(\zeta_3, \zeta_3^{-1})$. In particular, $1 + g + g^2 = 0$. Thus, $3p = p + p + p = p + g(p) + g^2(p) = (1 + g + g^2)(p) = 0$ for $p \in (S)^g$. Hence $S^g \subset (S)_3$ and $S^g \simeq (\mathbb{Z}_3)^k$ for some non negative integer k. On the other hand, by the Lefschetz fixed point formula, we have $\sharp S^g = \sum_{i=0}^{4} (-1)^i tr(g^* | H^i(S, \mathbb{C}))$. Recall that

$$H^1(S, \mathbb{C}) = \mathbb{C}dx \oplus \mathbb{C}dy \oplus \mathbb{C}d\overline{x} \oplus \mathbb{C}d\overline{y},$$

and

$$H^i(S, \mathbb{C}) = \wedge^i H^1(S, \mathbb{C}).$$

Now an explicit calculation based on $g = \operatorname{diag}(\zeta_3, \zeta_3^{-1})$ shows $tr(g^* | H^0(S, \mathbb{C})) = 1, -2, 3, -2, 1$ according to $i = 0, 1, 2, 3, 4$. Thus, $\sharp S^g = 9$. This implies $S^g \simeq (\mathbb{Z}_3)^2$.

Assume that $ord(g) = 4$. Since $S^g \subset S^{g^2} \simeq (\mathbb{Z}_2)^4$, it follows that $S^g \simeq (\mathbb{Z}_2)^k$ for some non negative integer k. As in the case of $\operatorname{ord}(g) = 3$, we can choose appropriate global coordinates (x, y) around 0 such that $g = \operatorname{diag}(\zeta_4, \zeta_4^{-1})$. Then, again using the Lefschetz fixed point formula, we calculate $\sharp S^g = 4$. This implies $S^g \simeq (\mathbb{Z}_2)^2$.

Finally assume that $\operatorname{ord}(g) = 6$. Then, it follows from the previous observation that $S^g \subset S^{g^2} \cap S^{g^3} \subset (S)_2 \cap (S)_3 = \{0\}$. q.e.d. of (1.7).

Now Claims (1.6), (1.7) and the fact that G is a finite Abelian group of the form $\mathbb{Z}_n$ or $\mathbb{Z}_n \times \mathbb{Z}_m$ $(n|m)$ together with the fundamental theorem on finite Abelian groups imply the assertion (1.3)(7).

The only remaining problem is to study $\mathrm{Sing}\,(S/G)$ for each G. If G is isomorphic to $\mathbb{Z}_m$, the result follows from Katsura's table ([Kt]). Next, consider the case when $\mathbb{Z}_n \times \mathbb{Z}_m$ for some n and m (with $n|m$). Since $S/G \simeq (S/\mathbb{Z}_n)/\mathbb{Z}_m$ and since $(S/\mathbb{Z}_n)$ is again an Abelian surface, the assertion follows from the first case.

Now we are done. Q.E.D. of (1.3).

§2. Good model over the global canonical covering

Let us fix a fibered Calabi-Yau threefold $\Phi : X \to W$ of type $\mathrm{II}_0 K$. Define $I := min\{n \in \mathbb{N} | \mathcal{O}_W(nK_W) \simeq \mathcal{O}_W\}$ and denote the global canonical cover of W by $\pi : T \to W$ ([Kw1, Z]). By our assumption, T is a projective K3 surface with only Du Val singularities. Set $W_0 := W - \mathrm{Sing}\,(W)$. It is well known by [Kw1, Z] that $\pi : T \to W$ is a cyclic Galois covering of order $I(W)$ and is étale over W_0. Moreover, there is a generator g of the Galois group $Gal(T/W)$ such that $g^*\omega_T = \zeta_I \omega_T$, where ω_T is a nowhere vanishing regular two form on T, that is, a generator of $H^0(\mathcal{O}_T(K_T))$.

We fix these notation till the end of Section 4.

Set $\Phi_T : X_T := X \times_W T \to T$. Then, the Galois group $Gal(T/W) = \langle g \rangle$ acts on this fibration by $g : (x, y) \mapsto (x, g(y))$ and induces an isomorphism

$$(\Phi : X \to W) \simeq (\Phi_T / : X_T \to T)/\langle g \rangle.$$

However, X_T itself has very bad singularities in general.

The goal of this section is to prove the following

Key Lemma (2.1). *There is a normal projective threefold Z such that*

(1) *Z has only $\mathbb{Q}-$factorial canonical singularities with $\mathcal{O}_Z(K_Z) \simeq O_Z$,*
(2) *Z is a quasi-product threefold ((1.1)) with two distinguished morphisms $f : Z \to T$ and $a : Z \to A$, where the latter map is the Albanese morphism of Z (see [Kw2] for the definition of the Albanese variety and the Albanese morphism for varieties with rational singularities), and*
(3) *there is a regular action of the Galois group of $\langle g \rangle$ on the fibration $f : Z \to T$ such that $W = T/\langle g \rangle$ and $(\Phi : X \to W)$ is birational to $(f : Z \to T)/\langle g \rangle$ over $W = T/\langle g \rangle$. Moreover, these are isomorphic over $W - \mathrm{Sing}\,(W)$.*

The plan of proof of Key Lemma is as follows. First, applying the log minimal model program, we find a birational model $f : Z \to T$ of $\Phi_T : X_T \to T$ with property (1) in (2.1). Then, we check that $f : Z \to T$ also satisfies (2) and (3).

In order to carry out this plan, we start by observing some general lemmas.

Proposition (2.2). *Let $\varphi : V \to S$ be a surjective morphism from a normal projective $\mathbb{Q}-$factorial threefold V to a normal projective surface S. Let $\{E_i\}_{i \in I}$ be the set of all two-dimensional irreducible components in fibers of φ. Set $E = \Sigma_{i \in I} E_i$. Assume that*

(1) *V is not covered by rational curves,*

(2) $K_V = \Sigma_{i\in I} a_i E_i$ *(as a Weil divisor on V) for some* $a_i \in \mathbb{Z}_{\geq 0}$,
(3) $(V, \epsilon E)$ *is klt for some positive small rational number* ϵ.

Then, there are a normal projective threefold $V^{(n)}$ *and a surjective morphism* $\varphi^{(n)} : V^{(n)} \to S$ *such that*

(4) $V^{(n)}$ *has only* $\mathbb{Q}-$*factorial canonical singularities with* $\mathcal{O}_{V^{(n)}}(K_{V^{(n)}}) \simeq \mathcal{O}_{V^{(n)}}$,
(5) $\varphi^{(n)} : V^{(n)} \to S$ *is birational to* $\varphi : V \to S$ *over* S *and is isomorphic except over a finite set* $\varphi(E)$, *and*
(6) $\varphi^{(n)} : V^{(n)} \to S$ *is an equi-dimensional elliptic fibration.*

Proof. First, we remark

CLAIM (2.3). $K_V + \epsilon E$ *is not nef unless* $E = 0$ *as a divisor.*

Proof of (2.3). Let H be a general very ample divisor on V. Then H is a normal surface and the restriction $\varphi|_H : H \to S$ is surjective. Since $(K_V + \epsilon E)|_H \equiv \Sigma_{i\in I}(a_i + \epsilon)E_i|_H$ and since $E_i|_H$ are contracted by φ_H, we get

$$((K_V + \epsilon E)^2 \cdot H) = ((K_V + \epsilon E)|_H)^2 = (\Sigma_{i\in I}(a_i + \epsilon)E_i|_H)^2 < 0$$

unless $E = 0$. q.e.d. of (2.3).

Let us apply the log minimal model program for a klt divisor $K_V + \epsilon E$. If $E \neq 0$, then $K_V + \epsilon E$ is not nef by (2.3). Thus, there is a log extremal ray R such that $(K_V + \epsilon E) \cdot C < 0$ for any curve C belonging to R. Let $cont_R : V \to W$ be the contraction morphism associated to R. This is a birational morphism by our assumption (1). Since $0 > (K_V + \epsilon E) \cdot C = \Sigma(a_i + \epsilon)(E_i \cdot C)$, there is a prime divisor E_i such that $E_i \cdot C < 0$. This implies $C \subset E_i$. Thus $cont_R$ is defined over S. Let $\phi : W \to S$ be the induced morphism.

If $cont_R$ is a divisorial contraction, setting $V^{(1)} := W$, $\varphi^{(1)} := \phi$ and changing E by its strict transform $E^{(1)}$ on $V^{(1)}$, we see that $\varphi^{(1)} : V^{(1)} \to S$ and $E^{(1)}$ satisfy all the assumptions in (2.2) (without any change of coefficients).

If $cont_R$ is a small contraction, then we apply a log flip for $cont_R$ to get $cont_R^+ : V^+ \to W$.

The existence of log flips for threefolds is guaranteed by [Sh].

Now, setting $V^{(1)} := V^+$, $\varphi^{(1)} := \phi \circ cont_R^+$ and changing E by its strict transform $E^{(1)}$ on $V^{(1)}$, we see that $\varphi^{(1)} : V^{(1)} \to S$ and $E^{(1)}$ also satisfy all the assumptions in (2.2).

Putting $V^{(0)} := V$, $\varphi^{(0)} := \varphi$ and $E^{(0)} := E$ and repeating this process, say, for $n(\geq 0)$ times, we finally get $\varphi^{(n)} : V^{(n)} \to S$ and the strict transform $E^{(n)}$ of E to $V^{(n)}$ such that

(1) $\varphi^{(n)} : V^{(n)} \to S$ and $E^{(n)}$ satisfy all the assumptions in (2.2), and
(2) $K_{V^{(n)}} + \epsilon E^{(n)}$ is nef.

This is due to the termination of log flips for threefolds shown by [Kw4].

Then $E^{(n)} = 0$ by (2.3). This implies the equi-dimensionality of $\varphi^{(n)}$. Note that all modifications are done over $\varphi(E)$. Thus $\varphi^{(n)} : V^{(n)} \to S$ and $\varphi : V \to S$ coincide over $S - \varphi(E)$. Set $V_0 := V - E - \mathrm{Sing}\,(V)$. Then the assumption (2) implies $\mathcal{O}_{V_0}(K_{V_0}) \simeq \mathcal{O}_{V_0}$. Let $\nu : V \cdot\cdot \to V^{(n)}$ be the birational map obtained by the above process. Since $\nu|_{V_0} : V_0 \to \nu(V_0)$ is an isomorphism, we have $\mathcal{O}_{\nu(V_0)}(K_{\nu(V_0)}) \simeq \mathcal{O}_{\nu(V_0)}$.

Since the codimension of $V^{(n)} - \nu(V_0)$ in $V^{(n)}$ is at least two by $E^{(n)} = 0$ and since $V^{(n)}$ is normal, this isomorphism gives $\mathcal{O}_{V^{(n)}}(K_{V^{(n)}}) \simeq \mathcal{O}_{V^{(n)}}$. Note that $V^{(n)}$ has only rational singularities, because $(V^{(n)}, E^{(n)}) = (V^{(n)}, 0)$ is klt. Thus $V^{(n)}$ has only rational Gorenstein singularities, that is, canonical singularities of index one. Now the remaining assertion is obvious. Q.E.D. of(2.2).

The next two lemmas are concerned with singular fibers of certain elliptic threefolds.

Lemma (2.4). *Let $\varphi : V \to S$ be a fiber space such that*

(1) *V is a normal projective threefold with only $\mathbb{Q}-$factorial terminal singularities and with $K_V \equiv 0$,*
(2) *S is a normal projective surface with only quotient singularities and with $K_V \equiv 0$.*

Then, $\varphi^{-1}(s)$ is a smooth elliptic curve if $s \in S - \mathrm{Sing}\,(S)$. In particular, φ is a smooth morphism over $S - \mathrm{Sing}\,(S)$.

Proof. We make use of the following theorem due to Nakayama.

Theorem (2.5)([Na1 also Na2]). *Let $f : V_{\Delta^2} \to \Delta^2$ be a relatively minimal projective elliptic fibration over a two-dimensional (small) polydisk*

$$\Delta^2 := \{(x, y) \in \mathbb{C}^2 \mid |x| < \epsilon, |y| < \epsilon\}.$$

Assume that f has (singular) fibers of type I_a ($a \geq 0$) over $(x = 0) - \{(0, 0)\}$ and those of type I_b ($b \geq 0$) over $(y = 0) - \{(0, 0)\}$. (Here we employed Kodaira's notation.) Then $f^{-1}((0, 0))$ is a (singular) fiber of type I_{a+b}. In particular, if f is smooth over $\Delta^2 - \{(0, 0)\}$, then $f^{-1}((0, 0))$ is a smooth elliptic curve and f is a smooth morphism over the whole Δ^2.

First, we show

Claim (2.6). *$\varphi : V \to S$ is an elliptic fibration and has singular fibers only over a finite set of points of S.*

Proof of (2.6). Note that a general fiber of φ is a smooth elliptic curve. Let H be a general very ample divisor on S. Set $V_H := \varphi^{-1}(H)$. Since V has only isolated singularities and since H is general, we may assume that $H \cap (\mathrm{Sing}\,(S) \cup \varphi(\mathrm{Sing}\,(V))) = \phi$ and both H and V_H are smooth. Let $\varphi|_{V_H} : V_H \to H$ be the induced elliptic fibration. Using the adjunction formula, we calculate $K_H \equiv H|_H$ and $K_{V_H} = (K_V + V_H)|_{V_H} \equiv \varphi^*(K_H)$. Comparing this with the canonical bundle formula of an elliptic surface (for example see [BPV]), we find that $\varphi|_{V_H}$ is a smooth morphism. This implies the result. q.e.d of (2.6).

Let $s \in S$ be an arbitrary smooth point of S and take a sufficiently small polydisk $\Delta^2 \subset S$ around s. By (2.6), φ is smooth over $\Delta^2 - \{s\}$. Now applying (2.5) for an elliptic fibration $\varphi|_{\varphi^{-1}(\Delta^2)} : \varphi^{-1}(\Delta^2) \to \Delta^2$, we get (2.4). Q.E.D. of (2.4).

Lemma (2.7). *Let $\varphi : V \to S$ be a fiber space such that*

(1) *V is a normal projective threefold with only canonical singularities and with $\mathcal{O}_V(K_V) \simeq \mathcal{O}_V$,*
(2) *S is a normal projective surface with only Du Val singularities and with $\mathcal{O}_S(K_S) \simeq \mathcal{O}_S$,*
(3) *φ is an equi-dimensional fibration, and*
(4) *φ is smooth except over a finite set of points of S.*

Then, the reduction of each fiber $\varphi^{-1}(s)_{\mathrm{red}}$ ($s \in S$) is a smooth elliptic curve. Moreover, if s is a smooth point of S, then, $\varphi^{-1}(s)$ itself is a smooth elliptic curve. In particular, φ is a smooth morphism over $S - \mathrm{Sing}\,(S)$.

Proof. Let $s \in S$ be an arbitrary point of S. Since S has only Du Val singularities, we can choose a small neighborhood U around s such that

$$U = \Delta^2/G, s = (0,0)(\mathrm{mod}\, G).$$

Here Δ^2 is a two dimensional small polydisk and G is a finite Gorenstein automorphism group of Δ^2 each of whose element fixes only the origin $(0,0)$. We may also assume by (4) that φ is smooth over $U - \{s\}$.

Letting $\varphi_U : V_U \to U$ be the restriction of φ, we consider the fiber product

$$\varphi_{\Delta^2} : V_{\Delta^2} := V_U \times_U \Delta^2 \to \Delta^2.$$

Since $\Delta^2 \to U$ is étale over $U - \{s\}$ and φ_U is smooth over $U - \{s\}$, it follows that $\varphi_{\Delta^2} : V_{\Delta^2} \to \Delta^2$ is smooth over $\Delta^2 - \{(0,0)\}$.

Take a resolution $\nu : V^{(1)} \to V_{\Delta^2}$ of V_{Δ^2} and set $\varphi^{(1)} := \varphi \circ \nu : V^{(1)} \to \Delta^2$. Note that φ and $\varphi^{(1)}$ coincide over $\Delta^2 - \{(0,0)\}$.

Applying a relatively minimal model program with respect to $K_{V^{(1)}}$ over Δ^2 ([Mo]), we get a relatively minimal model

$$\varphi^{(2)} : V^{(2)} \to \Delta^2$$

of $\varphi^{(1)} : V^{(1)} \to \Delta^2$. Since each fiber of $\varphi^{(1)}$ over $\Delta^2 - \{(0,0)\}$ is a smooth elliptic curve, $\varphi^{(2)}$ coincides with $\varphi^{(1)}$ (and then φ_{Δ^2}) over $\Delta^2 - \{(0,0)\}$. This together with (2.5) implies that $(\varphi^{(2)})^{-1}((0,0))$ is also a smooth elliptic curve and that $\varphi^{(2)}$ is smooth over whole Δ^2. In particular, $V^{(2)}$ is also smooth. Since φ_{Δ^2} and $\varphi^{(2)}$ are birational over Δ^2, the natural action of G on $\varphi_{\Delta^2} : V_{\Delta^2} \to \Delta^2$ induces a rational action on

$$\varphi^{(2)} : V^{(2)} \to \Delta^2.$$

On the other hand, since each fiber of $\varphi^{(2)}$ is an elliptic curve, it follows that $\varphi^{(2)}$ is a unique relatively minimal model. Thus this action of G on $\varphi^{(2)} : V^{(2)} \to \Delta^2$ is regular and induces

$$\overline{\varphi^{(2)}} : V^{(2)}/G \to \Delta^2/G = U.$$

This is birational to $\varphi_U : V_U \to U$ over U and is isomorphic over $U - \{s\}$. Denote this birational map over U by

$$\mu : V_U \cdots \to V^{(2)}/G.$$

Then, μ gives an isomorphism

$$V_U - \varphi_U^{-1}(s) \simeq V^{(2)}/G - (\overline{\varphi^{(2)}})^{-1}(s).$$

Since $\mathcal{O}_{V_U - \varphi_U^{-1}(s)}(K_{V_U}) \simeq \mathcal{O}_{V_U - \varphi_U^{-1}(s)}$ by our assumption (1) and since $(\overline{\varphi^{(2)}})^{-1}(s)$ is of codimension two in a normal variety it follows that

$$\mathcal{O}_{V^{(2)}/G}(K_{V^{(2)}/G}) \simeq \mathcal{O}_{V^{(2)}/G}.$$

This shows that the action of G on $V^{(2)}$ is Gorenstein. Since each element of G fixes the origin $(0,0)$ of Δ^2, G stabilizes a smooth elliptic curve $E := (\varphi^{(2)})^{-1}((0,0))$. Since G is also Gorenstein on Δ^2, so is on E. That is, G acts on E as a translation group. Thus $(\overline{\varphi^{(2)}})^{-1}(s)_{\mathrm{red}} = E/G$ is a smooth elliptic curve.

Now, in order to complete the first part of (2.7), it is enough to show that $\mu : V_U \cdots \to V^{(2)}/G$ is actually an isomorphism. But, now, this immediately follows from the facts that V_U has only rational singularities and that $V^{(2)}/G$ is $\mathbb{Q}-$factorial.

If s is a smooth point of S, then we can take $G = \{1\}$ and then $V_U = V^{(2)}$ over $U = \Delta^2$. This implies the last half of (2.7). Q.E.D. of (2.7).

The next lemma is a slight generalization of Kollár's result (in the three dimensional case), which should be known by specialists. However, because of the lack of suitable references, we give here a brief proof based on the Kollár's original result.

LEMMA (2.8). *Let $\varphi : V \to S$ be a fiber space such that*

(1) *V is a normal projective threefold with only canonical singularities,*
(2) *S is a normal surface with only Du Val singularities.*

Let ω_V and ω_S be the dualizing sheaves on V and S. Then, $R^1\varphi_*\omega_V \simeq \omega_S$. Assume furthermore that

(3) *$\mathcal{O}_V(K_V) \simeq \mathcal{O}_V$ and*
(4) *S is a K3 surface with only Du Val singularities.*

Then $h^1(\mathcal{O}_V) = 1$.

REMARK. Kollár proved the first part of (2.8) under the assumption that both V and S are smooth ([Ko1]).

Proof. We want to reduce our proof to the smooth case.

Consider the following commutative diagram,

$$\begin{array}{ccc} V' & \xrightarrow{\nu} & V \\ \Phi\big\downarrow & & \big\downarrow\varphi \\ S' & \xrightarrow[\mu]{} & S \end{array}$$

where $\mu : S' \to S$ is the minimal resolution of S' and $\nu : V' \to V$ is a resolution of both the singularities of V and indeterminacy of $\mu^{-1} \circ \varphi$.

Then $R^i\nu_*\omega_{V'} = 0$ for $i > 0$. Moreover, $\nu_*\omega_{V'} = \omega_V$ because V has only canonical singularities. Thus, from the Leray spectral sequence

$$R^p\varphi_*(R^q\nu_*\omega_{V'}) \Rightarrow R^{p+q}(\varphi \circ \nu)_*\omega_{V'}$$

we get

$$R^p\varphi_*\omega_V \simeq R^p(\varphi \circ \nu)_*\omega_{V'} \simeq R^p(\mu \circ \Phi)_*\omega_{V'}.$$

In particular,

$$R^1\varphi_*\omega_V \simeq R^1(\mu \circ \Phi)_*\omega_{V'}.$$

On the other hand, the edge sequence of another Leray spectral sequence

$$R^p\mu_*(R^q\Phi_*\omega_{V'}) \Rightarrow R^{p+q}(\mu \circ \Phi)_*\omega_{V'}$$

gives an exact sequence

$$0 \to R^1\mu_*(\Phi_*\omega_{V'}) \to R^1(\mu \circ \Phi)_*\omega_{V'} \to \mu_*(R^1\Phi_*\omega_{V'}) \to R^2\mu_*(\Phi_*\omega_{V'}).$$

Note that $R^2\mu_*(\Phi_*\omega_V) = 0$ and that $R^1\mu_*(\Phi_*\omega_{V'})$ is a torsion sheaf, because $\mu : S' \to S$ is a birational morphism between surfaces.

On the other hand, since V' is smooth, $R^1(\mu \circ \Phi)_*\omega_{V'}$ is a torsion free sheaf by [Ko1]. Then, chasing the above exact sequence, we get

$$R^1\mu_*(\Phi_*\omega_{V'}) = 0$$

and

$$R^1(\mu \circ \Phi)_*\omega_{V'} \simeq \mu_*(R^1\Phi_*\omega_{V'}).$$

Since V' and S' are smooth, Kollár's original result implies

$$R^1\Phi_*\omega_{V'} \simeq \omega_{S'}.$$

Thus,

$$R^1(\mu \circ \Phi)_*\omega_{V'} \simeq \mu_*\omega_{S'}.$$

Moreover, since S has only canonical singularities, it follows that

$$\mu_*\omega_{S'} \simeq \omega_S.$$

Thus,

$$R^1(\mu \circ \Phi)_*\omega_{V'} \simeq \omega_S.$$

Combining these, we get

$$R^1(\mu \circ \Phi)_*\omega_{V'} \simeq \omega_S.$$

This completes the proof of the first part.

We show the second part. Since $\omega_V \simeq \mathcal{O}_V$ and $\omega_S \simeq \mathcal{O}_S$, the first part of (2.8) gives

$$R^1\varphi_*\mathcal{O}_Z \simeq \mathcal{O}_S.$$

Substituting this into the edge sequence of the Leray spectral sequence

$$H^p(R^q\varphi_*\mathcal{O}_V) \Rightarrow H^{p+q}(\mathcal{O}_V),$$

we get an exact sequence

$$0 \to H^1(\mathcal{O}_S) \to H^1(\mathcal{O}_V) \to H^0(\mathcal{O}_S).$$

This implies

$$h^1(\mathcal{O}_V) \le h^1(\mathcal{O}_S) + h^0(\mathcal{O}_S) = 0 + 1 = 1.$$

We show that $h^1(\mathcal{O}_V) \ge 1$. Considering the pullback of the regular two forms by Φ and using Hodge theory, we calculate

$$h^2(\mathcal{O}_{V'}) = h^{2,0}(V') \ge h^{2,0}(S') = 1.$$

On the other hand, using the fact that V has only rational singularities and the Serre duality, we see that

$$h^2(\mathcal{O}_{V'}) = h^2(\mathcal{O}_V) = h^1(\mathcal{O}_V).$$

Combining these, we get the desired inequality $h^1(\mathcal{O}_V) \ge 1$. Q.E.D. of (2.8).

We return back to Key Lemma (2.1). This is now proved by a simple combination of the previous lemmas.

Proof of Key Lemma.

Set $W_0 := W - \mathrm{Sing}\,(W)$ as before and denote the restrictions of $\Phi : X \to W$ and $\pi : T \to W$ to W_0 by

$$\Phi_0 : X_0 := \Phi^{-1}(W_0) \to W_0$$

and

$$\pi_0 : T_0 := \pi^{-1}(W_0) \to W_0.$$

Note that Φ_0 is a smooth morphism by (2.4) and π_0 is an étale morphism by definition.

We consider the Cartesian product defined by Φ and π

$$\begin{array}{ccc} X_T := X \times_W T & \xrightarrow{\pi_X} & X \\ {\scriptstyle \Phi_T}\downarrow & & \downarrow{\scriptstyle \Phi} \\ T & \xrightarrow[\pi]{} & W \end{array}$$

and its restriction over W_0

$$\begin{array}{ccc} (X_T)_0 := X_0 \times_{W_0} T_0 & \longrightarrow & X_0 \\ \downarrow & & \downarrow \\ T_0 & \longrightarrow & W_0 \end{array}$$

Since W_0 is smooth and since each morphism in the second diagram is smooth or étale, it follows that

$$\mathrm{Sing}\,(X) \subset \Phi^{-1}(W - W_0),$$

and

$$\mathrm{Sing}\,(X_T) \subset \pi_X^{-1}(\mathrm{Sing}\,(X)) \subset (\pi_X \circ \Phi)^{-1}(W - W_0) = \Phi_T^{-1}(T - T_0).$$

In what follows, we apply several birational modifications on the first diagram keeping everything in the second diagram invariant.

Since all singularities in the first diagram are supported over $W - W_0$, we find a commutative diagram

$$\begin{array}{ccc} X'_T & \xrightarrow{\pi'_X} & X' \\ {\scriptstyle \nu_{X_T}}\downarrow & & \downarrow{\scriptstyle \nu_X} \\ X_T & \xrightarrow[\pi_X]{} & X \end{array}$$

such that

(1) X' and X'_T are smooth,
(2) $\nu_X : X' \to X$ is a birational modification only over $W - W_0$, and that
(3) $\nu_{X_T} : X'_T \to X_T$ is a birational modification only over $T - T_0$.

Let $\{E_i\}_{i \in I}$ be the set of all the two dimensional irreducible components of fibers of $\Phi'_T := \Phi_T \circ \nu_{X_T} : X'_T \to X_T \to T$. Set $E := \Sigma_{i \in I} E_i$. By construction, E is supported only over $T - T_0$.

Claim (2.10).

(1) *X'_T is not covered by rational curves.*
(2) *$K_{X'_T} = \Sigma_{i \in I} a_i E_i$ for some non-negative integers a_i.*
(3) *$(X'_T, \epsilon E)$ is klt if $\epsilon > 0$ is sufficiently small.*

Proof of (2.10). The assertions (1) and (3) are clear. We show the assertion (2). Since X has only terminal singularities, $\mathrm{Sing}\,(X) \subset X - X_0$, and $K_X = 0$ as a divisor, we see that

$$K_{X'} = \Sigma c_j E'_j,$$

where c_j are some positive integers and E'_j are some irreducible divisors supported in $\nu_X^{-1}(X - X_0)$.

On the other hand, since $\pi'_{X_T} : X'_T \to X'$ ramifies only at E, the ramification formula gives

$$K_{X'_T} = (\pi'_{X_T})^*(K_{X'}) + \Sigma_{i \in I} b_i E_i,$$

for some non-negative integers b_i. Since $(\pi'_{X_T})^* E'_i$ are effective divisors supported in E, substituting the first equality into the second, we get the result. q.e.d. of (2.10).

Now we can apply (2.2) for $\Phi'_T : X'_T \to T$ to get a fiber space $f : Z \to T$ such that

(1) Z has only $\mathbb{Q}-$factorial canonical singularities with $\mathcal{O}_Z(K_Z) \simeq \mathcal{O}_Z$,
(2) $f : Z \to T$ is birational to $\Phi_T : X_T \to T$ over T and is isomorphic over T_0,
(3) $f : Z \to T$ is an equi-dimensional elliptic fibration.

Recall that T is a K3 surface with only Du Val singularities, and that Φ_T is smooth over T_0.
Now using (2.7) and (2.8), we see that

(4) $f^{-1}(t)_{\mathrm{red}}$ is a smooth elliptic curve for each $t \in T$,

(5) $f^{-1}(t)$ itself is smooth if t is a smooth point of T (in particular, if $t \in T_0$),
(6) $h^1(\mathcal{O}_Z) = 1$.

Thus, it follows from (1) and (6) and [Kw2] that

(7)$A := \mathrm{Alb}\,(Z)$ is a smooth elliptic curve and the Albanese morphism $a : Z \to A$ is a fiber space.

By (2), the natural action of $\langle g \rangle$ on $\Phi_T : X_T \to T$ induces a rational action of G on $f : Z \to T$ which is regular over T_0. By virtue of (1) and (4), we can apply the same argument as in the last part of the proof of (2.7) to conclude

(8) $\langle g \rangle$ induces a regular action on $f : Z \to T$ and
(9) $(f : Z \to T)/\langle g \rangle$ is birational to $\Phi : X \to W$ and is isomorphic over $W_0 = T_0/\langle g \rangle$.

Now these statements (1) - (9) imply the Key Lemma. Q.E.D. of Key Lemma.

§3. Lifting the group action on a fiber space to its covering

In this section, we continue to employ the same notation given at the beginning of Section 2.

Let $f : Z \to T$ be the quasi-product threefold found in (2.1) for a fibered Calabi-Yau threefold $\Phi : X \to W$ of type $\mathrm{II}_0 K$.

Then $(f : Z \to T) \simeq (p_2 : E \times S \to S)/G$, where

(1) E is a smooth elliptic curve,
(2) S is either a (projective) K3 surface with only Du Val singularities or a smooth Abelian surface, given as (any) fiber of the Albanese morphism $a : Z \to A$,
(3) G is a finite commutative Gorenstein automorphism group of $E \times S$ as is described in Theorem (1.3).

We want to lift the action of $\langle g \rangle$ on $f : Z \to T$ to one on $p_2 : E \times S \to S$ in an equivariant way.

Lemma (3.1). *There is a point* 0 *on* A *such that* $\langle g \rangle$ *stabilizes* $a^{-1}(0)$.

Proof. Since the Albanese morphism is an intrinsically and uniquely defined object, $\langle g \rangle$ acts on the Albanese morphism $a : Z \to A$. This induces a fibration

$$\overline{a} : Z/\langle g \rangle \to A/\langle g \rangle.$$

On the other hand, since X and $Z/\langle g \rangle$ are birational and since both of them have only rational singularities, it follows that $h^1(\mathcal{O}_{Z/\langle g \rangle}) = h^1(\mathcal{O}_X) = 0$. This implies $A/\langle g \rangle = \mathbb{P}^1$. Thus, $A^{\langle g \rangle} \neq \phi$. Since A is an elliptic curve, this is equivalent to $A^g \neq \phi$. Hence we can choose such a point 0 in A^g. Q.E.D. of (3.1).

Let us take $a^{-1}(0)$ as S. Then g induces an action $g_S := g|_S : S \to S$. Since g acts on the fiber space $f : Z \to T$, $\langle g_S \rangle$ and $\langle g \rangle$ give an equivariant action on $q_T := f|_S : S \to T$. Note that q_T is nothing but the quotient map $S \to T = S/G$.

Lemma (3.2). *$g_S^*\omega_S = \zeta_I \omega_S$, where ω_S is a nowhere vanishing regular two form on S, that is, a generator of $H^0(S, \mathcal{O}_S(K_S))$.*

Proof. Let ω_T be a nowhere vanishing regular two form on T. Then, $\omega_S := q_T^*\omega_T$ is a nowhere vanishing regular two form on S. Thus,

$$g_S^*\omega_S = g_S^* \circ q_T^*\omega_T = q_T^* \circ g^*\omega_T = q_T^*\zeta_I\omega_T = \zeta_I\omega_S.$$

This implies the result. Q.E.D. of (3.2).

Lemma (3.3). *There is an automorphism $g_{E\times S}$ of $E \times S$ such that $g_{E\times S}$, g_S and g give an equivariant action on the commutative diagram*

$$\begin{tikzcd} E \times S \arrow[r, "p_2"] \arrow[d, "q'"'] & S \arrow[d, "q_T"] \\ Z \arrow[r, "f"'] & T \end{tikzcd}$$

where q and q' are natural quotient maps.

Proof. Let us consider the fiber product

$$\begin{tikzcd} Z \times_T S \arrow[r, "p_2"] \arrow[d, "p_1"'] & S \arrow[d, "q_T"] \\ Z \arrow[r, "f"'] & T \end{tikzcd}$$

Define the action of $\langle g' \rangle$ on $Z \times_T S$ by

$$g' : Z \times_T S \ni (u, v) \mapsto (g(u), g_S(v)) \in Z \times_T S.$$

Then, g', $\langle g_S \rangle$ and $\langle g \rangle$ give an equivariant action on this fiber product.

By the definition of fiber product, there is a surjective morphism $\nu : E \times S \to Z \times_T S$ which factors through the quotient map $q : E \times S \to Z = (E \times S)/G$ and the second projection $p_2 : E \times S \to S$.

Claim (3.4). *$\nu : E \times S \to Z \times_T S$ is the normalization of $Z \times_T S$.*

Proof of (3.4). Obvious. q.e.d. of (3.4).

Since normalization is an intrinsically and uniquely defined notion, the action $\langle g' \rangle$ on $Z \times_T S$ lifts to the action $\langle g_{E\times S} \rangle$ on $E \times S$ equivariantly with respect to $\nu : E \times S \to Z \times_T S$. This gives a desired action on $E \times S$. Q.E.D. of (3.3).

Corollary (3.5). $\mathrm{ord}\,(g_S) = \mathrm{ord}\,(g_{E\times S}) = I (:= \mathrm{ord}\,(g))$.

Proof. Since g_S is a restriction of g, it follows that $\mathrm{ord}\,(g_S) \le \mathrm{ord}\,(g)$. On the other hand, since $\tau : S \to T$ is surjective and since g_S and g induce an equivariant action on τ, we see that $\mathrm{ord}\,(g_S) \ge \mathrm{ord}\,(g)$. This implies $\mathrm{ord}\,(g_S) = \mathrm{ord}\,(g)$. Now it follows from the construction of $g_{E\times S}$ that $\mathrm{ord}\,(g_{E\times S}) = \mathrm{ord}\,(g') = \mathrm{ord}\,(g)$. Q.E.D. of (3.5).

Define $\tilde{G}$ to be the subgroup of $\mathrm{Aut}\,(E \times S)$ generated by G and $g_{E\times S}$ found in (3.3). Then $\tilde{G}$ acts on the fiber space $p_2 : E \times S \to S$. Thus, there is a (unique) group homomorphism $\rho : \tilde{G} \to \mathrm{Aut}\,(S)$ such that $p_2 \circ h = \rho(h) \circ p_2$. By construction, we have $\rho(G) = G_S$ and $\rho(g_{E\times S}) = g_S$. Corollary (3.5) shows that $\rho|_{\langle g_{E\times S}\rangle} : \langle g_{E\times S}\rangle \to \langle g_S\rangle$ is a group isomorphism as is $\rho|_G : G \to G_S$. Set $\tilde{G_S} = \rho(\tilde{G})$.

Lemma (3.6).

(1) *G_S is a normal subgroup of $\tilde{G_S}$.*
(2) *$\tilde{G_S} = G_S \rtimes \langle g_S\rangle$.*
(3) *G is a normal subgroup of $\tilde{G}$.*
(4) *$\tilde{G} = G \rtimes \langle g_{E\times S}\rangle$.*
(5) *$\rho : \tilde{G} \to \tilde{G_S}$ is an isomorphism.*

Proof. For the assertion (1), it is enough to show that there is an $h' \in G_S$ such that $g_S \circ h = h' \circ g_S$ for each $h \in G_S$. Let $s \in S$ be a point on S such that $g_S(s) \notin S^{G_S}$. Using $g \circ q_T = q_T \circ g_S$ and $T = S/G_S$, we calculate

$$q_T \circ g_S \circ h(s) = g \circ q_T \circ h(s) = g \circ q_T(s) = q_T \circ g_S(s).$$

Thus, for each $s \in S$, there is $h_s \in G_S$ such that $g_S \circ h(s) = h_s \circ g_S(s)$. Such an h_s is uniquely determined by s because $g_S(s) \notin S^{G_S}$. Thus, we find a continuous map $S - R \to G_S$ defined by $s \mapsto h_s$. Since G_S is discrete, the image must be one point, say h'. Then, $g_S \circ h = h' \circ g_S$ over $S - g_S^{-1}(S^{G_S})$. Taking the closure, we find that $g_S \circ h = h' \circ g_S$ whole over S. This finishes the proof of (1).

Applying the same argument for $E \times S \to (E \times S)/G = Z$ (instead of $T = S/G_S$), we can also show assertion (3).

We show assertion (2). By (1), we have $\tilde{G_S}/G_S = \langle g_S (\mathrm{mod}\, G_S)\rangle$. Consider the natural representation $\tilde{G_S}$ on $H^0(S, \mathcal{O}_S(K_S))$

$$\zeta : \tilde{G} \to \mathbb{C}^\times, h \mapsto \zeta(h)$$

defined by $h^*\omega_S = \zeta(h)\omega_S$. Since G_S is a Gorenstein automorphism group of S, this factors

$$\overline{\zeta} : \tilde{G_S}/G_S = \langle g_S(\mathrm{mod}\, G_S)\rangle \to \mathbb{C}^\times.$$

Since $\overline{\zeta}(g_S(\mathrm{mod}\, G_S)) = \zeta(g_S) = \zeta_I$ by (3.3), it follows that $\mathrm{ord}\,(g_S(\mathrm{mod}\, G_S)) \geq I = \mathrm{ord}\,(g_S)$. Thus, the natural surjective group homomorphism $\langle g_S\rangle \to \langle g_S(\mathrm{mod}\, G_S)\rangle$ must be isomorphism. This implies the assertion (2).

Finally, we show assertions (4) and (5).

By (3), we see that $\tilde{G}/G \simeq \langle g_{E\times S}(\mathrm{mod}\, G)\rangle$. Combining this with (3.5), we get

$$\sharp\tilde{G} = (\sharp G) \cdot (\sharp\langle g_{E\times S}(\mathrm{mod}\, G)\rangle) \leq (\sharp G) \cdot (\sharp\langle g\rangle).$$

On the other hand, by (2) and (3.5), we have

$$\sharp\tilde{G_S} = (\sharp G_S) \cdot (\sharp\langle g_S\rangle) = (\sharp G) \cdot (\sharp\langle g\rangle).$$

However, since $\tilde{G_S}$ is an image of $\tilde{G}$, it follows that

$$\sharp\tilde{G} \geq \sharp\tilde{G_S}.$$

Combining these three we get $\sharp\tilde{G} = \sharp\tilde{G_S}$. This implies that the surjective group homomorphism $\rho : \tilde{G} \to \tilde{G_S}$ is an isomorphism. Combining this together with (2), we get $\tilde{G} = G \rtimes \langle g_{E\times S}\rangle$. This completes the proof. Q.E.D. of (3.6).

From now on, we denote the equivariant actions $\tilde{G}$ and $\tilde{G_S}$ on the fiber space $p_2 : E \times S \to S$ simply by $\tilde{G}$. We also set $\tilde{g} := g_{E\times S}$ for consistency of notation. If no confusion seems to arise, we also identify g_S and G_S with $\tilde{g}$ and G (under the isomorphism ρ).

The following corollary is an immediate consequence of Lemma (3.6).

Corollary (3.7).

$$(f : Z \to T)/\langle g\rangle = (p_2 : E \times S \to S)/\tilde{G}.$$

Thus, the fiber space $\Phi : X \to W$ is birational to $(p_2 : E \times S \to S)/\tilde{G}$ over $W = S/\tilde{G}$ and is isomorphic over W_0.

Now this together with the next lemma and the corollary completes the proof of Main Theorem (2) modulo impossibility for S to be a smooth Abelian surface.

Lemma (3.8). *Assume that S is a K3 surface with only Du Val singularities. Then, the action of $\tilde{g}$ on $E \times S$ is written as follows:*

$$\tilde{g} : E \times S \ni (x, y) \mapsto (\zeta_I^{-1}x, g_S(y)) \in E \times S$$

for an appropriate origin 0 of E.

Proof. Since $\langle\tilde{g}\rangle$ acts on $p_2 : E \times S \to S$, there is a homomorphic map

$$c : S \to \mathrm{Aut}\,(E) = E \rtimes \mathrm{Aut}\,(E, \{0\})$$

defined by $s \mapsto (p_1((x, s)) \mapsto p_1(\tilde{g}(x, s)))$.

On the other hand, since $h^1(\mathcal{O}_S) = 0$ and S has only Du Val singularities, the Albanese variety of S is trivial. Thus c must be constant map. That is, $\tilde{g} = (g_E, g_S)$ for some $g_E \in \mathrm{Aut}\,(E)$. Since X is isomorphic to $(E \times S)/\tilde{G}$ over W_0 and since $(E\times S)/\tilde{G} \to W$ is equidimensional, $\mathcal{O}_X(K_X) \simeq \mathcal{O}_X$ implies $\mathcal{O}_{(E\times S)/\tilde{G}}(K_{(E\times S)/\tilde{G}}) \simeq \mathcal{O}_{(E\times S)/\tilde{G}}$. This means $\tilde{G}$ is a Gorenstein automorphism of $E \times S$. In particular, so is $\tilde{g}$. Combining this with $g_S^*\omega_S = \zeta_I\omega_S$, we get $g_E^*\omega_E = \zeta_I^{-1}\omega_E$. In particular, $E^{g_E} \neq \phi$. Now, choosing the origin 0 of E in E^{g_E}, we get the desired expressions of $\tilde{g}$. This completes the proof of (3.8). Q.E.D.

Combining (3.8) and (3.7), we get

Corollary(3.9). *Assume that S is a K3 surface with only Du Val singularities. Then,*

(1) *the global canonical index $I = I(W)$ of W is either $2, 3, 4$, or 6,*
(2) *if $\nu : S' \to S$ is a minimal resolution of S, then the action $\langle\tilde{g}\rangle$ on $E \times S$ lifts to $E \times S'$ in an equivariant way and $\Phi : X \to W$ is birational to*

$$(p_2 \circ (id. \times \nu) : E \times S' \to S)/\tilde{G}$$

over $W = S/\tilde{G}$ and is isomorphic over W_0.

§4. Impossibility for S to be a smooth abelian surface

We continue to employ the same notation given in the previous sections 2 and 3. In this section, we show that each surface S (found at the beginning of section 3) is not a smooth abelian surface if $\Phi : X \to W$ is a Calabi-Yau threefold of type $\mathrm{II}_0 K$. This completes the proof of Main Theorem (2).

Thoughout this section, assuming the contrary that S is a smooth abelian surface, we shall derive a contradiction.

For simplicity, we denote $\tilde{G_S}$, G_S and g_S by $\tilde{G}$, G and $\tilde{g}$ respectively. Under this notation, we have $T = S/G$, $W = T/\langle g \rangle = S/\tilde{G}$ and $I = \mathrm{ord}\,(g) = \mathrm{ord}\,(\tilde{g})$. As before, we denote by $q_T : S \to T$ the natural quotient morphism. This has an equivariant action of $\langle \tilde{g} \rangle$ and $\langle g \rangle$. Recall also that all the possibilities of G are listed up in (1.3)(4).

The next Lemma is shown by [O2].

Lemma (4.1). *I is either 2, 3, 4, 6, or 12.*

By virtue of this Lemma, the next two Claims will give a contradiction.

Key Claim (4.2). *I is not divided by 2.*

Key Claim (4.3). $I \neq 3$.

The following obvious lemma and its corollaries will be frequently used to prove these claims.

Lemma (4.4). *Let $q : S_1 \to S_2$ be a surjective finite morphism between normal projective surfaces with $K_{S_1} \equiv 0$ and $K_{S_2} \equiv 0$. Then q ramifies only at finitely many points.*

Corollary (4.5). *The quotient map $S \to W (= S/\tilde{G})$ ramifies only at finitely many points. In particular, $S^{\tilde{G_S}}$ is a finite set.*

Corollary (4.6). *Let h be a non-Gorenstein involution in $\tilde{G}$. Then, $S^h = \phi$. In particular, if $I = 2k$ is even, then $S^{\tilde{g}^k} = \phi$ and $S^{\tilde{g}} = \phi$.*

Proof. Assuming $S^h \neq \phi$, we take a point P in S^h. Since h is an involution with $h^*\omega_S = -\omega_S$, it follows that $h = \mathrm{diag}\,(-1, 1)$ under appropriate coordinates (x, y) of S around P. But then h would have a fixed curve $(x = 0)$, contradiction. q.e.d. of (4.6).

Corollary (4.7). *If I is either 2, 3, or 4, then $T^g \neq \phi$. If $I = pq$ where $p = 2$ or 4 and $q = 3$, then $T^{g^p} \neq \phi$ and $T^{g^q} \neq \phi$. Moreover, if I is either 2 or 4, then $(\phi \neq) T^g \subset \mathrm{Sing}\,(T)$.*

Proof. Since I is the least common multiple of the local canonical indices of W, the first part of the assertion is obvious. Assume that I is either 2 or 4. The first half part shows $T^g \neq \phi$. Assume the contrary that there is a smooth point Q in T^g. Then, arguing similarly as in (4.6), we see that $g^{I/2} = \mathrm{diag}\,(-1, 1)$ under appropriate local coordinates around P. Then, $g^{I/2}$ has a fixed curve. On the other hand, Lemma (4.4) shows $T \to W (= T/\langle g \rangle)$ has no ramification divisor, contradiction. q.e.d. of (4.7).

We return back to the key claims (4.2) and (4.3).

Proof of Key Claim (4.2).

Assume the contrary that $I = 2k$ for some integer k. We set $h := \tilde{g}^k$. Then h is a non-Gorenstein involution on S. Dividing into the following five cases, we shall derive a contradiction:
Case 1. $G \simeq \mathbb{Z}_3$ or $\mathbb{Z}_3 \times \mathbb{Z}_3$,
Case 2. $G \simeq \mathbb{Z}_6$,
Case 3. $G \simeq \mathbb{Z}_2$,
Case 4. $G \simeq \mathbb{Z}_2 \times \mathbb{Z}_2$,
Case 5. $G \simeq \mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_4$.

Case 1. Since g acts on the set B consisting of nine singular points of type A_3 on T ((1.3)(4)), $\langle \tilde{g} \rangle$ acts on $q_T^{-1}(B)$. Since $\sharp q_T^{-1}(B)$ is either 9 or 27, h has a fixed points. This contradicts (4.6).

Case 2. Consider the unique singular point Q of type A_5 on T ((1.3)(4)). Then, $q_T^{-1}(Q)$ consists of one point, say, P. Since $g(Q) = Q$, it follows that $\tilde{g}(P) = P$. But this contradicts (4.6).

Case 3. By (4.7), $T^{g^k} \neq \phi$. On the other hand, since g^k is a non-Gorenstein involution on T, the same argument as in (4.7) implies that $T^{g^k} \subset \mathrm{Sing}\,(T)$. Let $Q \in T^{g^k}$. Then Q is a singular point of type A_1 and then $q_T^{-1}(Q)$ consists of one point, say, P ((1.3)(4)). But then $h(P) = P$, contradiction.

Case 4. The same argument as in case 3 shows that $T^{g^k} \neq \phi$ and $T^{g^k} \subset \mathrm{Sing}\,(T)$. Let $Q \in T^{g^k}$. Then, Q is a singular point of type A_1 and $q_T^{-1}(Q)$ is written as $\{P, r(P)\}$ for some point P and a translation r in G ((1.3)(4)). Since h acts on this set, we have either $h(P) = P$ or $h(P) = r(P)$. The first equality contradicts (4.6). Consider the second case. Set $h' := r \circ h$. Then $h^*\omega_S = -\omega_S$. Since the translation subgroup of G is just $\langle r \rangle$ and since $h^{-1} \circ r \circ h$ is a translation in G (because G is a normal subgroup of $\tilde{G}$), it follows that $h^{-1} \circ r \circ h \in \langle r \rangle$ and then $\langle r, h \rangle = \langle r \rangle \times \langle h \rangle \simeq (\mathbb{Z}_2)^2$. Thus h' is a non-Gorenstein involution with $h'(P) = P$. But this contradicts (4.6).

Case 5. We treat the following three cases separately:
Case 5a. $3|I$, *Case 5b.* $I = 4$, and *Case 5c.* $I = 2$.

Case 5a. In this case, $I = 6m$ for some integer m. Set $j := \tilde{g}^m$. This is of order 6. Since g acts on the set consisting of 4 singular points of type A_3 on T ((1.3)(4)), j^2 acts on the inverse image of these points. This consists of either 4 or 8 points. Thus, j^2 has a fixed point among these points. Let P be such a fixed point. Then, $j^2(P) = P$. Since $(j^2)^*\omega_S = \zeta_3\omega_S$ and j^2 has at most finite fixed points by (4.5), an easy coordinate calculation shows that $j^2 = \mathrm{diag}\,(\zeta_3^2, \zeta_3^2)$ under appropriate global coordinates (x, y) around P. Thus, the eigen value of the matrix part of j is in $\{\zeta_3, -\zeta_3\}$. Thus, j has a fixed point on S, say Q. Since $h = j^3$, Q is also a fixed point of h. But this contradicts (4.6).

Case 5b. By (4.7), we can take a point Q in T^g. Again by (4.7) and (1.3)(4), Q is either a singular point of type A_3 or of type A_1.

If Q is a singular point of type A_3, then $q_T^{-1}(Q)$ is written as $\{P\}$ (in the case when $G \simeq \mathbb{Z}_4$) and $\{P, r(P)\}$ for a translation r in G (in the case when $G \simeq \mathbb{Z}_2 \times \mathbb{Z}_4$). In the first case, we have $\tilde{g}(P) = P$. But this contradicts (4.6). In the second case, we have either $\tilde{g}(P) = P$ or $\tilde{g}(P) = r(P)$. Since r is of order two, in each case, we get $h(P) = \tilde{g}^2(P) = P$, contradiction.

If Q is a singular point of type A_1, then $q_T^{-1}(Q)$ is written as $\{P, u^2(P)\}$ (if $G = \langle u \rangle \simeq \mathbb{Z}_4$) and $\{P, u^2(P), r(P), r \circ u^2(P)\}$ (if $G \simeq \mathbb{Z}_2 \times \mathbb{Z}_4$). In the second case, r is the unique translation in G and u is some (suitable) generator of G.

In anyway, we have $\tilde{g}(P) = P$ or $\tilde{g}(P) = t(P)$, where t is an involution in G. Thus, $h(P) = \tilde{g}^2(P) = P$, contradiction.

Case 5c. First consider the case $G = \langle u \rangle \simeq \mathbb{Z}_4$.

Since $\tilde{G} = \langle u \rangle \rtimes \langle \tilde{g} \rangle$ is of order 8, elementary group theory shows that $\tilde{G}$ is isomorphic to either

(1) D_8, the dihedral group of order 8, or
(2) $\mathbb{Z}_4 \times \mathbb{Z}_2$.

Assume first that $\tilde{G} \simeq D_8$. Then, $\tilde{g} \circ u$ is a non-Gorenstein involution. Take a point Q in T^g. Then, Q is a singular point either of type A_3 or of type A_1.

If Q is of type A_3, then $q_T^{-1}(Q) = \{P\}$, a one point set. But then $\tilde{g}(P) = P$, contradiction.

If Q is of type A_1, then $q_T^{-1}(Q)$ is written as $\{P, u(P)\}$ and $\tilde{g}$ stabilizes this set. If $\tilde{g}(P) = P$, then we get the same contradiction as before. If $\tilde{g}(P) = u(P)$, then $\tilde{g} \circ u(P) = P$. Since $\tilde{g} \circ u$ is a non-Gorenstein involution, we again get a contradiction. In any case, we found a contradiction if $\tilde{G} \simeq D_8$.

Next consider the case when $\tilde{G} \simeq \mathbb{Z}_4 \times \mathbb{Z}_2$, that is, $\tilde{G} = \langle u \rangle \times \langle \tilde{g} \rangle$. Then $\langle u \rangle \simeq \tilde{G}/\langle \tilde{g} \rangle$ acts on $\overline{p_2} : (E \times S)/\langle \tilde{g} \rangle \to S/\langle \tilde{g} \rangle$. Note that $(E \times S)/\langle \tilde{g} \rangle$ is also a smooth threefold, because $S^{[\langle \tilde{g} \rangle]} = \phi$ by (4.6) so that $(E \times S)^{[\langle \tilde{g} \rangle]} = \phi$.

CLAIM. $(E \times S/\langle \tilde{g} \rangle)^{[\langle u \rangle]} = \phi$.

Proof of Claim. Since u is of order 4, it is sufficient to show that

$$(E \times S/\langle \tilde{g} \rangle)^{u^2} = \phi.$$

Assume the contrary that $P \in (E \times S/\langle \tilde{g} \rangle)^{u^2}$. Set $\overline{p_2}(P) = Q$. Then $u^2(Q) = Q$. Thus u^2 acts on the fiber $E_Q := (\overline{p_2})^{-1}(Q)$. On the other hand, the fiber of $E \times S \to (E \times S/\langle \tilde{g} \rangle)$ over Q is written as $\{R, \tilde{g}(R)\}$ and u^2 also acts on this set. If $u^2(R) = \tilde{g}(R)$, then $u^2 \circ \tilde{g}(R) = R$ on S. But, since $u^2 \circ \tilde{g}$ is a non-Gorenstein involution on S, this contradicts (4.6). Thus $u^2(R) = R$. Let E_R be the fiber of $p_2 : E \times S \to S$ over R. Then the natural projection $E \times S \to E \times S/\langle \tilde{g} \rangle$ (of degree two) induces an isomorphism $E_R \simeq E_Q$, because $E_{\tilde{g}(R)}$ is also mapped to E_Q. Since u^2 gives an equivariant action on this isomorphism and since u^2 acts on E_R as a translation of order two by (1.3), we see that u^2 also acts on E_Q as a translation of order two. Thus $E_Q^{u^2} = \phi$. But this is absurd, because $P \in E_Q$ is a fixed point of u^2. q.e.d. of Claim.

Thus $Y := ((E \times S)/\langle \tilde{g} \rangle)/\langle u \rangle = (E \times S)/\tilde{G}$ is also a smooth threefold (with $\mathcal{O}_Y(K_Y) \simeq \mathcal{O}_Y$). Since X is birational to Y, X is connected with Y by flops. Then

X is also smooth and $\pi_1(X) \simeq \pi_1(Y)$ ([Ko2]). Thus X has a non-trivial finite étale covering, because so does Y. But this contradicts our assumption $\pi_1^{alg}(X) = \{1\}$. Therefore, we get a contradiction even in the case $G \simeq \mathbb{Z}_4$.

We consider the remaining case $G = \langle t \rangle \times \langle u \rangle \simeq \mathbb{Z}_2 \times \mathbb{Z}_4$. Reducing to the previous case $G \simeq \mathbb{Z}_4$, we find a contradiction.

Since the translation group of G is just $\langle t \rangle$ and since G is a normal subgroup of $\tilde{G}$, the same argument as before shows $\langle t \rangle$ is a normal subgroup of $\tilde{G}$. Thus $\tilde{G}/\langle t \rangle \simeq \langle u_1 \rangle \rtimes \langle \tilde{g_1} \rangle$, where $u_1 := u(\mathrm{mod}\, \langle t \rangle)$ and $\tilde{g_1} := \tilde{g}(\mathrm{mod}\, \langle t \rangle)$. Observe that u_1 is of order four and $\tilde{g_1}$ is of order two.

On the other hand, since $\langle t \rangle$ acts on $p_2 : E \times S \to S$, we get a new fiber space

$$\overline{p_2} : (E \times S)/\langle t \rangle \to S/\langle t \rangle,$$

on which $\langle u_1 \rangle \times \langle \tilde{g_1} \rangle$ gives an equivariant action. Since $\langle t \rangle$ is a translation group on both $E \times S$ and S, it follows that $(E \times S)/\langle t \rangle$ is an Abelian threefold and $S/\langle t \rangle$ is an Abelian surface. Set $S_1 := S/\langle t \rangle$ and $V := (E \times S)/\langle t \rangle$. Then, $T = S_1/\langle u_1 \rangle$ and $W = S_1/\langle u_1, \tilde{g}_1 \rangle$.

Observe that $\tilde{g}_1^* \omega_{S_1} = -\omega_{S_1}$, $u_1^* \omega_{S_1} = \omega_{S_1}$ and that u_1 acts on each fiber over $S_1^{u_1} (\neq \phi)$ as a translation of order 4. The last statement follows from (1.3) and a similar argument as is given in the last claim. Thus we can apply the same argument as in the previous case ($G \simeq \mathbb{Z}_4$) for $\overline{p_2} : (E \times S)/\langle t \rangle \to S/\langle t \rangle$ and $S_1 \to T \to W$ to get a contradiction. This finishes the proof of case 5c.

Now we have completed the proof of (4.2). Q.E.D. of (4.2).

Proof of Key Claim (4.3).

Assuming the contrary that $I = 3$ and dividing into the following five cases, we shall derive a contradiction.

Case 1. $G \simeq \mathbb{Z}_4$ or $\mathbb{Z}_2 \times \mathbb{Z}_4$,
Case 2. $G \simeq \mathbb{Z}_2$ or $\mathbb{Z}_2 \times \mathbb{Z}_2$,
Case 3. $G \simeq \mathbb{Z}_6$,
Case 4. $G \simeq \mathbb{Z}_3$,
Case 5. $G \simeq \mathbb{Z}_3 \times \mathbb{Z}_3$.

Case 1. Since g acts on the set of singular points of type A_3 and since this set consists of 4 points, g has a fixed point, say Q, in this set. Then, $\tilde{g}$ acts on $q_T^{-1}(Q)$. Since $q_T^{-1}(Q)$ consists of one or two points, $\tilde{g}$ has a fixed point in $q_T^{-1}(Q)$. Denote this point by 0. Since $\tilde{g}^* \omega_S = \zeta_3 \omega_S$, $\tilde{g}(0) = 0$ and since $\tilde{g}$ has only finitely many fixed points, we can apply [CC, also O2] to get $S \simeq E_{\zeta_3}^2$ and $\tilde{g} = \zeta_3^2$, the scalar multiplication by ζ_3^2. On the other hand, the stabilizer of 0 in G is a cyclic group of order 4. We denote this group by $\langle u \rangle$. Then $u = \mathrm{diag}\,(\zeta_4, \zeta_4^{-1})$ under appropriate global coordinates around 0. Set $H := \langle u, \tilde{g} \rangle$. Then, $H \subset \mathrm{Aut}\,(S, \{0\})$. Moreover H is a cyclic group of order 12, because $\tilde{g} = \zeta_3^2$ so that $u \circ \tilde{g} = \tilde{g} \circ u$. In particular $H \ni -1$. But this is impossible by Fujiki's classification ([Fu, Table 6]).

Case 2. Just by the same argument as in case 1, we see that $\tilde{g}$ has a fixed point 0 (over some singular point of type A_1 of T) and then $S = E_{\zeta_3}^2$ and $\tilde{g} = \zeta_3^2$. Set $\mathrm{Stab}_{\{0\}}(G) = \langle u \rangle$. This is a cyclic group of order two and $u = \mathrm{diag}\,(-1, -1)$ under appropriate global coordinates around 0. Thus $u \circ \tilde{g} = \tilde{g} \circ u$. Since $\tilde{G}$ gives an equivariant action on $p_2 : E \times S \to S$, $\tilde{g}$ and u act on the fiber $E := p_2^{-1}(0)$. Since $\tilde{g}$ is a Gorenstein automorphism of $E \times S$, the matrix part of $\tilde{g}$ on E is ζ_3^2 so that $\tilde{g}$ acts on E by

$$\tilde{g} : E \ni x \mapsto \zeta_3^2 x \in E,$$

if we fix an origin 0_E of E in $E^{\tilde{g}} (\neq \phi)$. On the other hand, by (1.3), the action of u on E is written as

$$u : E \ni x \mapsto x + P \in E,$$

where $P \in (E)_2 - \{0\}$. Since $u \circ \tilde{g} = \tilde{g} \circ u$ in $\tilde{G}$, we calculate

$$\tilde{g}(x) + \tilde{g}(P) = \tilde{g} \circ u(x) = u \circ \tilde{g}(x) = \tilde{g}(x) + P.$$

Thus, $P \in E^{\tilde{g}} = E^{\zeta_3} \subset (E)_3$. But this is impossible because $(E)_3 \cap ((E)_2 - \{0\}) = \phi$.

Case 3. Let Q be the unique singular point of type A_5 on T. Then, $q_T^{-1}(Q)$ consists of one point, say, 0. Since $g(Q) = Q$, it follows that $\tilde{g}(0) = 0$. Thus, just by the same argument as before, we get $\tilde{g} = \zeta_3^2$. Set $\mathrm{Stab}_{\{0\}}(G) = \langle u \rangle$. This is a cyclic group of order 6 and $u = \mathrm{diag}\,(\zeta_6, \zeta_6^{-1})$ under an appropriate global coordinates (x, y) around 0. it follows that $\tilde{g} \circ u^2 = \mathrm{diag}\,(1, \zeta_3)$. Then $\tilde{g} \circ u^2$ has a fixed curve $(y = 0)$, contradiction.

Case 4. Set $G = \langle u \rangle$. Since $\tilde{G} = \langle u \rangle \rtimes \langle \tilde{g} \rangle$ is of order 9, it follows that $\tilde{G} = \langle u \rangle \times \langle \tilde{g} \rangle$. Let Q be a point in T^g. Then $\sharp q_T^{-1}(Q)$ is either one or three. If $q_T^{-1}(Q) = \{P\}$, a one point set, then $\tilde{g}(P) = P$. If $q_T^{-1}(Q) = \{P_1, P_2, P_3\}$, then, $\tilde{g}(P_1) = P_j$ for some $j = 1, 2$, or 3. Since $\langle u \rangle$ acts on $\{P_1, P_2, P_3\}$ transitively, we find that $u^i(P_1) = P_j$ for some i. Set $h := u^{-i} \circ \tilde{g}$. Then, $h(P_1) = P_1$. Note that h is of order 3 and satisfies $h^* \omega_S = \zeta_3 \omega_S$ and $\tilde{G} = \langle u \rangle \times \langle h \rangle$. In addition, h and g give an equivariant action on $q_T : S \to T$. Thus, we may replace $\tilde{g}$ by h in the second case. Then $\tilde{g}(P_1) = P_1$ in each case. We regard this point P_1 as an origin of S and denote it by 0_S.

Since $\tilde{g}$ has only isolated fixed points ((4.5)), the same argument as before shows that $S = E_{\zeta_3}^2$ and $\tilde{g} = \zeta_3^2$. This implies $(S)^{\tilde{g}} \cap (S)^u = \phi$. (In fact, otherwise, choosing a point P in $(S)^{\tilde{g}} \cap (S)^u$, we find appropriate coordinates (x, y) around P such that $u = \mathrm{diag}\,(\zeta_3, \zeta_3^{-1})$. Then, $\tilde{g} \circ u = \mathrm{diag}\,(1, \zeta_3)$ has a fixed curve $(y = 0)$, contradiction.)

Since $\tilde{G}$ is a Gorenstein automorphism of $E \times S$ and gives an equivariant action on $p_2 : E \times S \to S$, $\tilde{g}$ induces an automorphism on the fiber $E := p_2^{-1}(0_S)$ whose matrix part is ζ_3^2. Thus $E = E_{\zeta_3}$ and then $E \times S = E_{\zeta_3}^3$. Moreover, choosing an origin 0_E of E in $E^{\tilde{g}}$, we get $\tilde{g} = \zeta_3^2$ on E. Now, taking $0 := (0_S, 0_E)$ as an origin of $E \times S = E_{\zeta_3}^3$, we have $\tilde{g} = \zeta_3^2$ on $E_{\zeta_3}^3$. Let us consider the quotient threefolds $(E_{\zeta_3})^3 / \langle \tilde{g} \rangle$ and its crepant resolution $\nu : Y \to (E_{\zeta_3})^3 / \langle \tilde{g} \rangle$. Note that $\langle u \rangle \simeq \tilde{G} / \langle \tilde{g} \rangle$ acts on $(E_{\zeta_3})^3 / \langle \tilde{g} \rangle$. Note also that ν is unique. (In fact, one of such ν is given by replacing each of 27 singular points of type $1/3(1, 1, 1)$ of $(E_{\zeta_3})^3 / \langle \tilde{g} \rangle$ by $\mathbb{P}^2$ and then has no flopping curves in the exceptional divisor.) Thus, $\langle u \rangle$ induces a regular action on Y.

Claim. *$\langle u \rangle$ acts freely on Y.*

Proof of Claim. Since $\mathrm{ord}\,(u) = 3$, it is sufficient to show that $Y^u = \phi$. Assume the contrary that $P \in Y^u$. Put $Q := \nu(P)$. Then $u(Q) = Q$. Denote the natural quotient map $E_{\zeta_3}^3 \to (E_{\zeta_3})^3 / \langle \tilde{g} \rangle$ by τ. Then, $Q \notin \tau((E_{\zeta_3}^3)^{\tilde{g}})$. (In fact, otherwise,

$\tau^{-1}(Q) = \{R\}(\subset (E_{\zeta_3}^3)^{\tilde{g}})$, a one point set. Thus, $u(R) = R$ and $\tilde{g}(R) = R$ on $(E_{\zeta_3})^3$. Set $R' := p_2(R)$. Then, $u(R') = R'$ and $\tilde{g}(R') = R'$, because $\tilde{G}$ gives an equivariant action on $p_2 : E \times S \to S$. But this contradicts $(S)^{\tilde{g}} \cap (S)^u = \phi$.)

Thus, $\tau^{-1}(Q)$ consists of three points, say, R_1, R_2 and R_3. Since $u(Q) = Q$, u acts on $\{R_1, R_2, R_3\}$. Since $\langle u \rangle$ acts freely on $E_{\zeta_3}^3$ by (1.3), we may assume without loss of generality that $u(R_1) = R_2$. On the other hand, $\{R_1, R_2, R_3\}$ is the orbit space of R_1 by $\langle \tilde{g} \rangle$, it follows that $\tilde{g}^i(R_1) = R_2$ for some $i = 1, 2$. Set again $R' := p_2(R_1)$. Then, $\tilde{g}^i(R') = u(R')(= p_2(R_2))$ so that $u^{-1} \circ \tilde{g}^i(R') = R'$. Since the matrix part of u^{-1} is $\mathrm{diag}\,(\zeta_3, \zeta_3^{-1})$ under some appropriate global coordinates around R', we calculate $u^{-1} \circ \tilde{g}^i = \mathrm{diag}\,(1, \zeta_3)$. Thus $u^{-1} \circ \tilde{g}^i$ has a fixed curve $(y = 0)$, contradiction. q.e.d. of Claim.

By this claim $Y/\langle u \rangle$ is a smooth threefold with $\mathcal{O}_{Y/\langle u \rangle}(K_{Y/\langle u \rangle}) \simeq \mathcal{O}_{Y/\langle u \rangle}$ and with non-trivial étale covering. On the other hand, by construction, our original Calabi-Yau threefold X is birational to Y and then is connected with Y by flops. Thus X is also smooth and $\pi_1(X) \simeq \pi_1(Y)$ by [Ko2]. This implies that X has also non-trivial finite étale covering. But this contradicts our assumption $\pi_1^{alg}(X) = \{1\}$.

Case 5. As in case (5c) in Claim (4.2), reducing to the previous case 4, we find a contradiction. Set $G = \langle t \rangle \times \langle u \rangle$, where t is a translation of order 3. Since the translation group of G is just $\langle t \rangle$, and G is a normal subgroup of $\tilde{G}$, the same argument as in case 4 in Claim (4.2) implies that $\langle t \rangle$ is a normal subgroup of $\tilde{G}$. Thus, $\tilde{G}/\langle t \rangle = \langle u_1 \rangle \times \langle \tilde{g_1} \rangle \simeq (\mathbb{Z}_3)^2$, where $u_1 := u(\mathrm{mod}\,\langle t \rangle)$ and $\tilde{g_1} := \tilde{g}(\mathrm{mod}\,\langle t \rangle)$.

By the way, since $\langle t \rangle$ acts on $p_2 : E \times S \to S$, we get a new fiber space

$$\overline{p_2} : (E \times S)/\langle t \rangle \to S/\langle t \rangle,$$

on which $\langle u_1 \rangle \times \langle \tilde{g_1} \rangle$ gives an equivariant action. Since $\langle t \rangle$ is a translation group on both $E \times S$ and S, $(E \times S)/\langle t \rangle$ is an Abelian threefold and $S/\langle t \rangle$ is an Abelian surface. Set $S_1 := S/\langle t \rangle$ and $V := (E \times S)/\langle t \rangle$. Then, $T = S_1/\langle u_1 \rangle$ and $W = S_1/\langle u_1, \tilde{g}_1 \rangle$. Moreover $\tilde{g}_1^* \omega_{S_1} = \zeta_3 \omega_{S_1}$ while $u_1^* \omega_{S_1} = \omega_{S_1}$. Now applying the same argument as in case 4 for $S_1 \to T \to W$, we find that $S_1 = E_{\zeta_3}^2$ and $\tilde{g}_1 = \zeta_3^2$ (after replacing $\tilde{g}_1$ by $u_1^i \circ \tilde{g}_1$ so that $S_1^{\tilde{g}_1} \neq \phi$ and then fixing the origin 0 of S_1 in $S_1^{\tilde{g}_1}(\neq \phi)$). Note that $\langle u_1, \tilde{g}_1 \rangle$ gives a Gorenstein action on V. Then letting $E := \overline{p_2}^{-1}(0)$ and applying the same argument as in case 4, we see that $E = E_{\zeta_3}$ and the action of $\tilde{g}_1$ on E is $\tilde{g}_1 = \zeta_3^2$ (after fixing an origin 0_E of E in $E^{\tilde{g}_1}(\neq \phi)$). Thus, regarding 0_E as an origin 0 of V, we get $\tilde{g}_1 = \zeta_3^2$ under appropriate global coordinates around 0. This together with [CC also O2] implies $V = E_{\zeta_3}^3$. Now again applying the same argument as in case 4 for $\overline{p_2} : V \to S_1$, we finally get a contradiction that X is birational to a smooth threefold Y with non-trivial finite étale covering.

Now this completes the proof of Claim (4.3).

Now we are done. Q.E.D. of Main Theorem (2).

References

[Be] A. Beauville, *Some remarks on Kähler manifolds with $c_1 = 0$, In Classification of algebraic and analytic manifolds (K. Ueno ed.)*, Progr. Math. **39**, (1983) 1-26.

[BM] E. Bombieri and D. Mumford, *Enriques' classification of surfaces in char. p. II, In Complex analysis and Algebraic Geometry, A collection of papers dedicated to K. Kodaira*, Iwanami Shoten Publisher, Tokyo, and Princeton University Press Cambridge, (1977) 23-42.

[BPV] W. Barth, C. Peters and A. Van de Ven,, *Compact complex surfaces*, Springer-Verlag, (1984).

[CC] F. Catanese and C. Ciliberto, *On the irregularity of cyclic coverings of algebraic surfaces, In Geometry of complex projective varieties, Cetrano (Italy) (A. Lanteri, M. Palleschi, D.C.Struppa ed.)*, Mediterranean Press, (1993) 89-115.

[Fu] A. Fujiki, *On finite automorphism groups of complex tori of dimension two*, Publ. RIMS Kyoto Univ. **24**, (1988) 1-99.

[Kt] T. Katsura, *Generalized Kummer surfaces and their unirationality in characteristic p*, J. Fac. Sci. Univ. Tokyo Sect. IA **34**, (1987) 1-41.

[Kw1] Y. Kawamata, *The cone of curves of algebraic varieties*, Ann. of Math. **119**, (1984) 603-633.

[Kw2] Y. Kawamata, *Minimal models and the Kodaira dimension of algebraic fiber spaces*, J. Reine Angew Math. **363**, (1985) 1-46.

[Kw3] Y. Kawamata, *The crepant blowing-ups of 3-dimensional canonical singularities and its application to degeneration of surfaces*, Ann. of Math. **127**, (1988) 93-163.

[Kw4] Y. Kawamata, *Termination of log-flips for algebraic 3-folds*, Intern. J. Math. **3**, (1992) 653-659.

[KMM] Y. Kawamata, K. Matsuda, K. Matsuki, *Introduction to the minimal model problem*, Adv. Stud. Pure Math. **10**, (1987) 283-360.

[Ko1] J. Kollár, *Higher direct images of dualizing sheaves I*, Ann. of Math. **123**, (1986) 11-42.

[Ko2] J. Kollár, *Flops*, Nagoya Math. J. **113**, (1989) 15-36.

[Ko3] J. Kollár et all, *Flips and abundance for algebraic threefolds, a summer seminar at the University of Utah (1991)*, Astérisque **211**, (1992).

[Mi] Y. Miyaoka, *The Chern classes and Kodaira dimension of a minimal threefolds*, Adv. St. Pure Math. **10**, (1987) 449-476.

[Mo] S. Mori, *Flip theorem and the existence of minimal models for 3-folds*, J. of AMS **1**, (1988) 117-253.

[Na1] N. Nakayama, *Elliptic fibrations over surfaces I, in Algebraic Geometry and Analytic Geometry ICM-90 Satellite Conference Proceedings*, Springer, (1991) 126-137.

[Na2] N. Nakayama, *Local structure of an elliptic fibration*, preprint (1992).

[Ni] V. V. Nikulin, *Finite groups of automorphisms of Kählerian surfaces of type K3*, Moscow Math. Soc., (1980) 71-137.

[O1] K. Oguiso, *On algebraic fiber space structures on a Calabi-Yau 3-fold*, Intern. J. Math. **4**, (1993) 439-465.

[O2] K. Oguiso, *On certain rigid fibered Calabi-Yau threefolds*, preprint, to appear Math. Z.

[O3] K. Oguiso, *On the complete classification of Calabi-Yau threefolds of Type* III_0, preprint, to appear in Proceedings Complex Geometry Trento 1995.

[Ro] S. Roan, *Minimal resolution of Georenstein orbifolds in dimension three*, preprint (1995).

[SW] N.I. Shepherd-Barron, P.M.H. Wilson, *Singular threefolds with numerically trivial first and second Chern class*, J. Alg. Geom. **3**, (1994) 265-281.

[Sh] V. V. Shokurov, *3-folds log flips*, Math. USSR. Izv. **26**, (1993) 95-202.

[W1] P. M. H. Wilson, *The existence of elliptic fibre space structures on Calabi-Yau threefolds*, Math. Ann. **300**, (1994) 693-703.

[W2] P. M. H. Wilson, *The role of c_2 in Calabi-Yau classification- a preliminary survey*, preprint, (1994).

[Z] De-Qi Zhang, *Logarithmic Enriques surfaces*, J. Math. Kyoto Univ. **31**, (1991) 419-466.

Keiji Oguiso
Department of Mathematical Sciences
University of Tokyo
Komaba, Tokyo
Japan
oguiso@ms.u-tokyo.ac.jp

An Invariant of Quadratic Forms over Schemes

Marek Szyjewski

Received: September 29, 1996

Communicated by Alexander S. Merkurjev

Abstract. A ring homomorphism $e^0 : W(X) \to EX$ from the Witt ring of a scheme X into a proper subquotient EX of the Grothendieck ring $K_0(X)$ is a natural generalization of the dimension index for a Witt ring of a field. In the case of an even dimensional projective quadric X, the value of e^0 on the Witt class of a bundle of an endomorphisms $\mathcal{E}$ of an indecomposable component $\mathcal{V}_0$ of the Swan sheaf $\mathcal{U}$ with the trace of a product as a bilinear form θ is outside of the image of composition $W(F) \to W(X) \to E(X)$. Therefore the Witt class of $(\mathcal{E}, \theta)$ is not extended.

Introduction

An important role in the quadratic form theory is played by the first (0-dimensional) cohomological invariant, the dimension index $e^0 : W(F) \to \mathbb{Z}/2\mathbb{Z}$, which maps a Witt class of a symmetric bilinear space $(\mathcal{V}, \beta)$ over a field F onto $\dim \mathcal{V} \mod 2$. A straightforward generalization of this map for symmetric bilinear spaces over rings or schemes, which assigns to a Witt class the rank of its supporting module or bundle, is commonly used. We define a better invariant e^0 in Section 1 below. It is a variant of the construction used in [8] and [9]. The map e^0 defined in Section 1 assigns to a Witt class of a symmetric bilinear space $(\mathcal{V}, \beta)$ a class $[\mathcal{V}]$ of $\mathcal{V}$ in the group EX, attached functorially to a scheme X. The group EX consist of the self-dual (i.e., stable under dualization) elements of the Grothendieck group $K_0(X)$ up to the split self-dual ones (i.e., sums of a class and its dual). Thus the rank $\mod 2$ may be obtained by passing to the generic stalk. The group EX carries much more information on the Witt group $W(X)$ than $\mathbb{Z}/2\mathbb{Z}$, and so does the map e^0 defined here when compared to the rank $\mod 2$. In particular, we use it here to show that certain Witt classes are not extended, i.e., are not of the form $(V \otimes \mathcal{O}_X, \beta \otimes 1)$ for a symmetric bilinear space (V, β) over a base field.

In the Section 1 basic facts on dualization in the Grothendieck group, definition and elementary properties of the group EX and map e^0 are given. Theorem 1.1 describes EX for a smooth curve X. In the geometric case (algebraically closed base field) the group EX appears to coincide with the Witt group $W(X)$ of curve X itself.

Moreover, it is shown that Witt classes of line bundles of order two in Picard group are not extended from the base field.

Section 2 contains a number of examples to show that EX may be actually computed: the affine space - Proposition 2.1.1, the projective space over a field - Proposition 2.1.3, the projective space over a scheme - Proposition 2.1.5.

The main objective of this paper is to prove that on the projective quadric of even dimension $d \neq 2$ defined by a hyperbolic form, there exist nonextended Witt classes. For this purpose, a close look at the Swan computation of the K-theory of a quadric hypersurface is needed. Section 3 contains all needed facts on Clifford algebras and modules, the construction of the Swan bundle, its behavior under dualization, and how to find a canonical resolution of a regular bundle.

In Section 4, we develop a combinatorial method for operations with resolutions using generating functions. Next we use the classical computation of the Chow ring of a split quadric X to establish the ring structure of $K_0(X)$. Theorem 4.3 gives the description of EX for a split quadric.

Thus, in Section 5, we show in Theorem 5.1 that, in case of even dimension $d > 2$ of a quadric the bundle of endomorphisms of each indecomposable component of the Swan bundle carries a canonical symmetric bilinear form, whose Witt class is not extended from the base field, since its invariant e^0 has a value outside the image of the composite map $W(F) \to W(X) \to EX$.

The first version of this paper contained only an explicit computation for a quadric of dimension 4. The referee made several suggestions for simplification of proofs and computations. These remarks led author to the present more general results. The author would like to thank very much the referee for generous assistance. The author is glad to thank Prof. W. Scharlau for helpful discussions and Prof. K. Szymiczek, who suggested several improvements of the exposition.

1 The group EX and the invariant e^0

1.1 Notation.

If X is a scheme with the structural sheaf $\mathcal{O}_X$, and $\mathcal{M}$, $\mathcal{N}$ are coherent locally free sheaves of $\mathcal{O}_X$-modules (vector bundles on X), $\phi : \mathcal{M} \to \mathcal{N}$ is a morphism, then we write

$$\mathcal{M}^\wedge = \mathcal{H}om_{\mathcal{O}_X}(\mathcal{M}, \mathcal{O}_X) \qquad \text{and} \qquad \phi^\wedge : \mathcal{N}^\wedge \to \mathcal{M}^\wedge$$

for the duals.

A *symmetric bilinear space* $(\mathcal{M}, \beta)$ consists of a coherent locally free sheaf $\mathcal{M}$ and a morphism $\beta : \mathcal{M} \to \mathcal{M}^\wedge$, which is self-dual, i.e. $\beta^\wedge = \beta$.

For a subbundle (a subsheaf which is locally a direct summand) $\iota : \mathcal{N} \to \mathcal{M}$ define its *orthogonal complement* $\mathcal{N}^\perp$ as a kernel of composition $\iota^\wedge \circ \beta$:

$$\mathcal{N}^\perp = \mathrm{Ker}(\mathcal{M} \xrightarrow{\beta} \mathcal{M}^\wedge \xrightarrow{\iota^\wedge} \mathcal{N}^\wedge).$$

Thus β induces an isomorphism $\mathcal{N}^\perp \cong (\mathcal{M}/\mathcal{N})^\wedge$.

There are two important special cases: the first, when $\mathcal{N}$ has trivial intersection with $\mathcal{N}^\perp$ or is non-singular, then β induces an isomorphism $\mathcal{N} \cong \mathcal{N}^\wedge$; the second, when $\mathcal{N} = \mathcal{N}^\perp$, and in this case $\mathcal{N}$ is said to be a *Lagrangian subbundle.*

A symmetric bilinear space $(\mathcal{M}, \beta)$ is said to be *metabolic* if it possesses a Lagrangian subbundle, i.e., if there exists an exact sequence

$$0 \to \mathcal{N} \xrightarrow{\iota} \mathcal{M} \xrightarrow{\iota^\wedge \circ \beta} \mathcal{N}^\wedge \to 0 \tag{1.1.1}$$

for some subbundle $\mathcal{N}$.

Direct sum and tensor product are defined in the set $B(X)$ of isomorphism classes of symmetric bilinear spaces, and in its Grothendieck ring $G(X)$ the set $M(X)$ of differences of classes of metabolic spaces forms an ideal.

The *Witt ring* $W(X)$ of X is the factor ring $G(X)/M(X)$. The *Witt class* of a symmetric bilinear space $(\mathcal{M}, \beta)$ is its coset in $W(X)$. Two symmetric bilinear spaces $(\mathcal{M}_1, \beta_1)$ and $(\mathcal{M}_2, \beta_2)$ are *Witt equivalent*, $(\mathcal{M}_1, \beta_1) \approx (\mathcal{M}_2, \beta_2)$ iff their Witt classes are equal, or - equivalently - iff $(\mathcal{M}_1 \oplus \mathcal{M}_2, \beta_1 + (-\beta_2))$ is metabolic. Each Witt class (an element of $W(X)$) contains a symmetric bilinear space and $X \mapsto W(X)$ is a contravariant functor on schemes, namely for arbitrary morphism $f : Y \to X$ of schemes the inverse image functor f^* induces a ring homomorphism $f^* : W(X) \to W(Y)$. In fact, $f^*(\mathcal{M}^\wedge) = (f^*(\mathcal{M}))^\wedge$ and f^* is an exact functor. In the affine case $X = \operatorname{Spec} R$, $Y = \operatorname{Spec} S$, $f^\# : R \to S$ a ring homomorphism, $f^* : W(X) \to W(Y)$ is simply the scalar extension $S \otimes_R - : W(R) \to W(S)$. Important special cases are localization or taking a stalk at a point $x \in X$, i.e., the inverse image for $\operatorname{Spec} \mathcal{O}_{X,x} \to X$, and the extension, i.e., taking the inverse image for the structure map $f : X \to \operatorname{Spec} F$ for a variety X over a field F. In the latter case a Witt class of the form $(f^*\mathcal{M}, f^*\beta) = (\mathcal{M} \otimes_F \mathcal{O}_X, \beta \otimes 1)$ for genuine bilinear space $(\mathcal{M}, \beta)$ over F is said to be *extended* or *induced* from the base field F.

1.2 Rank mod 2

In the affine case $X = \operatorname{Spec} R$, we write as usual $W(R)$ instead of $W(\operatorname{Spec} R)$. The classical situation is if $R = F$ is a field of characteristic different from two. In this case there is a ring homomorphism

$$e^0 : W(F) \to \mathbb{Z}/2\mathbb{Z}, \ e^0(\mathcal{M}, \beta) = \dim \mathcal{M} \mod 2,$$

known as dimension index. One may put the definition of e^0 into a K-theoretical framework as follows:

The map $e : (\mathcal{M}, \beta) \mapsto [\mathcal{M}]$ induces a ring homomorphism

$$G(F) \xrightarrow{e} K_0(F) \xrightarrow{\cong} \mathbb{Z}$$

which is surjective, since each vector space over F carries a symmetric bilinear form. Any metabolic form $(\mathcal{M}, \beta)$ is *hyperbolic*, i.e., the sequence 1.1.1 splits, and

$$(\mathcal{M}, \beta) \cong (\mathcal{N} \oplus \mathcal{N}, \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}).$$

Since each vector space is self-dual, $e(M(F)) = 2K_0(F) \cong 2\mathbb{Z}$, so e^0 is the induced ring homomorphism

$$W(F) \xrightarrow{e^0} K_0(F)/2K_0(F) \cong \mathbb{Z}/2\mathbb{Z}.$$

In general the forgetful functor $(\mathcal{M},\beta)\mapsto[\mathcal{M}]$ induces a ring homomorphism which in general neither is surjective nor maps $M(X)$ into $2K_0(X)$. We shall show below how to handle this using a proper subquotient of $K_0(X)$.

1.3 The involution ^ and the group $E(X)$

Denote by $\mathbf{P}(X)$ the category of locally free coherent $\mathcal{O}_X$-modules. The dualization functor ^ is an exact additive functor $^\wedge :\ \mathbf{P}(X)\to\mathbf{P}(X)^{op}$, which preserves tensor products and commutes with inverse image functors. Since

$$K_*(\mathbf{P}(X))=K_*(\mathbf{P}(X)^{op})=K_*(X),$$

the functor ^ induces a homomorphism on K-groups, known also as the Adams operation ψ^{-1} . We shall denote it by ^:

Definition 1.3.1. $^\wedge :\ K_*(X)\to K_*(X)$ is the homomorphism induced by the exact functor $^\wedge :\ \mathbf{P}(X)\to\mathbf{P}(X)^{op}$.

Proposition 1.3.2. *The homomorphism* $^\wedge :\ K_*(X)\to K_*(X)$ *enjoys the following properties:*

i) ^ *is an involution,* $^\wedge\circ{}^\wedge=1$;

ii) ^ *is a graded ring automorphism of* $K_*(X)$: $(\alpha\cdot\beta)^\wedge=\alpha^\wedge\cdot\beta^\wedge$;

iii) if $f:\ Y\to X$ *is a morphism of schemes, then* $f^*\circ{}^\wedge={}^\wedge\circ f^*$;

iv) if $i:\ Z\to X$ *is a closed immersion and* X *is regular of finite dimension, then* $(i^*(K_0(Z)))^\wedge=i^*(K_0(Z))$.

Proof. iv) Consider a finite resolution of $i^*(\mathcal{M})$ by vector bundles for a bundle $\mathcal{M}$ on Z. The stalk of this resolution at any point outside Z is exact, so its dual is exact. Hence the class of the alternating sum of the members of the resolution vanishes outside Z. □

We focus our attention on the Grothendieck group $K_0(X)$. The main object of this paper are the homology groups of the following complex:

$$\cdots\to K_0(X)\xrightarrow{1+^\wedge}K_0(X)\xrightarrow{1-^\wedge}K_0(X)\xrightarrow{1+^\wedge}K_0(X)\xrightarrow{1-^\wedge}\cdots \tag{1.3.1}$$

Definition 1.3.3.

$$EX=\mathrm{Ker}(1-{}^\wedge)/\,\mathrm{Im}(1+{}^\wedge)$$
$$E^-X=\mathrm{Ker}(1+{}^\wedge)/\,\mathrm{Im}(1-{}^\wedge).$$

We shall define a natural homomorphism $e^0:\ W(X)\to EX$. The group E^-X will play only a technical role here, although one may consider a natural map $L_{2k+1}(X)\to E^-X$. The EX is the group of "symmetric" or "self-dual" elements in $K_0(X)$ modulo "split self-dual" elements, i.e., elements of the form $[\mathcal{M}]+[\mathcal{M}^\wedge]$. The following observations are obvious:

Proposition 1.3.4. *i)* $\mathrm{Ker}(1-\wedge)$ *is a subring of* $K_0(X)$ *and the groups* $\mathrm{Im}(1+\wedge)$, $\mathrm{Ker}(1+\wedge)$, $\mathrm{Im}(1-\wedge)$ *are* $\mathrm{Ker}(1-\wedge)$*-modules;*

ii) EX *is a ring and* E^-X *is an* EX*-module;*

iii) an arbitrary morphism $f: Y \to X$ *of schemes induces a ring homomorphism* $f^*: EX \to EY$ *and an* EX*-module homomorphism* $f^*: E^-X \to E^-Y$ *;*

iv) for a regular Noetherian X, EX *and* E^-X *carry a natural filtration, induced by the topological filtration of* $K_0(X) = K'_0(X)$ *;*

v) $2EX = 0$ *and* $2E^-X = 0$.

□

Note that the forgetful functor $(\mathcal{M}, \beta) \mapsto [\mathcal{M}]$ induces a ring homomorphism $G(X) \mapsto K_0(X)$ which admits values in $\mathrm{Ker}(1-\wedge)$ and maps $M(X)$ onto $\mathrm{Im}(1+\wedge)$, since for a metabolic space $(\mathcal{M}, \beta)$ there is exact sequence 1.1.1, i.e., the equality $[\mathcal{M}] = [\mathcal{N}] + [\mathcal{N}^\wedge]$ holds in $K_0(X)$.

Definition 1.3.5. $e^0: W(X) \to EX$ is the ring homomorphism induced by the forgetful functor $(\mathcal{M}, \beta) \mapsto [\mathcal{M}]$.

This notion enjoys nice functorial properties.

Proposition 1.3.6. *Let* $f: X \to Y$ *be a morphism of schemes. Then the following diagram commutes:*

$$\begin{array}{ccc} W(X) & \xrightarrow{e^0} & EX \\ {\scriptstyle f^*}\big\uparrow & & \big\uparrow{\scriptstyle f^*} \\ W(Y) & \xrightarrow{e^0} & EY \end{array}$$

□

Example 1.3.7. Let X be an irreducible scheme with the function field $F(X)$, and let $j: \operatorname{Spec} F(X) \to X$ be the embedding of the generic point. Then there is a commutative diagram

$$\begin{array}{ccc} W(F(X)) & \xrightarrow{e^0} & E(F(X)) = \mathbb{Z}/2\mathbb{Z} \\ {\scriptstyle j^*}\big\uparrow & & \big\uparrow{\scriptstyle j^*} \\ W(X) & \xrightarrow{e^0} & EX \end{array}$$

and the composition $j^* \circ e^0 = e^0 \circ j^*$ is rank mod 2, usually used instead of e^0. Since $\mathcal{O}_X$ carries the standard symmetric bilinear form $< 1 >$, the surjection $j^*: EX \to \mathbb{Z}/2\mathbb{Z}$ splits canonically. The kernel of the map $j^*: EX \to \mathbb{Z}/2\mathbb{Z}$ has been used in [9]. It is easy to see that this kernel is a nilpotent ideal of a ring EX for a regular Noetherian X of finite dimension.

EXAMPLE 1.3.8. Retain the notation of example 1.3.7, and assume in addition that X is a variety over a field F, $\operatorname{char} F \neq 2$. Let $f : X \to \operatorname{SPEC} F$ be the structure map. Thus we have a commutative diagram:

$$\begin{tikzcd}
W(F(X)) \arrow[rrr, "e^0"] & & & \mathbb{Z}/2\mathbb{Z} \\
& W(X) \arrow[ul] \arrow[r, "e^0"] & EX \arrow[ur] & \\
W(F) \arrow[uu] \arrow[ur] \arrow[rrr, "e^0"] & & & \mathbb{Z}/2\mathbb{Z} \arrow[uu, "\mathrm{id}"'] \arrow[ul]
\end{tikzcd}$$

The values of $e^0 \circ f^*$ are inside the direct summand $\mathbb{Z}/2\mathbb{Z}[\mathcal{O}_X]$ of EX. If we produce a variety X with nontrivial (i.e., having more than two elements) EX, and a symmetric bilinear space with a nontrivial value of e^0, then the Witt class of this space must be non-extended.

1.4 CURVES

The case $\dim X = 1$ is exceptional for several reasons, so we treat it here as an illustration. The following theorem covers the classical case of (spectra of) Dedekind rings.

THEOREM 1.1. *Let X be an irreducible regular Noetherian scheme of dimension one. Then*

i) $EX = \mathbb{Z}/2\mathbb{Z}[\mathcal{O}_X] \oplus I$, where $I \cdot I = 0$ and I is canonically isomorphic to the group ${}_2\mathrm{Pic}(X)$ of the elements of order ≤ 2 in the Picard group;

ii) E^-X is canonically isomorphic to $\mathrm{Pic}(X)/2\,\mathrm{Pic}(X)$;

iii) the map $e^0 : W(X) \to EX$ is surjective.

Proof. The rank map (i.e., the restriction to the generic point) yields the splitting

$$K_0(X) = \mathbb{Z} \cdot [\mathcal{O}_X] \oplus \mathrm{F}^1 K_0(X)$$

where $0 \subset \mathrm{F}^1 K_0(X) \subset K_0(X)$ is the topological filtration on $K_0(X)$. The map $\wedge$ maps each direct summand onto itself.

Under assumptions on X the map $\bigwedge :\ \mathrm{F}^1 K_0(X) \to \mathrm{Pic}(X)$, induced by taking the highest exterior power of a bundle, is an isomorphism. An arbitrary element α of the group $\mathrm{F}^1 K_0(X)$ may be expressed as a difference of the classes of two bundles of the same rank r:

$$\alpha = [\mathcal{M}] - [\mathcal{N}].$$

The isomorphism $\bigwedge$ maps α onto the class of a line bundle $\mathcal{L}$,

$$\mathcal{L} = \bigwedge^r \mathcal{M} \otimes \bigwedge^r \mathcal{N}^\wedge$$

in $\mathrm{Pic}(X)$. The isomorphism $\bigwedge$ maps $[\mathcal{L}] - [\mathcal{O}_X]$ onto the class $\mathcal{L}$ in $\mathrm{Pic}(X)$, too. So, any element a of $\mathrm{F}^1 K_0(X)$ may be expressed as a difference of a line bundle and the trivial line bundle:

$$\alpha = [\mathcal{L}] - [\mathcal{O}_X].$$

Moreover, for arbitrary line bundles $\mathcal{L}_1,\ \mathcal{L}_2$

$$([\mathcal{L}_1] - [\mathcal{O}_X]) \cdot ([\mathcal{L}_2] - [\mathcal{O}_X]) = [\mathcal{L}_1 \otimes \mathcal{L}_2] - [\mathcal{O}_X].$$

Hence the involution $^\wedge$ acts on $\mathrm{F}^1 K_0(X)$ as taking the opposite, and it acts trivially on $\mathbb{Z} \cdot [\mathcal{O}_X]$. Therefore

$$\mathrm{Ker}(1 - {}^\wedge) = \mathbb{Z} \cdot [\mathcal{O}_X] \oplus 2\,\mathrm{F}^1 K_0(X) \quad , \qquad \mathrm{Im}(1 + {}^\wedge) = 2\mathbb{Z} \cdot [\mathcal{O}_X],$$
$$\mathrm{Ker}(1 + {}^\wedge) = \mathrm{F}^1 K_0(X) \quad , \qquad \mathrm{Im}(1 - {}^\wedge) = 2\,\mathrm{F}^1 K_0(X),$$

and assertions i), ii) follow.

To prove iii) note that a line bundle $\mathcal{L}$ which has order two in $\mathrm{Pic}(X)$ is isomorphic to its inverse $\mathcal{L}^\wedge$, so is automatically endowed with a nonsingular bilinear form $\mu : \mathcal{L} \to \mathcal{L}^\wedge$. This form must be symmetric locally at any point, hence is symmetric globally. Finally, e^0 maps the Witt class of $(\mathcal{L}, \mu) \oplus (\mathcal{O}_X, < 1 >)$ onto the class of $\mathcal{L}$ in $2\,\mathrm{Pic}(X)$ via $\bigwedge$. □

Remark 1.4.1. If R is a Dedekind ring, $X = \mathrm{Spec}\, R$, then $\mathrm{Pic}(X) = \mathrm{Pic}(R)$ is simply the ideal class group $H(R)$; the claim on the form of element of $F^1 K_0(X)$ is a consequence of the structural theorem for projective modules: if $\mathrm{rank}(P) = r$, then there exist fractional ideals $I_1, ..., I_r$ such that $P \cong I_1 \oplus ... \oplus I_r$; moreover, $P \cong R^{r-1} \oplus I_1 \cdot ... \cdot I_r \cong R^{r-1} \oplus \bigwedge^r P$. In this case $L^1(X) \cong \mathrm{Pic}(X)/2\,\mathrm{Pic}(X)$ and $L^1(X)$ is isomorphic to $E^- X$ via obvious generalization of e^0.

Remark 1.4.2. If ${}_2\mathrm{Pic}(X)$ is nontrivial, ${}_2\mathrm{Pic}(X) \neq 0$, then there exist non-extended Witt classes on X.

Corollary 1.4.3. *If X is a smooth projective curve of genus g over an algebraically closed field F, then*

i) if $\mathrm{char}\, F \neq 2$, *then* $EX \cong (\mathbb{Z}/2\mathbb{Z})^{1+2g}$;

ii) the degree map induces isomorphism $E^- X \cong \mathbb{Z}/2\mathbb{Z}$.

□

Remark 1.4.4. The result in Corollary 1.4.3. i) has been pointed out to author by W. Scharlau.

Remark 1.4.5. The proposition 2.1 of [3] states that for $F = \mathbb{C}$ the Witt group $W(X)$ of a smooth projective curve X is itself isomorphic to $(\mathbb{Z}/2\mathbb{Z})^{1+2g}$, but the proof remains valid for an arbitrary algebraically closed field F provided $\operatorname{char} F \neq 2$. So under assumptions of Corollary 1.4.3.i) the map $e^0 : W(X) \to EX$ is an isomorphism.

2 The map $e^0 : W(X) \to EX$ for certain quasiprojective X.

2.1

We shall show now that the group EX may be actually computed, and compare the result with known Witt rings. The simplest case is following:

Proposition 2.1.1. *If R is a regular ring, and $X = \mathbb{A}^n_R$, the affine space, then the inverse image functor f^* for the structure map $f : X \to \operatorname{Spec} R$ induces isomorphisms $W(R) \to W(X)$, $ER \to EX$, $E^-R \to E^-X$.*

Proof. By the homotopy property of K-theory, the map $f^* : K_0(R) \to K_0(X)$ is an isomorphism and commutes with $^\wedge$, so the assertion on E and E^- follows. The assertion on $W(X)$ is a consequence of the Karoubi theorem, see [6], Ch. VI.2, Corollary 2.2.2. □

Now let X be a quasiprojective variety over a field F, $\operatorname{char} F \neq 2$, with the structure map $f : X \to \operatorname{Spec} F$. Consider the commutative diagram

$$\begin{array}{ccc} W(X) & \xrightarrow{e^0} & EX \\ {\scriptstyle f^*}\uparrow & & {\scriptstyle f^*}\uparrow \\ W(F) & \xrightarrow[e^0]{} & EF \end{array} \tag{2.1.1}$$

We shall refer to "left f^*" and "right f^*" in 2.1.1 for various X.
Next, fix a projective embedding $i : X \to \mathbb{P}^n_F$ and denote:

$$1 = [\mathcal{O}_X] \text{ - the unit element in } K_0(X); \tag{2.1.2}$$

$$\mathcal{O}_X(-1) = i^* \mathcal{O}_{\mathbb{P}^n_F}(-1); \tag{2.1.3}$$

$$H = 1 - [\mathcal{O}_X(-1)] \text{ - the class of hyperplane section in } K_0(X). \tag{2.1.4}$$

We summarize some technicalities as follows:

Lemma 2.1.2. *If $d = \dim X$, then*

i) $H^{d+1} = 0$;

ii) $[\mathcal{O}_X(1)] = (1-H)^{-1} = \sum_{i=0}^{d} H^i \text{in } K_0(X)$ *(here $H^0 = 1$);*

iii) $H^\wedge = \dfrac{-H}{1-H} = -\sum_{i=1}^{d} H^i$;

iv) $(H^k)^\wedge = \left(\dfrac{-H}{1-H}\right)^k = (-1)^k H^k \sum_{i=0}^{d-k} \binom{k+i-1}{i} H^i$;

v) $(H^d)^\wedge = (-1)^d H^d$.

Proof. $H = 1 - [\mathcal{O}_X(-1)]$, so $[\mathcal{O}_X(-1)] = 1 - H$, $[\mathcal{O}_X(1)] = (1-H)^{-1}$, H being nilpotent. Thus $H^\wedge = 1 - [\mathcal{O}_X(1)] = ([\mathcal{O}_X(-1)] - 1)\cdot[\mathcal{O}_X(1)] = -H\cdot(1-H)^{-1}$ and $(H^k)^\wedge = (-H)^k(1-H)^{-k}$. □

In the case $i = \mathrm{id}$, $X = \mathbb{P}^d_F$, the family $1,\ H,\ \dots,\ H^d$ forms a basis of a free Abelian group $K_0(X)$, which allows us to compute EX, E^-X:

Proposition 2.1.3. *If $X = \mathbb{P}^d_F$, the projective space, then:*

i) both vertical arrows in the diagram 2.1.1 are isomorphisms;

ii) $E^-X = \mathbb{Z}/2\mathbb{Z}\cdot[H^d]$ for odd d and $E^-X = 0$ for even d.

Proof. The left f^* in the diagram 2.1.1 is an isomorphism by Arason's theorem [1]. Note that the statements on EX, E^-X are valid for $d = 0$, and - by Theorem 1.1 above - for $d = 1$. Consider $Y = \mathbb{P}^{d-1}_F$ and a closed embedding $k : Y \to X$ of Y as a hyperplane in X. There is an exact sequence

$$0 \to \mathbb{Z}\cdot H^d \to K_0(X) \xrightarrow{k^*} K_0(Y) \to 0$$

since $k^*\,\mathcal{O}_X(i) = \mathcal{O}_Y(i)$. Thus we have a short exact sequence of complexes:

$$\begin{array}{ccccccc}
\cdots \xrightarrow{1-\wedge} & K_0(Y) & \xrightarrow{1+\wedge} & K_0(Y) & \xrightarrow{1-\wedge} \cdots \\
 & {\scriptstyle k^*}\uparrow & & {\scriptstyle k^*}\uparrow & \\
\cdots \xrightarrow{1-\wedge} & K_0(X) & \xrightarrow{1+\wedge} & K_0(X) & \xrightarrow{1-\wedge} \cdots \\
 & \uparrow & & \uparrow & \\
\cdots \xrightarrow{1-(-1)^d} & \mathbb{Z}\cdot H^d & \xrightarrow{1+(-1)^d} & \mathbb{Z}\cdot H^d & \xrightarrow{1-(-1)^d} \cdots
\end{array}$$

and an induced exact sequence in homology. For even d this looks like

$$\cdots \to 0 \to E^-X \to E^-Y \to \mathbb{Z}/2\mathbb{Z}\cdot[H^d] \xrightarrow{\partial} EX \to EY \to 0 \to \cdots$$

and if - by induction - the proposition holds for Y, then ∂ maps the generator of $E^-Y = \mathbb{Z}/2\mathbb{Z}\cdot[H^{d-1}]$ onto $H^d \mod 2\mathbb{Z}\cdot H^d$, so the proposition holds for X: $E^-X = 0$, $k^* :\ EX \to EY$ is an isomorphism. In case of an odd d we have an exact sequence

$$\cdots \to 0 \to EX \to EY \xrightarrow{\partial} \mathbb{Z}/2\mathbb{Z}\cdot[H^d] \to E^-X \to E^-Y \to 0 \to \cdots$$

in homology. By induction $EY = \mathbb{Z}/2\mathbb{Z}\cdot[\mathcal{O}_X]$, $\partial = 0$, so $k^* :\ EX \to EY$ is an isomorphism. Thus $\mathbb{Z}/2\mathbb{Z}\cdot[H^d] \to E^-X$ is an isomorphism, since $E^-Y = 0$. □

Remark 2.1.4. The idea of this proof is due to the referee.

Proposition 2.1.5. *For an arbitrary variety Y let $X = \mathbb{P}^d_F \times Y$ and let $p_1 : X \to \mathbb{P}^d_F$, $p_2 :\ X \to Y$ be the projections. Then*

$$EX = (p_1^*(E(\mathbb{P}^d_F)) \otimes p_2^*(EY)) \oplus (p_1^*(E^-(\mathbb{P}^d_F)) \otimes p_2^*(E^-Y))$$
$$E^-X = (p_1^*(E(\mathbb{P}^d_F)) \otimes p_2^*(E^-Y)) \oplus (p_1^*(E^-(\mathbb{P}^d_F)) \otimes p_2^*(EY)).$$

Proof. By the projective bundle theorem p_1^*, p_2^* yield the identification $K_0(X) = K_0(\mathbb{P}_F^d) \otimes K_0(Y)$. Denote

$$A = \operatorname{Ker}(K_0(\mathbb{P}_F^d) \xrightarrow{1-\wedge} K_0(\mathbb{P}_F^d)), \qquad B = (1-\wedge)K_0(\mathbb{P}_F^d).$$

The complex 1.3.1 for $X = \mathbb{P}_F^d \times Y$ may be included into the short exact sequence of complexes:

$$\begin{tikzcd}
\cdots \arrow[r, "1-\wedge"] & B \otimes K_0(Y) \arrow[r, "1+\wedge"] & B \otimes K_0(Y) \arrow[r, "1-\wedge"] & \cdots \\
\cdots \arrow[r, "1-\wedge"] & K_0(X) \arrow[u, "(1-\wedge)\otimes 1"] \arrow[r, "1+\wedge"] & K_0(X) \arrow[u, "(1-\wedge)\otimes 1"] \arrow[r, "1-\wedge"] & \cdots \\
\cdots \arrow[r, "1-\wedge"] & A \otimes K_0(Y) \arrow[u] \arrow[r, "1+\wedge"] & A \otimes K_0(Y) \arrow[u] \arrow[r, "1-\wedge"] & \cdots
\end{tikzcd}$$

Note that $1 \pm \wedge$ restricted to $A \otimes K_0(Y)$ coincides with $1 \otimes (1 \pm \wedge)$ and induces $1 \otimes (1 \mp \wedge)$ on $B \otimes K_0(Y)$. Therefore the exact hexagon in homology

$$\begin{tikzcd}
 & EX \arrow[dr] & \\
A \otimes EY \arrow[ur] & & B \otimes E^-Y \arrow[d, "\partial_2"] \\
B \otimes EY \arrow[u, "\partial_1"] & & A \otimes E^-Y \arrow[dl] \\
 & E^-X \arrow[ul] &
\end{tikzcd}$$

breaks into short split exact sequences:

$$0 \to E(\mathbb{P}_F^d) \otimes E^-Y \to E^-X \to E^-(\mathbb{P}_F^d) \otimes EY \to 0 \tag{2.1.5}$$

$$0 \to E(\mathbb{P}_F^d) \otimes EY \to EX \to E^-(\mathbb{P}_F^d) \otimes E^-Y \to 0. \tag{2.1.6}$$

□

EXAMPLE 2.1.6. Put $d = 1$, $Y = \mathbb{P}_F^1$, i.e., $X = \mathbb{P}_F^1 \times \mathbb{P}_F^1$. Then

$$EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [H \boxtimes H] \tag{2.1.7}$$

$$E^-X = \mathbb{Z}/2\mathbb{Z} \cdot [H \boxtimes 1] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [1 \boxtimes H] \tag{2.1.8}$$

where $\boxtimes$ is induced by operation $\mathcal{F} \boxtimes \mathcal{G} = p_1^*\mathcal{F} \otimes p_2^*\mathcal{G}$. Since Witt ring is an invariant of birational equivalence in the class of smooth projective surfaces over a field F, $\operatorname{char} F \neq 2$ ([2], Theorem 3.4) and $X = \mathbb{P}_F^1 \times \mathbb{P}_F^1$ is birationally equivalent to $\mathbb{P}_F^2$, the left f^* in the diagram 2.1.1 is an isomorphism while the right f^* is not. This example shows that $e^0 : W(X) \to EX$ need not be surjective in general.

Remark 2.1.7. Probably there exists a skew symmetric bilinear space $(\mathcal{M}, \beta)$ on $X = \mathbb{P}^1_F \times \mathbb{P}^1_F$ such that $[\mathcal{M}] = [H \boxtimes H]$ in EX.

Remark 2.1.8. $X = \mathbb{P}^1_F \times \mathbb{P}^1_F$ may be embedded into $\mathbb{P}^3_F$ by Segre immersion as a quadric surface $x_0x_1 - x_2x_3 = 0$. In fact in the preliminary version of this paper this example was given using Swan's description of the K-theory of a quadric. The idea to use inverse images for projections was pointed out to author by the referee.

Remark 2.1.9. Note that we know $W(X)$ and EX for three quadrics of maximal index:

X	equation	$W(X)$	EX	E^-X
two points	$z_0^2 - z_1^2 = 0$	$W(F) \times W(F)$	$\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$	0
$\mathbb{P}^1_F$	$z_0^2 - z_1^2 + z_2^2 = 0$	$W(F)$	$\mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z}$
$\mathbb{P}^1_F \times \mathbb{P}^1_F$	$x_0x_1 - x_2x_3 = 0$	$W(F)$	$\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$	$\mathbb{Z}/2\mathbb{Z} \oplus \mathbb{Z}/2\mathbb{Z}$

We shall compute EX and E^-X for all projective quadrics of maximal index. To do this, some preparational work is required.

3 The Swan K-theory of a split projective quadric.

To compute EX and E^-X, we need some facts on dualization of vector bundles on quadrics. All needed information is known in fact, since indecomposable components of a Swan sheaf correspond to spinor representations. Nevertheless we give here complete proofs of the needed facts.

We shall apply the results of [11] in the simplest possible case of a split quadric: X is a projective quadric hypersurface over a field F, $\operatorname{char} F \neq 2$, defined by the quadratic form of maximal index.

3.1 Notation

Consider a vector space V with basis $v_0,\ v_1,\ \dots,\ v_{d+1}$ over a field F, $\operatorname{char} F \neq 2$. Denote $z_0,\ z_1,\ \dots,\ z_{d+1}$ the dual basis of $V^\wedge$. Let q be the quadratic form

$$q = \sum_{i=0}^{d+1} (-1)^i z_i^2.$$

Moreover, let $e_i = \frac{1}{2}(v_{2i} - v_{2i+1}),\ f_i = \frac{1}{2}(v_{2i} + v_{2i+1})$ for all possible values of i. Thus if d is even, $d = 2m$, then $e_0,\ f_0,\ e_1,\ f_1,\ \dots,\ e_m,\ f_m$ form a basis of V with the dual basis $x_0,\ y_0,\ x_1,\ y_1,\ \dots,\ x_m,\ y_m$ and

$$q = \sum_{i=0}^{m} x_i y_i.$$

If d is odd, $d = 2m+1$, then $f_0,\ e_1,\ f_1,\ \dots,\ e_m,\ f_m,\ v_{d+1}$ form a basis of V with the dual basis $x_0,\ y_0,\ x_1,\ y_1,\ \dots,\ x_m,\ y_m,\ z_{d+1}$ and

$$q = \sum_{i=0}^{m} x_i y_i + z_{d+1}^2.$$

We shall compute EX and $E^{-}X$ for a d-dimensional projective quadric X defined by equation $q = 0$ in $\mathbb{P}_F^{d+1}$, i.e., for

$$X = \operatorname{Proj} S(V^{\wedge})/(q) \cong \operatorname{Proj} F[z_0,\ z_1,\ \ldots,\ z_{d+1}]/(q).$$

3.2 The Clifford algebra

In case of an odd $d = 2m + 1$ the even part $C_0 = C_0(q)$ of the Clifford algebra $C(q)$ is isomorphic to the matrix algebra $M_N(F)$, where $N = 2^{m+1}$. In particular, $K_p(C_0) \cong K_p(F)$.

In case of an even $d = 2m$, the algebra C_0 has the center $F \oplus F \cdot \delta$, where $\delta = v_0 \cdot v_1 \cdot \ldots \cdot v_{d+1}$ and $\delta^2 = 1$. Thus $\frac{1}{2}(1+\delta)$, $\frac{1}{2}(1-\delta)$ are orthogonal central idempotents of C_0, so

$$C_0 = \frac{1}{2}(1+\delta)C_0 \oplus \frac{1}{2}(1-\delta)C_0$$

where each direct summand is isomorphic to the matrix algebra $M_{2^m}(F)$. For even $d = 2m$, consider the principal antiautomorphism $\Im : C_0 \to C_0$:

$$\Im(w_1 \cdot w_2 \cdot \ldots \cdot w_k) = (-1)^k w_k \cdot w_{k-1} \cdot \ldots \cdot w_1 \text{ for } w_1, w_2,\ \ldots, w_k \in V.$$

Note that

$$\Im(\delta) = (-1)^{m+1}\delta. \tag{3.2.1}$$

Moreover, for every anisotropic vector $w \in V$, the reflection $\alpha \mapsto -w\alpha w^{-1}$ in V induces an automorphism ρ_w of C_0, which interchanges δ with its opposite:

$$\rho_w(\delta) = -\delta. \tag{3.2.2}$$

Regarding subscripts $i \mod 2$ denote

$$P_i = (1 + (-1)^i\delta)C_0 \text{ for even } d.$$

Lemma 3.2.1. *In case of an even $d = 2m$:*

i) the involution $\Im$ of the algebra C_0 provides an identification of the left C_0-module $P_i{}^{\wedge} = \operatorname{Hom}_F(P_i, F)$ with the right C_0-module P_{i+m+1};

ii) for any anisotropic vector $w \in V$, the reflection ρ_w interchanges P_i's: $\rho_w(P_i) = P_{i+1}$.

□

3.3 Swan K-theory of a quadric

Recall some basic facts and notation of [11]. Denote by C_1 the odd part of the Clifford algebra $C(q)$. We shall use $\mod 2$ subscripts in C_i. Recall the definition of the Swan bundle $\mathcal{U}$. Put

$$\phi = \sum_{i=0}^{d+1} z_i \otimes v_i, \qquad \phi \in \Gamma(X, \mathcal{O}_X(1) \otimes V).$$

The complex

$$\cdots \xrightarrow{\phi\cdot} \mathcal{O}_X(-n)\otimes C_{n+d+1} \xrightarrow{\phi\cdot} \mathcal{O}_X(1-n)\otimes C_{n+d} \\ \xrightarrow{\phi\cdot} \mathcal{O}_X(2-n)\otimes C_{n+d-1} \xrightarrow{\phi\cdot} \cdots \tag{3.3.1}$$

is exact and locally splits ([11], Prop. 8.2.(a)).

Definition 3.3.1.

$$\mathcal{U}_n = \mathrm{Coker}(\mathcal{O}_X(-n-2)\otimes C_{n+d+3} \xrightarrow{\phi\cdot} \mathcal{O}_X(-n-1)\otimes C_{n+d+2}),$$
$$\mathcal{U} = \mathcal{U}_{d-1}.$$

Since the complex 3.3.1 is - up to a twist - periodical with period two, we have

$$\mathcal{U}_{n+2} = \mathcal{U}_n(-2).$$

Consider the exact sequences

$$\mathcal{O}_X(-n-2)\otimes C_{n+d+3} \xrightarrow{\phi\cdot} \mathcal{O}_X(-n-1)\otimes C_{n+d+2} \to \mathcal{U}_n \to 0$$

for two consecutive values n; twist the first one by 1. For any anisotropic vector $w \in V$ the isomorphism given by right multiplication by $1\otimes w$ fits into the commutative diagram:

$$\begin{array}{ccccccc}
\mathcal{O}_X(-n-2)\otimes C_{n+d+4} & \xrightarrow{\phi\cdot} & \mathcal{O}_X(-n-1)\otimes C_{n+d+3} & \longrightarrow & \mathcal{U}_{n+1}(1) & \longrightarrow & 0 \\
\cong\big\downarrow\cdot 1\otimes w & & \cong\big\downarrow\cdot 1\otimes w & & & & \\
\mathcal{O}_X(-n-2)\otimes C_{n+d+3} & \xrightarrow{\phi\cdot} & \mathcal{O}_X(-n-1)\otimes C_{n+d+2} & \longrightarrow & \mathcal{U}_n & \longrightarrow & 0.
\end{array}$$

Thus we have proved the following lemma:

Lemma 3.3.2.

$$\mathcal{U}_{n+1} \cong \mathcal{U}_n(-1) \quad \textit{and} \quad \mathcal{U}_n \cong \mathcal{U}_0(-n)$$

for arbitrary integer n.

□

There is an exact sequence

$$0 \to \mathcal{U}_0 \xrightarrow{\phi} \mathcal{O}_X \otimes C_0 \to \mathcal{U}_{-1} \to 0 \tag{3.3.2}$$

where an isomorphism $\cdot(1\otimes w)$ was used to replace $\mathcal{O}_X \otimes C_1$ by $\mathcal{O}_X \otimes C_0$ for even d.

Lemma 3.3.3. $\mathrm{End}_X(\mathcal{U}_n) \cong C_0$ *acts on* $\mathcal{U}_n$ *from the right.*

Proof. [11], Lemma 8.7. □

3.4

We are now ready to compute $\mathcal{U}_n{}^\wedge$.

Lemma 3.4.1. *$\mathcal{U}_n{}^\wedge \cong \mathcal{U}_n(2n+1)$, in particular $\mathcal{U}^\wedge \cong \mathcal{U}(2d-1)$.*

Proof. We have chosen a basis $v_0,\ v_1,\ \dots,\ v_{d+1}$ of V in 3.1 above. The set of naturally ordered products of several v_i's in an even number forms a basis of C_0. Define a quadratic form Q on C_0 as follows: let the distinct basis products be orthogonal to each other and

$$Q(v_{i_1}\cdot v_{i_2}\cdot\ldots\cdot v_{i_k}) = q(v_{i_1})\cdot q(v_{i_2})\cdot\ldots\cdot q(v_{i_k}).$$

The form Q is nonsingular and defines - by scalar extension - a nonsingular symmetric bilinear form Δ on $\mathcal{O}_X\otimes C_0$. Since $(q(v_i))^2 = 1$, a direct computation shows that $\mathrm{Im}(\mathcal{O}_X(-1)\otimes C_1 \xrightarrow{\phi} \mathcal{O}_X\otimes C_0) = \phi\cdot\mathcal{U}_0 \cong \mathcal{U}_0$ is a totally isotropic subspace of $\mathcal{O}_X\otimes C_0$. Therefore

$$\mathcal{U}_0 \cong \phi\cdot\mathcal{U}_0 = (\phi\cdot\mathcal{U}_0)^\perp \cong ((\mathcal{O}_X\otimes C_0)/(\phi\cdot\mathcal{U}_0))^\wedge \cong \mathcal{U}_{-1}{}^\wedge$$

follows quickly from sect. 1.1 above and the exactness of 3.3.2. Thus

$$\mathcal{U}_0{}^\wedge \cong \mathcal{U}_{-1} \cong \mathcal{U}_0(1)$$

and, in general

$$\mathcal{U}_n{}^\wedge \cong (\mathcal{U}_0(-n))^\wedge \cong \mathcal{U}_0{}^\wedge(n) \cong \mathcal{U}_0(n+1) \cong \mathcal{U}_n(2n+1).$$

□

Remark 3.4.2. This argument was pointed out to the author by the referee.

Corollary 3.4.3. *i)* $[\mathcal{U}^\wedge] = [\mathcal{U}(2d-1)]$ *and* $[\mathcal{U}(d-1)] + [\mathcal{U}(d-1)]^\wedge = 2d+1$ *in* $K_0(X)$*;*

ii) $\mathrm{rank}\,\mathcal{U} = \frac{1}{2}\dim C_0 = 2^d$.

□

In case of an even $d = 2m$ the algebra $\mathrm{End}_X(\mathcal{U}) = C_0$ splits into the direct product of subalgebras defined in 3.2 above: $C_0 = P_0 \times P_1$.

Definition 3.4.4. In case of an even d:

$$\mathcal{U}'_n = \mathcal{U}_n\otimes_{C_0}P_0,\ \mathcal{U}''_n = \mathcal{U}_n\otimes_{C_0}P_1,$$
$$\mathcal{U}' = \mathcal{U}\otimes_{C_0}P_0,\ \mathcal{U}'' = \mathcal{U}\otimes_{C_0}P_1.$$

Note that $\mathcal{U}_n = \mathcal{U}'_n \oplus \mathcal{U}''_n$, $\mathcal{U} = \mathcal{U}' \oplus \mathcal{U}''$. $\mathcal{U}'_0$ and $\mathcal{U}''_0$ correspond to spinor representation and we shall copy here the standard argument on dualization (compare [4], sect. 4.3).

In case of an even $d = 2m$ another property of ϕ and the quadratic form Q introduced in the proof of Lemma 3.4.1 may be verified by direct computation:

LEMMA 3.4.5. *In case of an even $d = 2m$*

i) *if m is even, then $P_i = (1 \pm \delta)C_0$ are orthogonal to each other, hence self-dual;*

ii) *if m is odd, then $P_i = (1 \pm \delta)C_0$ are totally isotropic, hence dual to each other;*

iii) $\phi(1 \pm \delta) = (1 \mp \delta)\phi$.

□

COROLLARY 3.4.6. *In case of an even $d = 2m$*

i) $\mathcal{U}'^{\wedge} \cong \mathcal{U}'(2d-1)$ and $\mathcal{U}''^{\wedge} \cong \mathcal{U}''(2d-1)$ *for even m;*

ii) $\mathcal{U}'^{\wedge} \cong \mathcal{U}''(2d-1)$ and $\mathcal{U}''^{\wedge} \cong \mathcal{U}'(2d-1)$ *for odd m;*

iii) $\mathrm{End}_X(\mathcal{U}') \cong \mathrm{End}_X(\mathcal{U}'') \cong M_{2^m}(F)$;

iv) *the exact sequence 3.3.2 splits into two exact parts*

$$0 \to \mathcal{U}_0' \xrightarrow{\phi\cdot} \mathcal{O}_X \otimes P_0 \to \mathcal{U}_0''(1) \to 0$$
$$0 \to \mathcal{U}_0'' \xrightarrow{\phi\cdot} \mathcal{O}_X \otimes P_1 \to \mathcal{U}_0'(1) \to 0$$

□

The standard way to determine indecomposable components is tensoring with the simple left module over an appropriate endomorphism algebra. We will use (from here onwards) superscript for the direct sum of identical objects.

DEFINITION 3.4.7.

i)	in case of an odd $d = 2m+1$	$\mathcal{V} = \mathcal{U} \otimes_{C_0} F^{2^{m+1}}$;
ii)	in case of an even $d = 2m$	$\mathcal{V}_0 = \mathcal{U}' \otimes_{M_{2^m}(F)} F^{2^m}$,
		$\mathcal{V}_1 = \mathcal{U}'' \otimes_{M_{2^m}(F)} F^{2^m}$.

For convenience we will use mod 2 subscripts in $\mathcal{V}_i$. Since $M_n(F) = (F^n)^n$ as a left $M_n(F)$-module, indecomposable components inherit properties of the Swan bundle: we have

PROPOSITION 3.4.8. *a) In case of an odd $d = 2m+1$:*

i) $\mathcal{U} \cong \mathcal{V}^{2^{m+1}}$;

ii) $\mathcal{V}^{\wedge} = \mathcal{V}(2d-1)$;

iii) $\mathrm{End}_X(\mathcal{V}) \cong F$ *and* $\mathrm{rank}\,\mathcal{V} = 2^m$;

iv) $[\mathcal{V}(d-1)] + [\mathcal{V}(d)] = 2^m$ *in* $K_0(X)$.

b) In case of an even $d = 2m$:

i) $\mathcal{U}' = \mathcal{V}_0^{2^m}$ and $\mathcal{U}'' = \mathcal{V}_1^{2^m}$;

ii) $\mathcal{V}_i{}^{\wedge} = \mathcal{V}_{i+m}(2d-1)$;

iii) $\mathrm{End}_X(\mathcal{V}_i) \cong F$ *and* $\mathrm{rank}\,\mathcal{V}_i = 2^{m-1}$

iv) $[\mathcal{V}_i(d-1)] + [\mathcal{V}_{i+1}(d)] = 2^m$ *in* $K_0(X)$.

□

In particular there is no global morphism $\mathcal{V}_i \to \mathcal{V}_{i+1}$.

Corollary 3.4.9. *In case of an even $d = 2m$ following identities hold in $K_0(X)$:*

i) $([\mathcal{V}_0] - [\mathcal{V}_1]) \cdot H = 0$;

ii) $([\mathcal{V}_0] - [\mathcal{V}_1]) \cdot [\mathcal{O}_X(n)] = [\mathcal{V}_0] - [\mathcal{V}_1]$;

iii) $([\mathcal{V}_0] - [\mathcal{V}_1])^\wedge = (-1)^m([\mathcal{V}_0] - [\mathcal{V}_1])$.

Proof. Proposition 3.4.8.b) *iv*) yields

$$[\mathcal{V}_0(d-1)] + [\mathcal{V}_1(d)] = [\mathcal{V}_1(d-1)] + [\mathcal{V}_0(d)].$$

Tensoring with $\mathcal{O}_X(-d)$ one obtains

$$[\mathcal{V}_0] - [\mathcal{V}_1] = ([\mathcal{V}_0] - [\mathcal{V}_1]) \cdot [\mathcal{O}_X(-1)],$$

hence i) and ii). Thus iii) results from 3.4.8. b) ii). □

Proposition 3.4.10. *$K_*(X)$ is a free $K_*(F)$-module of the rank $2m+2$; moreover*

i) in case of an odd $d = 2m+1$ the classes $[\mathcal{O}_X]$, $[\mathcal{O}_X(-1)]$, ..., $[\mathcal{O}_X(1-d)]$, $[\mathcal{V}]$ form a basis of $K_*(X)$;

ii) in case of an even $d = 2m$ the classes $[\mathcal{O}_X]$, $[\mathcal{O}_X(-1)]$, ..., $[\mathcal{O}_X(1-d)]$, $[\mathcal{V}_0]$, $[\mathcal{V}_1]$ form a basis of K_*X.

Proof. Apply Theorem 9.1 of [11]. □

We have expressed the action of $^\wedge$ on $K_0(X)$ in terms of a twist. We need a plain expression in order to determine EX and E^-X.

3.5

We recall here several facts known from section 6 of [11] needed for establishing plain formulas for the action of $^\wedge$.

Every regular sheaf $\mathcal{F}$ on X has a *canonical resolution* (infinite in general):

$$\cdots \to \mathcal{O}_X(-p)^{k_p} \to \cdots \to \mathcal{O}_X(-1)^{k_1} \to \mathcal{O}_X{}^{k_0} \to \mathcal{F} \to 0$$

where superscript k_p means a direct sum of k_p copies. One may compute the coefficients k_p and the differentials recursively as follows: put $\mathcal{Z}_{-1} = \mathcal{F}$. Since a regular sheaf is generated by its global sections, put $k_p = \dim \Gamma(X, \mathcal{Z}_{p-1}(p))$ and define $\mathcal{Z}_p$ as the twisted kernel in

$$0 \to \mathcal{Z}_p(p) \to \mathcal{O}_X^{k_p} \to \mathcal{Z}_{p-1}(p) \to 0.$$

Then $\mathcal{Z}_p(p+1)$ is a regular sheaf. Therefore the sequence

$$0 \to \mathcal{Z}_p(p) \to \mathcal{O}_X{}^{k_p} \to \cdots \mathcal{O}_X(p-1)^{k_1} \to \mathcal{O}_X(p)^{k_0} \to \mathcal{F}(p) \to 0$$

is exact. Twisting it by 1 one obtains an exact sequence of regular sheaves

$$0 \to \mathcal{Z}_p(p+1) \to \mathcal{O}_X(1)^{k_p} \to \cdots \to \mathcal{O}_X(p)^{k_1} \to \mathcal{O}_X(p+1)^{k_0} \to \mathcal{F}(p+1) \to 0.$$

Since the functor of global sections is exact on regular sheaves, there is following recurrence for $k_{p+1} = \dim \Gamma(X, \mathcal{Z}_p(p+1))$:

$$\dim \Gamma(X, \mathcal{F}(p+1)) - k_0 \cdot \dim \Gamma(X, \mathcal{O}_X(p+1)) + \cdots \\ + (-1)^{p-1} k_p \cdot \dim \Gamma(X, \mathcal{O}_X(1)) + (-1)^p k_{p+1} = 0. \quad (3.5.1)$$

In case of a -1- regular $\mathcal{F}$ to obtain an expression for $[\mathcal{F}] \in K_0(X)$ in terms of the basis from Proposition 3.4.10 one truncates the canonical resolution of $\mathcal{F}$:

$$0 \to \mathcal{Z}_{d-1} \to \mathcal{O}_X(1-d)^{k_{d-1}} \to \cdots \to \mathcal{O}_X(-1)^{k_1} \to \mathcal{O}_X{}^{k_0} \to \mathcal{F} \to 0$$

and replaces $\mathcal{Z}_{d-1}$ by $\mathcal{U} \otimes_{C_0} \mathrm{Hom}_X(\mathcal{U}, \mathcal{Z}_{d-1}) \cong \mathcal{Z}_{d-1}$. Then in $K_0(X)$

$$[\mathcal{F}] = \sum_{i=1}^{d-1} [\mathcal{O}_X(-i)] + [\mathcal{U} \otimes_{C_0} \mathrm{Hom}_X(\mathcal{U}, \mathcal{Z}_{d-1})].$$

Depending on the parity of d we have there

$$[\mathcal{U} \otimes_{C_0} \mathrm{Hom}_X(\mathcal{U}, \mathcal{Z}_{d-1})] = a[\mathcal{V}]$$

or

$$[\mathcal{U} \otimes_{C_0} \mathrm{Hom}_X(\mathcal{U}, \mathcal{Z}_{d-1})] = a[\mathcal{V}_0] + b[\mathcal{V}_1],$$

where the integers a, b in turn depend on the decomposition of $\mathrm{Hom}_X(\mathcal{U}, \mathcal{Z}_{d-1})$ into a direct sum of simple left C_0 - modules. Conversely, for a given $\mathcal{F}$ the equality

$$[\mathcal{F}] = \sum_{i=1}^{d-1} [\mathcal{O}_X(-i)] + W$$

holds, where W is either $a[\mathcal{V}]$ or $a[\mathcal{V}_0] + b[\mathcal{V}_1]$, then k_0 is the Euler characteristic $\sum (-1)^i \dim \mathrm{H}^i(X, \mathcal{F})$ of $\mathcal{F}$. So if $\mathcal{F}$ is regular, then $k_0 = \dim \Gamma(X, \mathcal{F})$. Next, $\mathcal{Z}_0(1) = \mathrm{Ker}(\mathcal{O}_X(1)^{k_0} \to \mathcal{F}(1))$ is regular, and iterating yields that for a regular $\mathcal{F}$ the congruence

$$[\mathcal{F}] \equiv [\mathcal{O}_X(-i)] \mod \mathrm{Im}(K_0(C_0) \to K_0(X))$$

holds if and only if integers k_i satisfy 3.5.1. In case of an odd $d = 2m+1$, in order to express a class $[\mathcal{F}]$ of a regular sheaf $\mathcal{F}$ in terms of the basis of Proposition 3.4.10, it is enough to know the dimensions of $\Gamma(X, \mathcal{F}(i))$ for $i = 0, 1, 2, \ldots, d-1$ to determine the k_i's and the rank $\mathcal{F}$ to determine the coefficient a of $[\mathcal{V}]$. An analogous statement remains valid for an arbitrary sheaf $\mathcal{F}$ with Euler characteristic of $\mathcal{F}(i)$ in place of $\dim \Gamma(X, \mathcal{F}(i))$. In case of an even $d = 2m$, in view of Corollary 3.4.9 ii) and Proposition 3.4.8 ii), the bundles $\mathcal{V}_0$ and $\mathcal{V}_1$ have the same Euler characteristic, rank and even the highest exterior power. Thus, without special considerations, one can express a class $[\mathcal{F}]$ in terms of basis of the Proposition 3.4.10 only up to a multiple of $[\mathcal{V}_0] - [\mathcal{V}_1]$.

4 The group EX for a split projective quadric

4.1 A Poincaré series of a sheaf

We introduce here a method for the determination of the coefficients of the canonical resolution of a large enough class of regular sheaves. A *Poincaré series* $\Pi_{\mathcal{F}}(t)$ of a sheaf $\mathcal{F}$ is the formal power series

$$\Pi_{\mathcal{F}}(t) \stackrel{def}{=} \sum_{i=0}^{\infty} \dim \Gamma(X, \mathcal{F}(i)) \cdot t^i \in \mathbb{Z}[[t]].$$

The Poincaré series $\Pi_S(t)$ of a variety S is the Poincaré series of its structural sheaf:

$$\Pi_S(t) \stackrel{def}{=} \Pi_{\mathcal{O}_S}(t).$$

In particular if $S = \operatorname{Proj} A$ for a graded algebra A, then $\Pi_S(t)$ is the usual Poincaré series of A.

Example 4.1.1. If S is the projective space, $S = \mathbb{P}^n_F$, then $\dim \Gamma(S, \mathcal{O}_S(i)) = \binom{n+i}{i}$, so

$$P_n(t) \stackrel{def}{=} \Pi_S(t) = \sum_{i=0}^{\infty} \binom{n+i}{i} \cdot t^i = (1-t)^{-n-1}.$$

Example 4.1.2. Let f be a homogeneous polynomial of degree k in homogeneous coordinates in $\mathbb{P}^{d+1}_F = \operatorname{Proj} B$, $B = F[x_0,\ x_1,\ \ldots,\ x_{d+1}]$, $A = B/(f)$, $S = \operatorname{Proj} A$ - a hypersurface $f = 0$ in $\mathbb{P}^{d+1}_F$. Since the exact sequence

$$0 \to B_n \xrightarrow{f\cdot} B_{n+k} \to A \to 0$$

splits for every n, the following equality holds:

$$\Pi_S(t) = P_{d+1}(t) - t^k P_{d+1}(t).$$

Thus $\Pi_S(t) = \dfrac{1-t^k}{(1-t)^{d+2}} = \dfrac{1+t+\ldots+t^{k-1}}{(1-t)^{d+1}}$.

Lemma 4.1.3. *For a projective quadric X of dimension d*

$$Q_d(t) \stackrel{def}{=} \Pi_X(t) = \frac{1+t}{(1-t)^{d+1}}.$$

□

Proposition 4.1.4. *If $0 \to \mathcal{F}' \to \mathcal{F} \to \mathcal{F}'' \to 0$ is an exact sequence of $\mathcal{O}_X$ - modules and either $\mathcal{F}'$, $\mathcal{F}''$ are regular or $\mathcal{F}$, $\mathcal{F}'(1)$ are regular, then*

$$\Pi_{\mathcal{F}}(t) = \Pi_{\mathcal{F}'}(t) + \Pi_{\mathcal{F}''}(t).$$

Proof. By [7], Sect. 8, Lemma 1.2 either $\mathcal{F}'$, $\mathcal{F}$, $\mathcal{F}''$ are regular or $\mathcal{F}'(1)$, $\mathcal{F}$, $\mathcal{F}''$ are regular. Hence each exact sequence of sheaves

$$0 \to \mathcal{F}'(i) \to \mathcal{F}(i) \to \mathcal{F}''(i) \to 0$$

induces an exact sequence of global sections. □

4.2 The generating function for the canonical resolution

The recursive method of finding a canonical resolution

$$\cdots \to \mathcal{O}_X(-p)^{k_p} \to \cdots \to \mathcal{O}_X(-1)^{k_1} \to {\mathcal{O}_X}^{k_0} \to \mathcal{F} \to 0$$

of a regular sheaf $\mathcal{F}$, described in 3.5 above, namely the identity 3.5.1, yields following identities for the generating function $G_{\mathcal{F}}(t) \overset{def}{=} \sum_{i=0}^{\infty} k_i t^i$:

$$\Pi_{\mathcal{F}}(t) = G_{\mathcal{F}}(-t) \cdot \Pi_X(t) \qquad \text{and} \qquad G_{\mathcal{F}}(t) = \frac{\Pi_{\mathcal{F}}(-t)}{Q_d(-t)}.$$

Example 4.2.1. The generating function for the canonical resolution of the sheaf $\mathcal{O}_X(1)$:

$$\Pi_{\mathcal{O}_X(1)}(t) = \frac{\Pi_{\mathcal{O}_X}(t) - 1}{t},$$

so

$$G_{\mathcal{O}_X(1)}(t) = \frac{\Pi_{\mathcal{O}_X}(-t)}{Q_d(-t)} = \frac{Q_d(-t) - 1}{-tQ_d(-t)} = \frac{\frac{1-t}{(1+t)^{d+1}} - 1}{-t\frac{1-t}{(1+t)^{d+1}}} = \frac{(1+t)^{d+1} - (1-t)}{t(1-t)}.$$

Example 4.2.2. For a linear section $H^l = (1 - [\mathcal{O}_X(-1)])^l$ of codimension l in X

$$G_{H^l} = (1+t)^l.$$

Example 4.2.3. Continue the notation of 3.1. Since X splits, it contains linear subvarieties $S_k = \operatorname{Proj} F[x_0, \ldots, x_k]$ given by the following equations:
a) in case of an even $d = 2m$:

$$y_0 = \ldots = y_m = x_{k+1} = \ldots = x_m = 0 \text{ for } k < m \text{ and}$$
$$y_0 = \ldots = y_m = 0 \text{ for } k = m;$$

b) in case of an odd $d = 2m+1$:

$$y_0 = \ldots = y_m = z_d = x_{k+1} = \ldots = x_m = 0 \text{ for } k < m \text{ and}$$
$$y_0 = \ldots = y_m = z_d = 0 \text{ for } k = m.$$

S_k is isomorphic to $\mathbb{P}^k_F$, in particular its structural sheaf $\mathcal{L}_k$ is regular. Therefore

$$G_{\mathcal{L}_k}(t) = \frac{P_k(-t)}{Q_d(-t)} = \frac{(1+t)^{-k-1}}{\frac{1-t}{(1+t)^{d+1}}} = \frac{(1+t)^{d-k}}{1-t}.$$

Lemma 4.2.4.

$$2G_{\mathcal{L}_k} - G_{\mathcal{L}_{k-1}} = (1+t)^{d-k}.$$

□

4.3 The generating function for a truncated canonical resolution

Truncating a generating function $G_{\mathcal{F}}$ one obtains a polynomial $T_{\mathcal{F}}$. For $l < d$ the canonical resolution for H^l is itself truncated:

$$T_{H^l} = (1+t)^l \qquad \text{for} \qquad l < d.$$

The sequence (c_i) of coefficients of the canonical resolution of the sheaf $\mathcal{L}_k$ stabilizes from the degree $d-k$ onwards:

$$G_{\mathcal{L}_k} = \frac{(1+t)^{d-k}}{1-t} = (1+t)^{d-k} \cdot \sum_{i=0}^{\infty} t^i = \sum_{i=0}^{\infty} c_i t^i$$

so

$$c_{d-k} = c_{d-k+1} = \ldots = 2^{d-k}.$$

Thus

$$T_{\mathcal{L}_k}(t) = \frac{(1-t)^{d-k} - 2t^d}{1-t}.$$

Proposition 4.3.1. *If, for fixed k, $\mathcal{L}_k$ is a structural sheaf of a linear subvariety S_k of dimension k in X, then in $K_0(X)$:*
a) in case of an odd $d = 2m+1$

$$[\mathcal{L}_k] = \sum_{i=0}^{d-1} \left(\sum_{p=0}^{i} \binom{d-k}{p} \right) (-1)^i [\mathcal{O}_X(-i)] + 2^{m-k}[\mathcal{V}];$$

b) in case of an even $d = 2m$ for a suitable integer a

$$[\mathcal{L}_k] = \sum_{i=0}^{d-1} \left(\sum_{p=0}^{i} \binom{d-k}{p} \right) (-1)^i [\mathcal{O}_X(-i)] + a[\mathcal{V}_0] + (2^{m-k} - a)[\mathcal{V}_1].$$

Proof. Substituting $t = -[\mathcal{O}_X(-1)]$ into the expansion for $T_{\mathcal{L}_k}(t)$ yields, depending on the parity of d, the expressions

$$[\mathcal{L}_k] = \sum_{i=0}^{d-1} \left(\sum_{p=0}^{i} \binom{d-k}{p} \right) (-1)^i [\mathcal{O}_X(-i)] + a[\mathcal{V}];$$

$$[\mathcal{L}_k] = \sum_{i=0}^{d-1} \left(\sum_{p=0}^{i} \binom{d-k}{p} \right) (-1)^i [\mathcal{O}_X(-i)] + a[\mathcal{V}_0] + b[\mathcal{V}_1].$$

for suitable integers a, b. Thus

$$0 = \operatorname{rank}[\mathcal{L}_k] = \begin{cases} T_{\mathcal{L}_k}(-1) + (-1)^d a \cdot 2^m = & \\ \quad = (-1)^d (2^m a - 2^{d-k-1}) & \text{for } d = 2m+1 \\ T_{\mathcal{L}_k}(-1) + (-1)^d (a+b) \cdot 2^{m-1} = & \\ \quad = (-1)^d (2^{m-1}(a+b) - 2^{d-k-1}) & \text{for } d = 2m. \end{cases}$$

□

4.4 The topological filtration

Now we shall find a basis of $K_0(X)$ which is convenient for computations. Since the quadric X is regular, $K_0'(X) = K_0(X)$ and one may transfer the topological filtration

$$\mathrm{F}^p\, K_0'(X) = \text{ subgroup generated by } \left\{ [\mathcal{F}] : \begin{array}{c} \text{the stalk } \mathcal{F}_x = 0 \text{ for all generic points} \\ x \text{ of subvarieties of codimension } < p \end{array} \right\}$$

of $K_0'(X)$ to $K_0(X)$. We omit the standard proof of following

Proposition 4.4.1. *For a split projective quadric X the Chow groups $A^p(X)$ are isomorphic to the corresponding factors of the topological filtration:*

$$A^p(X) \cong \mathrm{F}^p\, K_0(X)/\, \mathrm{F}^{p+1}\, K_0(X).$$

□

Continue the notation of 3.1. Recall the classical computation of the Chow ring of a split projective quadric.

Proposition 4.4.2. *For a split projective quadric X of dimension d*
a) in case of an even $d = 2m$

$$A^p(X) \cong \mathbb{Z} \textit{ for } p \neq m,\ 0 \leq p \leq 2m \textit{ and } A^m(X) \cong \mathbb{Z} \oplus \mathbb{Z};$$

b) in case of an odd $d = 2m+1$

$$A^p(X) \cong \mathbb{Z} \textit{ for all } p, 0 \leq p \leq 2m.$$

Explicit generators are given as follows:
Case $d = 2m$:
i) for $p > m$, a class of any linear subvariety of dimension $d - p$, e.g., $S_{d-p} : y_0 = \ldots = y_m = x_{d-p+1} = \ldots = x_m = 0$;
ii) for $p < m$, a class H^p of a linear section of codimension p;
iii) for $p = m$, $A^m(X)$ is generated by two classes of linear subvarieties $S_m' :\ x_0 = \ldots = x_m = 0$ and $S_m'' :\ y_0 = x_1 = \ldots = x_m = 0$; the classes in $A^m(X)$ remain unchanged if an even number of x_i, y_i are exchanged in these equations.
Case $d = 2m+1$:
i) for $p > m$, a class of any linear subvariety of dimension $d - p$, e.g. $S_{d-p} : y_0 = \ldots = y_m = z_{d+1} = x_{d-p+1} = \ldots = x_m = 0$;
ii) for $p \leq m$, a class H^p of a linear section of codimension p.

For a sketch of proof and references see [10], Thm. 13.3.

Now we can give an explicit description of the ring structure and the action of the involution ^ on $K_0(X)$. To do this denote $L_p = [\mathcal{L}_p]$ the class of the structural sheaf of the linear subvariety S_p of dimension p. Moreover, in case of an even $d = 2m$ denote by L_m' and L_m'' the class of the structural sheaf of S_m' and S_m'' respectively.

THEOREM 4.1. *Let X be a split projective quadric of dimension d. Then*

i) *in case of an odd $d = 2m+1$ classes $1,\ H,\ H^2,\ \dots,\ H^m,\ L_m,\ \dots,\ L_0$ form a basis of the free Abelian group $K_0(X)$;*

ii) *in case of an even $d = 2m$ classes $1,\ H,\ H^2,\ \dots,\ H^{m-1},\ L'_m, L''_m,\ L_{m-1},\ \dots,\ L_0$ form a basis of the free Abelian group $K_0(X)$;*

iii) *in case of an even $d = 2m$ classes may be chosen as follows:*

$$L'_m = \sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{m}{p}\right)(-1)^i[\mathcal{O}_X(-i)] + [\mathcal{V}_0],$$

$$L''_m = \sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{m}{p}\right)(-1)^i[\mathcal{O}_X(-i)] + [\mathcal{V}_1]$$

and for dimensions $k < m$

$$L_k = \sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{d-k}{p}\right)(-1)^i[\mathcal{O}_X(-i)] + 2^{m-k-1}([\mathcal{V}_0] + [\mathcal{V}_1]);$$

iv) *if $d = 2m$, then $H^m = L'_m + L''_m - L_{m-1}$;*

v) $H \cdot L_p = L_{p-1} \quad , \quad H \cdot L'_m = H \cdot L''_m = L_{m-1}$;

vi) $H^{d-k} = 2L_k - L_{k-1}$ *for* $k \le \dfrac{d-1}{2}$, $H^d = 2L_0$, $H^{d+1} = 0$;

vii) $L_p \cdot L_q = L_p \cdot L'_m = L_p \cdot L''_m = 0$;

viii) *if $d = 2m$ and m is even, then ${L'_m}^2 = {L''_m}^2 = L_0$, $L'_m \cdot L''_m = 0$, if $d = 2m$ and m is odd, then ${L'_m}^2 = {L''_m}^2 = 0$, $L'_m \cdot L''_m = L_0$.*

Proof. First of all note that the classes H^k, L_k for $k \le \frac{d-1}{2}$ and the pair $\{L'_m,\ L''_m\}$ are determined uniquely by the conditions of irreducibility of the underlying subvariety and to form a basis of some appropriate $A^p(X)$. In fact, by Proposition 4.4.1 this is clear for $\mathrm{F}^d K_0(X) \cong A^d(X)$. Thus, the general case results by induction. Statements i) and ii) follow from Proposition 4.4.1 and 4.4.2. To verify iii), note that the reflection ρ_{v_1} fixes $v_0,\ v_2,\ \dots,\ v_{d+1}$ and changes v_1 into the opposite (3.2 above). Thus, this reflection induces an automorphism of the symmetric algebra $S(\mathcal{V}^\wedge)$, which interchanges x_0 with y_0 and fixes other coordinates and q. Therefore it induces an automorphism of $S(\mathcal{V}^\wedge)/(q)$, $X = \operatorname{Proj} S(\mathcal{V}^\wedge)/(q)$, a semilinear automorphism of $\mathcal{O}_X(n)$ for all n, and an automorphism of $K_0(X)$. By Lemma 3.2.1 ii), the reflection ρ_{v_1} interchanges the P_i's with each other. So the induced automorphism of $\mathcal{U}$ interchanges direct summands $\mathcal{U}' = U \otimes P_0$ and $\mathcal{U}'' = U \otimes P_1$ of $\mathcal{U}$ and their indecomposable components $\mathcal{V}_0$, $\mathcal{V}_1$. Therefore, the induced automorphism of $K_0(X)$ fixes the basic elements $[\mathcal{O}_X]$, $[\mathcal{O}_X(-1)]$, $\dots$, $[\mathcal{O}_X(1-d)]$ and interchanges $[\mathcal{V}_0]$ with $[\mathcal{V}_1]$. By the uniqueness statement this automorphism fixes $L_0,\ \dots,\ L_{m-1}$. The explicit description given in Proposition 4.4.2 ii) shows that this automorphism interchanges L'_m with L''_m. Hence, by the explicit formula of Proposition 4.3.1 ii) for $k<m$

$$L_k=[\mathcal{L}_k]=\sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{d-k}{p}\right)(-1)^i[\mathcal{O}_X(-i)]+a[\mathcal{V}_0]+(2^{m-k}-a)[\mathcal{V}_1]$$

the integer a must be equal to 2^{m-k-1}. This same argument for $k=m$ yields

$$L'_m=\sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{d-k}{p}\right)(-1)^i[\mathcal{O}_X(-i)]+a[\mathcal{V}_0]+(1-a)[\mathcal{V}_1]$$

and

$$L''_m=\sum_{i=0}^{d-1}\left(\sum_{p=0}^{i}\binom{d-k}{p}\right)(-1)^i[\mathcal{O}_X(-i)]+(1-a)[\mathcal{V}_0]+a[\mathcal{V}_1].$$

Since the statement ii) of the theorem holds, the integer a must be 0 or 1 (this follows from the regularity of the structural sheaves of S'_m and S''_m , too). Statements iv) - vii) are obvious consequences of the uniqueness and the explicit equations of Proposition 4.4.2. For to prove *viii)* assume, without loss of generality, that L'_m is the class of S'_m and L''_m is the class of S''_m. Consider the class L_m of the subvariety $S_m:\ y_0=\ldots=y_m=0$. In case of an even m the classes L''_m and L_m coincide, and S_m meets S'_m transversally at the empty set of points, so $L'_m\cdot L''_m=0$. Moreover, S_m meets S''_m transversally at the rational point S_0, so ${L''_m}^2=L_0$. Analogously, ${L'_m}^2=L_0$.
In case of an odd m $L'_m=L_m$, so ${L'_m}^2={L''_m}^2=0$, $L'_m\cdot L''_m=L_0$. ◻

Theorem 4.2. *For a split projective quadric X of dimension d, the involution $^\wedge$ acts as follows:*

i) $\quad {L_k}^\wedge=(-1)^{d-k}\cdot\sum_{i=0}^{k}\binom{d-k-2+i}{i}L_{k-i}$ *for* $k\le\frac{d-1}{2}$;

ii) *in case of an even $d=2m$:*

$${L'_m}^\wedge=(-1)^m\cdot\left(L'_m+\sum_{i=1}^{m}\binom{m-2+i}{i}L_{m-i}\right),$$

$${L''_m}^\wedge=(-1)^m\cdot\left(L''_m+\sum_{i=1}^{m}\binom{m-2+i}{i}L_{m-i}\right),$$

$$\begin{aligned}H^{k\wedge}=&(-1)^k\cdot\left(\sum_{j=k}^{m-1}\binom{j-1}{k-1}H^j+\binom{m-1}{k-1}(L'_m+L''_m)\right)\\&+(-1)^k\cdot\left(\sum_{j=m+1}^{d}\left(\binom{j-1}{k-1}-\binom{j-1}{k-2}\right)\cdot L_{d-j}\right);\end{aligned}$$

iii) in case of an odd $d = 2m + 1$

$$H^{k\wedge} = (-1)^k \cdot \left(\sum_{j=k}^{m-1} \binom{j-1}{k-1} H^j + 2\binom{m-1}{k-1} L_m \right) + (-1)^k \left(\sum_{j=m+1}^{d} \left(\binom{j-1}{k-1} - \binom{j-1}{k-2} \right) \cdot L_{d-j} \right).$$

Proof. i) Since $H^{d-k} = 2L_k - L_{k-1} = 2L_k - H \cdot L_k$ by the Theorem 4.1 iv), vi),

$$L_k = \frac{H^{d-k}}{2-H} \quad \text{and} \quad H^\wedge = \frac{-H}{1-H}$$

by Lemma 2.1.2 iii),

$$L_k{}^\wedge = \frac{\left(\frac{-H}{1-H}\right)^{d-k}}{2 + \frac{H}{1-H}} = \frac{(-H)^{d-k}}{(2-H)(1-H)^{d-k-1}} = (-1)^{d-k} L_k \frac{1}{(1-H)^{d-k-1}}$$

$$= (-1)^{d-k} L_k \sum_{i=0}^{d} \binom{d-k-2+i}{i} H^i = (-1)^{d-k} \sum_{i=0}^{d} \binom{d-k-2+i}{i} L_{k-i}.$$

ii) To obtain the formula for $H^{k\wedge}$ substitute $H^{d-k} = 2L_k - L_{k-1}$ and $H^m = L'_m + L''_m - L_{m-1}$ into the formula of Lemma 2.1.2 iv). Analogously one proves iii). In case $d = 2m$

$$L'_m + L''_m = H^m + L_{m-1} = H^m + \frac{H^{m+1}}{2-H} = \frac{2H^m}{2-H},$$

so

$$(L'_m + L''_m)^\wedge = 2\left(\frac{-H}{1-H}\right)^m \frac{1}{2 + \frac{H}{1-H}} = (-1)^m \cdot 2 \cdot \frac{H^m}{2-H} \cdot \frac{1}{(1-H)^{m-1}}$$

$$= (-1)^m \cdot (L'_m + L''_m) \frac{1}{(1-H)^{m-1}} = (-1)^m \cdot (L'_m + L''_m) \sum_{i=0}^{d} \binom{m-2+i}{i} H^i,$$

and thus

$$(L'_m + L''_m)^\wedge = (-1)^m \cdot \left(L'_m + L''_m + 2\sum_{i=1}^{m} \binom{m-2+i}{i} L_{m-i} \right).$$

On the other hand, by Theorem 4.1 iii)

$$L'_m - L''_m = [\mathcal{V}_0] - [\mathcal{V}_1]$$

and by Corollary 3.4.9 iii)

$$(L'_m - L''_m)^\wedge = (-1)^m \cdot ([\mathcal{V}_0] - [\mathcal{V}_1]) = (-1)^m \cdot (L'_m - L''_m).$$

The formula for $L'_m{}^\wedge$ and $L''_m{}^\wedge$ follows directly, since we know their sum and difference. $\square$

Consider the matrix A of the involution $^\wedge$ in the free Abelian group $K_0(X)$ with respect to the basis given in Theorem 4.1 i), ii). We shall write it in a slightly unusual way:

$$A = [a_{i,j}] \qquad , \qquad 0 \le i, j \le 2m+1.$$

In case of an even $d = 2m$ we regard A as a block matrix B, arranging two central rows and two central columns into separate blocks:

$$b_{m,m} = \left[\begin{array}{cc} a_{m,m} & a_{m,m+1} \\ a_{m+1,m} & a_{m+1,m+1} \end{array} \right] \in \mathrm{Hom}(\mathbb{Z}^2, \mathbb{Z}^2),$$
$$b_{m,i} = \left[\begin{array}{c} a_{m,i} \\ a_{m+1,i} \end{array} \right] \in \mathrm{Hom}(\mathbb{Z}, \mathbb{Z}^2),$$
$$b_{i,m} = [a_{i,m}\ a_{i,m+1}] \in \mathrm{Hom}(\mathbb{Z}^2, \mathbb{Z}) \quad \text{for } i \ne m,$$
$$b_{i,j} = \begin{cases} a_{i,j} & \text{for } i, j < m \\ a_{i,j+1} & \text{for } i < m, j > m \\ a_{i+1,j} & \text{for } i > m, j < m \\ a_{i+1,j+1} & \text{for } i, j > m. \end{cases}$$

As one may expect in view of Proposition 1.3.2 iv), the matrix A is triangular in the odd dimensional case and the matrix B is triangular in the even dimensional case. We summarize the most important results of Theorem 4.2 as follows:

COROLLARY 4.4.3. *a) In case of an odd $d = 2m+1$ the matrix A is triangular with*

$$a_{i,i} = (-1)^i \quad \textit{for } i = 0, 1, \ldots, 2m+1,$$
$$a_{i,0} = 0 \quad \textit{for } i > 0,$$
$$a_{i+1,i} = \begin{cases} (-1)^i i & \textit{for } i = 0, 1, \ldots, m-1 \\ (-1)^m \cdot 2m & \textit{for } i = m \\ (-1)^i (i-1) & \textit{for } i = m+1, \ldots, 2m. \end{cases}$$

b) In case of an even $d = 2m$ the matrix B is block triangular with

$$b_{i,i} = (-1)^i \textit{for } i \ne m,$$
$$b_{m,m} = (-1)^m \left[\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right]$$
$$b_{i,0} = 0 \textit{for } i > 0,$$
$$b_{i+1,i} = \begin{cases} (-1)^i i & \textit{for } i = 0, 1, \ldots, m-2 \\ (-1)^i (i-1) & \textit{for } i = m+1, \ldots, 2m, \end{cases}$$

$$b_{m,m-1} = (-1)^{m-1}(m-1) \left[\begin{array}{c} 1 \\ 1 \end{array} \right],$$
$$b_{m+1,m} = (-1)^m (m-1)\,[1\ 1]\,,$$
$$b_{2m,m} = (-1)^m \binom{2m-2}{m}\,[1\ 1]\,.$$

□

Theorem 4.3. *Let X be a split projective quadric of dimension d.*
a) If $d = 2m+1$ is odd, then

$$EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X],$$

$$E^- X = \begin{cases} \mathbb{Z}/2\mathbb{Z} \cdot [L_m] & \text{for even } m \\ \mathbb{Z}/2\mathbb{Z} \cdot [H^m] & \text{for odd } m; \end{cases}$$

b) If $d = 2m$ is even, then

$$EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [L_0],$$

$$E^- X = \begin{cases} 0 & \text{for even } m \\ \mathbb{Z}/2\mathbb{Z} \cdot [L'_m] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [L''_m] & \text{for odd } m. \end{cases}$$

Proof. Consider the complex 1.3.1:

$$\cdots \to K_0(X) \xrightarrow{1+^\wedge} K_0 X \xrightarrow{1-^\wedge} K_0(X) \xrightarrow{1+^\wedge} K_0(X) \to \cdots$$

with the topological filtration

$$K_0(X) = \mathrm{F}^0\, K_0(X) \supset \mathrm{F}^1\, K_0(X) \supset \cdots \supset \mathrm{F}^d\, K_0(X) \supset \mathrm{F}^{d+1}\, K_0(X) = 0,$$

and the corresponding spectral sequence

$$\mathrm{E}_1^{p,q} = \operatorname{Ker}(1 - (-1)^{p+q} \cdot {}^\wedge) / \operatorname{Im}(1 + (-1)^{p+q} \cdot {}^\wedge) \Longrightarrow E^{(-1)^{p+q}} X$$

where $E^1 X = EX$, $E^{-1} X = E^- X$. The E_1 - term has period 2 with respect to q.
a) In case of an odd $d = 2m+1$ the term E_1 looks like

$$\begin{array}{ccccccccccc} \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots \\ {\scriptstyle q=1} & 0 & & 0 & & \cdots & & 0 & & 0 \\ {\scriptstyle q=0} & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{\partial_0} & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{\partial_1} & \cdots & \to & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{\partial_{d-1}} & \mathbb{Z}/2\mathbb{Z} \end{array}$$

The differential ∂_i is induced by the multiplication by the entry $a_{i+1,i}$ of the matrix A of ${}^\wedge$. Thus, for each even q, we have complex $\mathrm{E}_1^{\cdot,q}$:

$$\mathbb{Z}/2\mathbb{Z} \xrightarrow{0\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{1\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{2\cdot} \cdots$$
$$\xrightarrow{(m-1)\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{2m\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{m\cdot} \cdots$$
$$\to \mathbb{Z}/2\mathbb{Z} \xrightarrow{(2m-1)\cdot} \mathbb{Z}/2\mathbb{Z}$$

Therefore for even m we have $\mathrm{E}_2\, 0, q = E_2^{m+1,q} = \mathbb{Z}/2\mathbb{Z}$ and $E_2^{i,q} = 0$ for other values of i. Since the left (the zeroth) column of A has zero entries except $a_{0,0} = 1$, all the differentials starting from $\mathrm{E}_r^{0,q}$ are trivial. So $EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X]$, $E^- X = \mathbb{Z}/2\mathbb{Z} \cdot [L_m]$. Analogously, for an odd m, we have $E_2^{0,q} = E_2^{m,q} = \mathbb{Z}/2\mathbb{Z}$ and $E_2^{i,q} = 0$ for other values of i, so $EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X]$, $E^- X = \mathbb{Z}/2\mathbb{Z} \cdot [H^m]$.

b) In case of an even $d = 2m$, the term E_1 looks like

$$\begin{array}{ccccccccccccc}
\vdots & & \vdots & \cdots & \vdots & & \vdots & & \vdots & & \vdots & \cdots & \vdots & & \vdots \\
0 & & & \cdots & 0 & & & & 0 & & 0 & \cdots & & & 0 \\
\mathbb{Z}/2\mathbb{Z} & \xrightarrow{\partial_0} & \cdots & & \mathbb{Z}/2\mathbb{Z} & \xrightarrow{\partial_{m-1}} & & & (\mathbb{Z}/2\mathbb{Z})^2 & \xrightarrow{\partial_m} & \mathbb{Z}/2\mathbb{Z} & \cdots & & \xrightarrow{\partial_{d-1}} & \mathbb{Z}/2\mathbb{Z}.
\end{array}$$

The differential ∂_i is induced by the corresponding block of the matrix B of $^\wedge$. For each even q we have a complex $E_1^{\cdot,q}$:

$$\mathbb{Z}/2\mathbb{Z} \xrightarrow{0\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{1\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{2\cdot} \cdots$$
$$\xrightarrow{(m-2)} \mathbb{Z}/2\mathbb{Z} \xrightarrow{(m-1)\begin{bmatrix}1\\1\end{bmatrix}} (\mathbb{Z}/2\mathbb{Z})^2 \xrightarrow{(m-1)[1\ \ 1]\cdot} \mathbb{Z}/2\mathbb{Z} \xrightarrow{m\cdot} \cdots$$
$$\to \mathbb{Z}/2\mathbb{Z} \xrightarrow{(2m-2)\cdot} \mathbb{Z}/2\mathbb{Z}$$

Thus for even m and even q only $\mathrm{E}_2^{0,q} = \mathrm{E}_2^{2m,q} = \mathbb{Z}/2\mathbb{Z}$ are nonzero. By the dimension argument the sequence degenerates from E_2 onwards. Hence

$$EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [L_0],\ E - X = 0 \text{ for even } m.$$

For odd m and even q only $\mathrm{E}_2^{0,q} = \mathrm{E}_2^{2m,q} = \mathbb{Z}/2\mathbb{Z}$ and $E_2^{m,q} = (\mathbb{Z}/2\mathbb{Z})^2$ are nonzero. There is no nonzero differential starting from $\mathrm{E}_r^{0,q}$ since the entries of the left (zeroth) column of B are 0 except $b_{0,0} = 1$. All the differentials but $\mathrm{E}_m^{m,q} \to \mathrm{E}_m^{2m,q-m+1}$ must be zero. This exceptional one is zero too, since it is induced by $b_{2m,m} = (-1)^m \cdot \binom{2m-2}{m}[1\ \ 1]$, and $\binom{2m-2}{m}$ is even for odd m. Therefore the spectral sequence degenerates, and finally

$$EX = \mathbb{Z}/2\mathbb{Z} \cdot [\mathcal{O}_X] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [L_0],\ E^- X = \mathbb{Z}/2\mathbb{Z} \cdot [L'_m] \oplus \mathbb{Z}/2\mathbb{Z} \cdot [L''_m] \text{ for odd } m.$$

The theorem is proved. ◻

5 Non-extended Witt classes on certain split projective quadrics

We shall show here that if the dimension d of a split projective quadric X is even and greater than two, then the invariant $e^0 : W(X) \to EX \cong (\mathbb{Z}/2\mathbb{Z})^2$ is surjective.

5.1 For an arbitrary locally free coherent sheaf $\mathcal{M}$ the sheaf $\mathcal{E} = \mathcal{E}nd_{\mathcal{O}_X}(\mathcal{M}) = \mathcal{M} \otimes \mathcal{M}^\wedge$ is self-dual and supports a canonical symmetric bilinear form θ, which reduces to the trace of a product on stalks:

$$\theta(\alpha)(\beta) = \mathrm{tr}(\alpha \cdot \beta) \ \text{ for } \alpha, \beta \in \mathcal{E}_x, x \in X$$

or if $\mu : \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{E} \to E$ is the multiplication map, then $\theta : E \to \mathcal{E}^\wedge$ is adjoint of $\mathrm{tr} \circ \mu : \mathcal{E} \otimes_{\mathcal{O}_X} \mathcal{E} \to \mathcal{O}_X$.

Theorem 5.1. *If X is a split projective quadric of dimension $d = 2m$, $m > 1$, then, for an indecomposable component $\mathcal{V}_0$ of the Swan sheaf $\mathcal{U}$,*

$$e^0(\mathcal{E}nd_{\mathcal{O}_X}(\mathcal{V}_0), \theta) = [L_0].$$

Thus $(\mathcal{E}nd_{\mathcal{O}_X}(\mathcal{V}_0), \theta)$ represents a non-extended Witt class in $W(X)$.

Proof. The case $m = 0$ is special, so assume $m > 0$. We shall compute the class of $[\mathcal{V}_0] \cdot [\mathcal{V}_0]^\wedge = [\mathcal{V}_0(d)] \cdot [\mathcal{V}_0(d)]^\wedge$ in EX. We know from Proposition 3.4.8 b) iv) that, for $d = 2m$,

$$[\mathcal{V}_0(d)] + [\mathcal{V}_1(d-1)] = 2^m.$$

On the other hand by Corollary 3.4.9 ii) and Theorem 4.1 iii)

$$[\mathcal{V}_0(d-1)] - [\mathcal{V}_1(d-1)] = [\mathcal{V}_0] - [\mathcal{V}_1] = L'_m - L''_m.$$

Thus

$$[\mathcal{V}_0(d)](1 + [\mathcal{O}_X(-1)]) = [\mathcal{V}_0(d)] + [\mathcal{V}_1(d-1)] = 2^m + L'_m - L''_m.$$

or

$$[\mathcal{V}_0(d)](2 + H) = 2^m + L'_m - L''_m.$$

The rules of multiplication in $K_0(X)$, given in Theorem 4.1 and Lemma 2.1.2 yield that multiplying both sides of this equality by

$$\begin{aligned}
\sum_{i=0}^{d} 2^{d-i}H^i &= \sum_{i=0}^{m-1} 2^{d-i}H^i + 2^m \cdot (L'_m + L''_m - L_{m-1}) + \sum_{j=1}^{m} 2^{m-j}H^{m+j} \\
&= \sum_{i=0}^{m-1} 2^{d-i}H^i + 2^m \cdot (L'_m + L''_m - L_{m-1}) + \sum_{j=1}^{m} 2^{m-j}(2L_{m-j} - L_{m-j-1}) \\
&= \sum_{i=0}^{m-1} 2^{d-i}H^i + 2^m \cdot (L'_m + L''_m)
\end{aligned}$$

we obtain

$$\begin{aligned}
2^{d+1}[\mathcal{V}_0(d)] &= [\mathcal{V}_0(d)](2^{d+1} + H^{d+1}) \\
&= \left(\sum_{i=0}^{m-1} 2^{d-i}H^i + 2^m \cdot (L'_m + L''_m)\right) \cdot (L'_m + L''_m - L_{m-1}) \\
&= \left(2^{m+1}\sum_{i=0}^{m-1} 2^{m-i-1}H^i + 2^m \cdot (L'_m + L''_m)\right) \cdot (L'_m + L''_m - L_{m-1}) \\
&= 2^{d+1}\sum_{i=0}^{m-1} 2^{m-i-1}H^i + 2^d \cdot (L'_m + L''_m) + 2^d \cdot (L'_m - L''_m) \\
&= 2^{d+1} \cdot \sum_{i=0}^{m-1} 2^{m-i-1}H^i + 2^{d+1} \cdot L'_m.
\end{aligned}$$

Since $K_0(X)$ is torsion free, $[\mathcal{V}_0(d)] = \sum_{i=0}^{m-1} 2^{m-i-1}H^i + L'_m$. Thus

$$[\mathcal{V}_0(d)]^\wedge = \sum_{i=0}^{m-1} 2^{m-i-1}H^{i\wedge} + {L'_m}^\wedge.$$

Note that $(\alpha+\beta)\cdot(\alpha+\beta)^\wedge \equiv \alpha\cdot\alpha^\wedge + \beta\cdot\beta^\wedge \mod \operatorname{Im}(1+{}^\wedge)$, since $\alpha^\wedge\cdot\beta+\alpha\cdot\beta^\wedge$ is a member of $\operatorname{Im}(1+{}^\wedge)$. Also $2\alpha\cdot\alpha^\wedge \equiv 0 \mod \operatorname{Im}(1+{}^\wedge)$. Therefore

$$\begin{aligned}[\mathcal{E}] &= [\mathcal{V}_0]\otimes[\mathcal{V}_0{}^\wedge] = [\mathcal{V}_0(d)]\cdot[\mathcal{V}_0(d)]^\wedge \\ &\equiv \sum_{i=0}^{m-1} 2^{2(m-i-1)} H^i H^{i\wedge} + L'_m L'_m{}^\wedge \mod \operatorname{Im}(1+{}^\wedge).\end{aligned}$$

If $m=1$, then the first summand equals 1 while the second is 0. For $m>1$

$$\begin{aligned}[\mathcal{E}] &\equiv \sum_{i=0}^{m-1} 2^{d-2i-2} H^i H^{i\wedge} + L'_m L'_m{}^\wedge \\ &= \sum_{i=0}^{m-2} 2^{d-2i-2} H^i H^{i\wedge} + H^{m-1}H^{m-1\wedge} + L'_m L'_m{}^\wedge \\ &\equiv H^{m-1}H^{m-1\wedge} + L'_m L'_m{}^\wedge && \text{since } 2\alpha\alpha^\wedge \equiv 0 \\ &\equiv H^{m-1}H^{m-1}\left(1\binom{m-1}{1}H + \binom{m}{2}H^2\right) + L'_m L'_m{}^\wedge && \text{by Lemma 2.1.2;} \\ &\equiv H^{d-2} + (m-1)H^{d-1} + \frac{m(m-1)}{2}H^d + L'_m L'_m{}^\wedge \\ &\equiv 2L_2 - L_1 + 2(m-1)L_1 - (m-1)L_0 + m(m-1)L_0 \\ &\qquad + L'_m L'_m{}^\wedge && \text{by Theorem 4.1 vi);} \\ &\equiv 2L_2 + (d-3)L_1 - (m-1)^2 L_0 + L'_m L'_m{}^\wedge \\ &\equiv L_2 + L_2{}^\wedge - (m-2)L_0 \\ &\qquad + (-1)^m L'_m(L'_m + \text{terms of higher codim}) && \text{by Theorem 4.2 i);} \\ &\equiv \begin{cases} (2-m)L_0 + L_0 \text{ for even } m \\ (2-m)L_0 \text{ for odd } m \end{cases} && \text{by Theorem 4.1 viii);} \\ &\equiv L_0 \mod \operatorname{Im}(1+{}^\wedge).\end{aligned}$$

Anyway, $e^0((\mathcal{V}_0),\theta) = [L_0]$ for $m>1$. □

5.2 In the particular case $d=4$ there exists another symmetric bilinear form ϑ on $\mathcal{E} = \mathcal{V}_0 \otimes_{\mathcal{O}_X} \mathcal{V}_0{}^\wedge$: the tensor product of exterior multiplications

$$\mathcal{V}_0\otimes_{\mathcal{O}_X}\mathcal{V}_0 \to \bigwedge^2 \mathcal{V}_0 \cong \mathcal{O}_X(-7) \text{ and } \mathcal{V}_0{}^\wedge\otimes_{\mathcal{O}_X}\mathcal{V}_0{}^\wedge \to \bigwedge^2 \mathcal{V}_0{}^\wedge \cong \mathcal{O}_X(7).$$

On the stalks the associated quadratic form is the determinant map. Since the value of e^0 depends only on supporting bundle, $(\mathcal{E},\vartheta)$ is non-extended as well as $(\mathcal{E},\theta)$.

The symmetric bilinear space $(\mathcal{E},\vartheta)$ has the following interesting property: it is not metabolic (since it has a nontrivial e^0) but is hyperbolic on stalks, i.e., *locally hyperbolic*. In fact, any stalk $\mathcal{V}_{0,x}$ at $x\in X$ is a free rank two $\mathcal{O}_{X,x}$ - module, so any stalk of $(\mathcal{E},\vartheta)$ is $(M_2(\mathcal{O}_{X,x}), det)$, which is hyperbolic. Thus, there is no local invariant to detect the symmetric bilinear space $(\mathcal{E},\vartheta)$ and such a global invariant as e^0 is useful. If -1 is a sum of two squares in F, then $(\mathcal{E},\theta)$ is locally hyperbolic, too.

Note that the case $d = 4$ is of particular interest, since the split four-dimensional quadric X is the smallest non-trivial Graßmann variety $G_2(4)$. Thus on general Graßmann varieties there may exist non-extended Witt classes contrary to the case of projective spaces, i.e., Graßmann varieties $G_1(n)$.

References

[1] J. Kr. Arason, *Der Wittring projektiver Räume*, Math. Ann. 253 (1980), 205–212.

[2] F. Fernández-Carmena, *On the injectivity of the map of the Witt group of a scheme into the Witt group of its function field*, Math. Ann. 277 (1987), 453–468. Corr: Math. Ann. 280 (1988), 697–698.

[3] F. Fernández-Carmena, *The Witt group of a smooth complex surface*, Math. Ann. 277 (1987), 469–481.

[4] M. M. Kapranov, *On the derived categories of coherent sheaves on some homogeneous spaces*, Invent. Math. 92 (1988), 479–508.

[5] M. Knebusch, *Symmetric bilinear forms over algebraic varieties*, Conference on quadratic forms 1976, ed. G. Orzech, Queens papers in pure and applied math. 46(1977), 103–283.

[6] T. Y. Lam, *Serres conjecture*, Lect. Notes in Math. 635, Springer, Berlin 1978.

[7] D. Quillen, *Higher algebraic K-theory I*, Lect. Notes in Math. 341, Springer, Berlin 1973.

[8] A. Ranicki, *Algebraic L-theory*, Proc. London Math. Soc. III ser. 27 (1973), 101–158.

[9] A. Ranicki, *The algebraic theory of surgery*, Proc. London Math. Soc. III ser. 40 (1980), 87–192.

[10] R. G. Swan, *Vector bundles, projective modules and the K-theory of spheres*, ed. W. Bowder, Algebraic topology and algebraic K-theory, Princeton 1987 Ann. Math. Studies 113 (1985), 432–522.

[11] R. G. Swan, *K-theory of quadric hypersurfaces*, Ann. Math. Studies 122 (1985), 113–153.

Marek Szyjewski
Instytut Matematyki
Uniwersytet Śląski
PL 40007 Katowice, ul. Bankowa 14,
Poland
szyjewski@gate.math.us.edu.pl

Bifurcation from Relative Equilibria of Noncompact Group Actions: Skew Products, Meanders, and Drifts

Bernold Fiedler, Björn Sandstede,
Arnd Scheel and Claudia Wulff

Received: October 4, 1996

Communicated by Alfred K. Louis

Abstract. We consider a finite-dimensional, typically noncompact Riemannian manifold M with a differentiable proper action of a possibly noncompact Lie group G. We describe G-equivariant flows in a tubular neighborhood U of a relative equilibrium $G \cdot u_0, u_0 \in M$, with compact isotropy H of u_0, by a skew product flow $\dot{g} = g\mathbf{a}(v)$, $\dot{v} = \varphi(v)$. Here $g \in G$, $\mathbf{a} \in \mathrm{alg}(G)$. The vector v is in a linear slice V to the group action. The induced local flow on $G \times V$ is equivariant under the action of $(g_0, h) \in G \times H$ on $(g, v) \in G \times V$, given by $(g_0, h)(g, v) = (g_0 g h^{-1}, hv)$. The original flow on U is equivalent to the induced flow on $\{id\} \times H$-orbits in $G \times V$.

Applications to relative equivariant Hopf bifurcation in V are presented, clarifying phenomena like periodicity, meandering, and drifting. Specific illustrations involving Euclidean groups G are meandering spirals, in the plane, and drifting twisted scroll rings, in three-dimensional Belousov-Zhabotinsky media.

1991 Mathematics Subject Classification: 58F35, 57S20, 55R91

1 Introduction

Going beyond rigidly rotating spirals, meandering and drifting spiral wave patterns have been observed in Belousov-Zhabotinsky media [UNUM93], [JSW89], [BE93] and in low pressure CO-oxidation on platinum monocrystals [NvORE93]. Mathematically speaking, the wave patterns are described by concentration vectors $u = u(t, x)$ depending on time t and location $x \in \mathbb{R}^2$. The partial differential equations, which model the dynamics of the solutions $u(t, x)$, are equivariant with respect to the standard affine action of the planar Euclidean group $E(2)$.

The *Euclidean group* $E(N), N = 2, 3, \ldots,$ is a semidirect product $E(N) = O(N) \times \mathbb{R}^N$ of the orthogonal group $O(N)$ with the Abelian translation group $\mathbb{R}^N$. The composition for $(R, S), (R', S') \in O(N) \times \mathbb{R}^N$ is defined by

$$(R, S) \circ (R', S') := (RR', S + RS'); \tag{1.1}$$

this rule is compatible with the standard affine representation

$$(R, S)x := Rx + S \tag{1.2}$$

on $x \in \mathbb{R}^N$. *Equivariance* of our dynamical system means that $u(t, \cdot)$ is a solution if, and only if, $(R, S)u(t, \cdot)$ is a solution for any (R, S). Here the linear representation of (R, S) in the state space X of solution x-profiles $u(t, \cdot)$ is given by

$$((R, S)u(t, \cdot))\,(x) := u(t, (R, S)^{-1}x). \tag{1.3}$$

The inverse $(R, S)^{-1}x$ is, of course, given explicitly by

$$(R, S)^{-1} = (R^{-1}, -R^{-1}S). \tag{1.4}$$

A *spiral wave* $u(t, \cdot)$ is a special time periodic solution, for which the time orbit is contained in a single group orbit. After a fixed shift of x-coordinates, it can be written as

$$u(t, \cdot) = (R(t), 0)u(0, \cdot). \tag{1.5}$$

The rotations $R(t) \in SO(N)$ are given as a periodic one-parameter subgroup

$$R(t) = \exp(\mathbf{r}_0 t) \tag{1.6}$$

generated by $\mathbf{r}_0$ in the Lie algebra $so(N)$ of anti-symmetric matrices. In the terminology of [Fie88], non-stationary spiral waves are called *rotating waves*; see also section 3. The term "spiral" arises from the above applied context, where the concentration patterns largely follow Archimedean spirals. Quite analogously, a *meandering wave* $u(t, \cdot)$ is a special solution of the form

$$u(t, \cdot) = (R(t), S(t))\, v(t, \cdot), \tag{1.7}$$

where this time $v(t, \cdot)$ is a nonstationary time periodic solution and the shifts $S(t)$ remain bounded. If the shifts $S(t)$ are unbounded, we call the solution $u(t, \cdot)$ *drifting.*

Numerically, meandering and drifting one-armed spirals have been observed in planar ($N = 2$) models by Barkley [Bar94]. Emphasizing the lack of a theoretical framework, based on Euclidean $E(2)$ equivariance, he also presented an ad-hoc heuristic ODE model exhibiting meandering and drifting solutions.

The first mathematically rigorous analysis of these phenomena has recently been achieved by Wulff, see [Wul96]. Her result is based on a careful Lyapunov-Schmidt reduction in a scale of Banach spaces. This resolves the difficulties of non-differentiability and, in some cases, *non-continuity* of the group action (1.3) on the infinite-dimensional Banach space $u(t, \cdot) \in X$. For technically related earlier results, restricted to compact group actions, see [Ren82] and [Ran82]. It has recently been shown, for the first time, that a center manifold reduction to a finite-dimensional globally group-invariant and locally time-invariant C^{k+1} manifold $M \subseteq X$ can also be achieved in an $E(2)$-equivariant context, if the nonlinearity of the differential equation governing the dynamics of the spiral waves is smooth; see [SSW96a], [SSW96b]. The reduction is based on the assumption that the linearization at the spiral wave does not exhibit continuous spectrum near the imaginary axis. Most notably, the group action becomes differentiable on M, albeit its possible noncontinuity on X. Communicated by one of the present authors, this idea is already being used successfully to investigate meandering of multi-armed spirals [GLM96]. The method of *center bundles,* there, is similar in spirit to a previous approach to bifurcation from relative equilibria of compact group actions [Kru90].

In the present paper we give an alternative, new description of the flow near relative equilibria inside a finite-dimensional Riemannian C^{k+1}-manifold M, typically noncompact, with a C^{k+1}-smooth action of a possibly noncompact Lie group G. Our principal aim is to represent the flow as a skew product flow on a trivial disk bundle $G \times V$ over G, see (1.19). The alternative approach by [GLM96], instead, works on a center bundle over the coset space G/H with respect to some discrete isotropy subgroup H. In our approach this amounts to working in the space $G \times_H V$ of H-orbits on $G \times V$, as defined in (1.15), (2.6) below.

Also, we will allow for general compact isotropies H, rather than requiring H to be finite or even trivial. In the following, the reader may find some background in Lie groups helpful; see for example [Bre72], [BtD85], [tD91], [Hel62], [Pal61], or [Die72].

To set up, we *assume* g in the Lie group G to act as a C^{k+1}-diffeomorphism $u \mapsto gu$ on the finite-dimensional Riemannian C^{k+1}-manifold M, such that the map

$$\begin{array}{rcl} \rho : G \times M & \to & M \\ (g,u) & \mapsto & gu = \rho(g,u) \end{array} \tag{1.8}$$

is C^{k+1}. Of course, we assume that G acts on M, that is $(gg')u = g(g'u)$ for all $g, g' \in G$ and $u \in M$. We also require the action to be *proper*, that is, the map $\tilde{\rho}(g,u) := (gu, u) \in M \times M$ is closed (mapping closed sets to closed sets) with compact preimages $\tilde{\rho}^{-1}(u_1, u_2)$, for any $u_1, u_2 \in M$. As a caveat, we note that $G = \mathbb{R}$ activing by shift on $BC_{unif}(\mathbb{R}, \mathbb{R})$, for example, does *not* define a proper $\mathbb{R}$ action. Still, the action of $G = SE(2)$ on a center manifold M is proper [SSW96b]. Picking $u_1 = u_2 = u_0$, in particular, we observe that the isotropy subgroup

$$H = H(u_0) := \{g \in G \mid gu_0 = u_0\} \tag{1.9}$$

is compact, for any $u_0 \in M$. Indeed, $H \times \{u_0\} = \tilde{\rho}^{-1}(u_0, u_0)$ is compact. Although M, G are allowed to be compact, in principle, we note here that the interesting new cases arise for noncompact M and G.

We fix u_0 and its isotropy H, henceforth. We construct the disk V of the trivial bundle $G \times V$ as a geometric cross section to the action of G near u_0. Using the Haar measure on the compact Lie group H, we may first assume the given Riemannian metric on M to be H-invariant, without loss of generality; see [Bre72], section VI.2. In particular, any $h \in H$ acts linearly and orthogonally on the tangent space $T_{u_0}M$ to M in u_0, by the derivative of $u \mapsto \rho(h,u)$ at $u = u_0$. Similarly, ρ induces a C^k-action of G on the C^k tangent bundle TM; we cannot assume G to act as an isometry on tangent spaces in general, if G is non-compact. It should be noted, however, that the special action (1.3) of the Euclidean group, arising in spiral wave motion, is an isometry in the usual L^p and $W^{k,p}$ spaces. In that case, G would automatically act as an isometry on a center manifold M; see [SSW96a], [SSW96b].

We will construct V as a linear version of a *slice* to the action of G in an arbitrarily small G-invariant neighborhood U, called a *tube*, around the G-orbit

$$G \cdot u_0 := \{gu_0 \mid g \in G\} \tag{1.10}$$

of u_0 as follows. Let $\mathrm{alg}(G) = T_{\mathrm{id}}\, G$ denote the Lie algebra of G and

$$T_{u_0}(Gu_0) = \mathrm{alg}(G) \cdot u_0 \tag{1.11}$$

the tangent space to the group orbit $G \cdot u_0$ at u_0. The Lie algebra of G acts on $u \in M$ by the derivative of $g \mapsto \rho(g,u)$ at $g = \mathrm{id}$. Now let the desired disk V of the bundle $G \times V$ be defined as the open ϵ_0-ball, centered at u_0, inside the orthogonal complement

$$V \subset (T_{u_0}(G \cdot u_0))^{\perp} \subseteq T_{u_0}M \tag{1.12}$$

to the orbit tangent space $T_{u_0}(G \cdot u_0)$ in $T_{u_0}M$. Note that the isotropy H of u_0 acts linearly and orthogonally on V, as it does on $T_{u_0}M$ and $T_{u_0}(G \cdot u_0)$.

To define the slice to the G-action and the G-invariant tube U around $G \cdot u_0$, let $\psi : (T_{u_0}M)_{\mathrm{loc}} \to M$ denote a local C^{k+1}-chart of M which is H-equivariant, that is

$$\psi(hv) = h\psi(v), \tag{1.13}$$

for all $v \in (T_{u_0}M)_{\mathrm{loc}}$ and $h \in H$. Here $(T_{u_0}M)_{\mathrm{loc}}$ denotes an ϵ_0-ball in $T_{u_0}M$. In fact we construct ψ^{-1}, first, such that $\psi^{-1}(u_0) = u_0$, and then achieve H-equivariance, by Haar measure, preserving the property that ψ^{-1} is a diffeomorphism; see for example [tD91], section I.5. Then $\psi(V) \subset M$ is a slice to the G-action at $u_0 \in \psi(V)$, and

$$U := G \cdot \psi(V) \tag{1.14}$$

is an open G-invariant tube around the G-orbit $G \cdot u_0$. For convenience, we also call the ϵ_0-disk $V \subset T_{u_0}M$ around u_0 a (linear) slice. We will take license to identify $u_0 \in V$ with the origin in $\mathbb{R}^l = T_{u_0}V$ sometimes.

To describe the dynamics in the tube U well, we consider the C^k-action of the direct product Lie group $G \times H$ on the Cartesian product $G \times V$, given by

$$(g_0, h)(g, v) := (g_0 g h^{-1}, hv). \tag{1.15}$$

Because the derivative of this action at (id, u_0) is surjective, by the choice (1.12) of V, the G-equivariant map

$$\begin{array}{rcl} \bar{\tau} : G \times V & \to & U \supset G \cdot u_0 \\ (g,v) & \mapsto & g\psi(v) \end{array} \tag{1.16}$$

is a submersion for small radius ϵ_0 of the disk V. In fact, the triple $(G \times V, U; \bar{\tau})$ identifies the trivial product $G \times V$ as a (generally nontrivial) C^{k+1} principal fiber bundle over U with fiber, alias structure group, H. For more details, we refer to section 2.

Returning to dynamics, consider a G-equivariant C^k vector field f on the "center" manifold M, that is

$$gf(u) = f(gu), \tag{1.17}$$

for all $u \in M, g \in G$. Of course, here we define $gf(u)$ by the induced (differential) C^k-action of G on the tangent space TM. We seek a representation of the (local) G-equivariant flow

$$\dot{u} = f(u) \tag{1.18}$$

on M near the G-orbit $G \cdot u_0$ by the *skew product flow*

$$\begin{array}{rcl} \dot{g} & = & g\mathbf{a}(v) \\ \dot{v} & = & \varphi(v) \end{array} \tag{1.19}$$

on $G \times V$. Here the maps $\mathbf{a} : V \to \mathrm{alg}(G)$ and $\varphi : V \to T_{u_0}V$ are requested to be of class C^k and H-equivariant in the following sense:

$$\begin{array}{rcl} \mathbf{a}(hv) & = & \mathrm{Ad}(h)\mathbf{a}(v) = h\mathbf{a}(v)h^{-1}, \\ \varphi(hv) & = & h\varphi(v), \end{array} \tag{1.20}$$

for all $h \in H$ and all $v \in V$. Here $\mathrm{Ad}(h)$ denotes the standard adjoint representation on the Lie algebra, and $h\varphi$ is again understood to be differential on the linear ball $V \subseteq T_{u_0}M$.

Theorem 1.1 *Let f be a G-equivariant C^k vector field on the Riemannian C^{k+1}-manifold $M, k \geq 1$, with proper C^{k+1}-action of G on M. Let $u_0 \in M$.*

Then the isotropy H of u_0 is compact. Moreover, there exists a disk slice V, an open G-invariant tube U around the group orbit $G \cdot u_0$, and H-equivariant C^k-maps $\mathbf{a}, \varphi$, as in (1.20), such that the projection $u := \bar{\tau}(g, v) \in U$ of any solution (g, v) of the skew product system (1.19) satisfies the original differential equation (1.18) in U. The projection $\bar{\tau}$ is defined in (1.16).

Conversely, for the local $G \times H$-equivariant flow defined on $(g, v) \in G \times V$ by any C^k vector field (1.19), which is H-equivariant in the sense of (1.20), the projection $u := \bar{\tau}(g, v) \in U$ induces a G-equivariant C^k vector field f on U such that (1.17), (1.18) hold.

We do not think that this theorem is particularly surprising: our proof, given in section 2, is essentially based on a coordinatization of U by the space $G \times_H V$ of the orbits in $G \times V$ under the action of the group $\{\mathrm{id}\} \times H$. This point of view is due to [Pal61] and is concisely presented in the beautiful topology textbook [tD91], section I.5.

We do think, however, that our theorem is particularly useful: in the present paper, it enables us to analyze drifting and meandering solutions on the "center manifold" M. To be precise, we fix nomenclature.

Definition 1.2 *Consider $u_0 \in M$ with isotropy H and a G-equivariant vector field f on M, as in the theorem, with lifted skew product $\dot{g} = g\mathbf{a}(v)$, $\dot{v} = \varphi(v)$ as in (1.19), (1.20).*

We call u_0 a relative equilibrium, *if*

$$\varphi(u_0) = 0 \tag{1.21}$$

In other words, $u_0 \in M$ is a relative equilibrium if, and only if, the time orbit of u_0 remains inside the group orbit $G \cdot u_0$:

$$\{(u(t)|t \in \mathbb{R}\} \subseteq G \cdot u_0. \tag{1.22}$$

Equivalently, $G \cdot u_0$ is a flow invariant manifold.

Next, take any solution $u(t) \in U$. Suppose that $u(t)$ is neither stationary nor periodic. Then, we call $u(t)$ meandering *if*

$$\{(g(t), v(t)) \mid t \in \mathbb{R}\} \subset G \times V \tag{1.23}$$

is globally defined and relatively compact. If, in contrast, the G-component

$$\{g(t) \mid t \in \mathbb{R}\} \tag{1.24}$$

is globally defined but not relatively compact, then we call $u(t)$ drifting.

Equilibria, as well as rotating waves (spirals) are examples of relative equilibria. The reference point $u_0 \in M$ is not required to be a relative equilibrium in theorem 1.1, although it will typically be in applications, and may be forced to be, by nontrivial H-equivariance of the skew product.

While the notion (1.22) of a relative equilibrium u_0 is intrinsically flow-defined, the definition (1.21) refers to a specific $G \times V$ lifting with respect to the isotropy H of u_0, as stated. For example, to apply condition (1.21) to any given point $\tilde{u}_0 \in U$ other than u_0, the vector field (1.19) has to be constructed with respect to $\tilde{u}_0$ instead of u_0. This subtlety, however, is irrelevant for small tubular neighborhoods U, as long as H is finite.

In the very special case $G = \{\mathrm{id}\}$ the maximal isolated invariant set, in the sense of [Con78], of an isolating neighborhood $V = U$ of u_0 consists precisely of the equilibria, the periodic solutions, and the meanders in $U = V$. A similar statement holds for the case of compact G.

As mentioned above, we prove our theorem in section 2. In section 3 we discuss H-equivariant Hopf bifurcation in V, in general. Section 4 collects some useful facts on actions of the Euclidean groups $SE(N)$ before we proceed sorting out drifts and meanders for $N = 2$, in section 5. We conclude, in section 6, with a slow-fast analysis of drifting circular filaments of scroll waves, so-called twisted scroll rings, in $N = 3$ dimensions.

Acknowledgment. We gratefully acknowledge helpful advice by Christian Leis. We are also much indebted to Marty Golubitsky for sharing his insight into the behavior of multi-armed spirals, as well as for some inciting competitiveness. Regina Löhr helped us with careful and very patient typesetting.

2 Tubes, slices, and skew products

In this section we prove theorem 1.1. So, let its assumptions hold. We specifically recall that

$$\dot{u} = f(u) \tag{2.1}$$

is a G-equivariant C^k vector field on the Riemannian C^{k+1}-manifold M with proper C^{k+1}-action of the Lie group G on M. Given $u_0 \in M$ with isotropy H, tube U, and linear slice V, we will relate (2.1) to a $G \times H$-equivariant C^k skew product flow

$$\begin{array}{rcl} \dot{g} & = & g\mathbf{a}(v) \\ \dot{v} & = & \varphi(v) \end{array} \tag{2.2}$$

on $G \times V$. The $G \times H$-action on $G \times V$ is defined as

$$(g_0, h)(g, v) = (g_0 g h^{-1}, hv); \tag{2.3}$$

see in particular (1.15)–(1.20). Talking about the H-action on $G \times V$, below, we will mean the action of $\{\mathrm{id}\} \times H$. Similarly, G-action refers to $G \times \{\mathrm{id}\}$.

Our proof can be outlined as follows. First, we check $G \times H$-equivariance of (2.2). After a brief digression clarifying the structure of $G \times V$ over U as a principal H bundle, we project (2.2) from $G \times V$ down to the tube $U \subset M$ by the submersion

$$\bar{\tau}(g, v) = g\psi(v), \tag{2.4}$$

defined in (1.16), to obtain a G-equivariant C^k vector field (2.1) on U from the skew product (2.2). To complete the proof, we finally lift a given G-equivariant C^k vector field f on U back to a $G \times H$-equivariant C^k skew product (2.2) on $G \times V$, such that the skew product projects down to the prescribed f, by $\bar{\tau}$.

Checking $G \times H$-equivariance of the skew product (2.2) on $G \times V$ is easy: fix $(g_0, h) \in G \times H$ and $(g, v) \in G \times V$. Then (2.3), (1.20) imply

$$\begin{array}{l} (g_0, h)(g\mathbf{a}(v), \varphi(v)) = (g_0 g\mathbf{a}(v)h^{-1}, h\varphi(v)) = \\ = (g_0 g h^{-1} h\mathbf{a}(v)h^{-1}, h\varphi(v)) = ((g_0 g h^{-1})\mathbf{a}(hv), \varphi(hv)). \end{array} \tag{2.5}$$

In other words, $(g_0, h)(g(t), v(t))$ is a solution of the skew product (2.2) if, and only if, $(g(t), v(t))$ is. This proves $G \times H$-equivariance of the skew product on $G \times V$.

In passing, we note that the skew product (2.2) with equivariance condition (1.20) is the most general form of a $G \times H$-equivariant C^k vector field on $G \times V$. Indeed, (left) G-equivariance forces the $\dot{g}$ component to be of the form $g\mathbf{a}(v)$ with $\mathbf{a}(v) \in \mathrm{alg}(G)$. Moreover, the $\dot{v}$ component must be independent of g. Then H-equivariance provides the equivariance conditions (1.20).

We briefly digress now, to clarify the structure of $G \times V$ as an H principal C^{k+1} bundle over U with fiber, alias structure group, H. Our presentation essentially follows [Pal61] and the textbook [tD91].

Identifying H-orbits of the free H-action on $G \times V$, we obtain the H orbit space

$$G \times_H V \quad := \quad G \times V \;/\; \{\mathrm{id}\} \times H. \tag{2.6}$$

It turns out that the $G(\times\{\mathrm{id}\})$-equivariant C^{k+1}-submersion $\bar{\tau}$ factorizes over the H orbit space $G \times_H V$, such that

$$\bar{\tau} : G \times V \xrightarrow{p} G \times_H V \xrightarrow{\tau} U. \tag{2.7}$$

Here p is the canonical G-equivariant C^{k+1}-submersion which projects (g, v) onto its H-orbit; it induces the structure of a C^{k+1}-manifold on $G \times_H V$ because the free H-action on $G \times V$ is C^{k+1}. In fact, $(G \times V, G \times_H V, p)$ is a G-equivariant C^{k+1} principal fiber bundle with compact fiber, alias structure group, H. The G-equivariant map τ, called *tube map,* is a C^{k+1} diffeomorphism onto the open tube U around the group orbit $G \cdot u_0$. We emphasize that these results are by no means original. They are essentially due to [Pal61], and are concisely summarized in the textbook [tD91], sections I.5, II.6.

After our bundle digression, we now project the skew product (2.2) down to M with the submersion $\bar{\tau}$, aiming at the second part of our theorem. Let $u \in M$. Since the C^{k+1}-submersion $\bar{\tau} : G \times V \to U$ is surjective, there exists (g, v) such that $\bar{\tau}(g, v) = u$. By the bundle digression, any other (g', v') in $\bar{\tau}^{-1}(u)$ is on the same H-orbit: there exists $h \in H$ such that

$$(g', v') = (gh^{-1}, hv). \tag{2.8}$$

We define $f(u)$ via the differential $D\bar{\tau}(g, v)$ of $\bar{\tau}$ with respect to g and v,

$$f(u) := D\bar{\tau}(g, v) \cdot (g\mathbf{a}(v), \varphi(v)). \tag{2.9}$$

To show that f is well-defined, we use the action of H on $G \times V$. In fact, we prefer an explicit calculation even though we could also argue "elegantly" with the H bundle structure. We start from

$$\bar{\tau}(gh^{-1}, hv) = \bar{\tau}(g, v), \tag{2.10}$$

for all $h \in H$. Differentiating with respect to g and v, we obtain

$$D\bar{\tau}(gh^{-1}, hv) \cdot (g\mathbf{a}h^{-1}, h\varphi) = D\bar{\tau}(g, v) \cdot (g\mathbf{a}, \varphi), \tag{2.11}$$

for any $\mathbf{a} \in \text{alg}(G), \varphi \in T_{u_0}V$. Therefore, $f(u)$ does not depend on the choice of $(g', v') \in \bar{\tau}^{-1}(u)$, because (2.8)–(2.11) and equivariance (1.20) imply

$$\begin{aligned} &D\bar{\tau}(g', v') \cdot (g'\mathbf{a}(v'), \varphi(v')) = \\ &= D\bar{\tau}(gh^{-1}, hv) \cdot (gh^{-1}\mathbf{a}(hv), \varphi(hv)) = \\ &= D\bar{\tau}(gh^{-1}, hv) \cdot (gh^{-1}h\mathbf{a}(v)h^{-1}, h\varphi(v)) = \\ &= D\bar{\tau}(gh^{-1}, hv) \cdot (g\mathbf{a}(v)h^{-1}, h\varphi(v)) = \\ &= D\bar{\tau}(g,\ v) \cdot (g\mathbf{a}(v), \varphi(v)) = f(u). \end{aligned} \tag{2.12}$$

This proves that $f(u)$ is indeed well-defined on $u \in U$, by (2.9).

Because $\bar{\tau} \in C^{k+1}$ and $\mathbf{a}, \varphi \in C^k$, it is obvious that f is a C^k vector field on the tube U. It remains to check G-equivariance (1.17) of f. Fixing $(g, v) \in \bar{\tau}^{-1}(u)$, this follows directly from G-equivariance of $\bar{\tau}$ and of the skew product $(g\mathbf{a}(v), \varphi(v))$. Explicitly, we have $\bar{\tau}(g_0 g, v) = g_0\bar{\tau}(g, v) = g_0 u$, and hence

$$\begin{aligned} f(g_0 u) &= D\bar{\tau}(g_0 g, v) \cdot (g_0 g\mathbf{a}(v), \varphi(v)) = \\ &= g_0 D\bar{\tau}(g, v) \cdot (g\mathbf{a}(v), \varphi(v)) = \\ &= g_0 f(u). \end{aligned} \tag{2.13}$$

This proves the second part of our theorem: the submersion $\bar{\tau}$ projects any $G \times H$-equivariant C^k vector field (2.2) on $G \times V$ down to a G-equivariant C^k vector field f on U.

It remains to, conversely, lift f from $U \subset M$ up to a skew product on $G \times V$, such that the lift projects back onto the prescribed f, by $\bar{\tau}$. Since the fiber is the isotropy group H, this is trivial if H happens to be discrete, that is, finite. Then we can simply lift the flow in U, and f, back to any sheet h_0 of the covering space $G \times V$ of U, by the local diffeomorphism $\bar{\tau}_{h_0}^{-1}$. Lifting f back to any other sheet h_1, locally near u_0, where $h_1 = h^{-1}h_0$ for some $h \in H$, we see that

$$(gh^{-1}, hv) = (\bar{\tau}_{h_0}^{-1} \circ \bar{\tau}_{h_1})(g, v) \tag{2.14}$$

induces, by linearization with respect to (g, v), the claimed H-equivariance of the lifted vector fields, where $h \in H$ acts freely as a permutation of the sheets in the covering space $G \times V$. The trivial case $H = \mathrm{id}$ was first presented by one of the authors, explaining meandering and drifting spirals [Fie95]; for a recent version see also [BHN96].

We now return to the general case of compact isotropy H. It is convenient to describe the lift of f in slightly more abstract notation. Let $w = (g_0, v) \in W := G \times V$ with left action $gw := (gg_0, v)$ of G and right action $wh := (g_0h^{-1}, hv)$ of H describe the action of the direct product $G \times H$ on W. Note that G, H act freely, separately. It remains to construct a $G \times H$-equivariant C^k vector field F on the total space W of our principal H bundle

$$\bar{\tau} : G \times V \to U, \tag{2.15}$$

such that F projects down to f by $\bar{\tau}$, that is,

$$D\bar{\tau}(w)F(w) = f(\bar{\tau}(w)) \tag{2.16}$$

for all $w \in W$.

We first define F on the linear slice $w \in \{\mathrm{id}\} \times V \subseteq G \times V$. Let P_v denote the orthogonal projection, with respect to the $H \times H$-invariant Riemannian metric on W, in the tangent space $T_{(\mathrm{id},v)}W = \mathrm{alg}(G) \times V$ onto the orthogonal complement $\left(T_{(\mathrm{id},v)}((\mathrm{id}, v) \cdot H)\right)^{\perp}$ of the right H-action. So P_v projects onto the second summand of the orthogonal decomposition

$$T_{(\mathrm{id},v)}W = T_{(\mathrm{id},v)}((\mathrm{id}, v) \cdot H) \oplus \left(T_{(\mathrm{id},v)}(\mathrm{id}, v) \cdot H)\right)^{\perp}. \tag{2.17}$$

Then we define the lifted vector field F at $w = (\mathrm{id}, v)$ as

$$F(w) := P_v(D\bar{\tau}(w))^{-1} f(\bar{\tau}(w)). \tag{2.18}$$

Note that F is now well defined on $\{\mathrm{id}\} \times V$. Indeed $\bar{\tau}^{-1}(u) = wH$, for $u = \bar{\tau}(w)$, and

$$D\bar{\tau}(w) : T_wW \to T_{\bar{\tau}(w)}U \tag{2.19}$$

is surjective. Hence the kernel of $D\bar{\tau}(w)$ is given by

$$\ker D\bar{\tau}(w) = T_w(w \cdot H) \tag{2.20}$$

in the H principal fiber bundle $\bar{\tau} : G \times V \to U$, and P_v annihilates that kernel. Thus (2.18) defines $F(w)$ properly on $w \in \{\mathrm{id}\} \times V$. Moreover, F is of class C^k on $\{\mathrm{id}\} \times V$, as are $P_v, D\bar{\tau}$, and f.

We extend F to $W = G \times V$ by the left action of G on W, defining

$$F(g, v) := gF(\mathrm{id}, v) \in T_{(g,v)}W, \tag{2.21}$$

for all $g \in G$. The vector field F is still C^k, by smoothness of the free G-action. By construction, F is G-equivariant.

We verify the projection property (2.16) next. Because $\bar{\tau}, F$, and f are all G-equivariant, it is sufficient to verify (2.16) at $w = (\mathrm{id}, v)$, that is,

$$D\bar{\tau}(\mathrm{id}, v)F(\mathrm{id}, v) = f(\bar{\tau}(\mathrm{id}, v)). \tag{2.22}$$

This follows trivially from definition (2.18) of F at $w = (\mathrm{id}, v)$, because P_v projects onto a complement of $\ker D\bar{\tau}(w)$.

To complete the proof of theorem 1.1, it remains to show equivariance of F under the right action of H, that is

$$F(gh^{-1}, hv) = F(g, v) \cdot h, \tag{2.23}$$

for all $g \in G, h \in H, v \in V$. By left G-equivariance of F, this is equivalent to showing

$$F(hw \cdot h) - hF(w) \cdot h = 0 \tag{2.24}$$

for any $w = (\mathrm{id}, v) \in \{\mathrm{id}\} \times V, h \in H$. To show (2.24), we first differentiate the relation $\bar{\tau}(gw \cdot h) = g\bar{\tau}(w)$ with respect to w to obtain

$$D\bar{\tau}(gw \cdot h)(g\tilde{w} \cdot h) = gD\bar{\tau}(w)\tilde{w}, \tag{2.25}$$

for any $g \in G, h \in H$, and $(w, \tilde{w}) \in TW$. Putting $w = (\mathrm{id}, v), g := h$, and $\tilde{w} := F(w)$, this implies

$$\begin{aligned} &D\bar{\tau}(hw \cdot h)(hF(w) \cdot h) = hD\bar{\tau}(w)F(w) = \\ &= hf(\bar{\tau}(w)) = f(h\bar{\tau}(w)) = f(\bar{\tau}(hw)) = \\ &= f(\bar{\tau}(hw \cdot h)), \end{aligned} \tag{2.26}$$

so that $hF(w) \cdot h$ is indeed a candidate for $F(hw \cdot h)$ in (2.24): the difference lies in the kernel of $D\bar{\tau}(hw \cdot h)$.

To complete the proof of (2.24), and of theorem 1.1, we finally show that $hF(w) \cdot h$ is orthogonal to $\ker D\bar{\tau}(hw \cdot h)$, as is $F(hw \cdot h)$ by definition (2.18), at $hw \cdot h = (\mathrm{id}, hv)$. Indeed, by invariance of the Riemannian metric on W with respect to the action of the compact group $H \times H$, we conclude from (2.18) at w and (2.25) that

$$\begin{aligned} hF(w) \cdot h \quad &\in h(\ker D\bar{\tau}(w))^{\perp} \cdot h = \\ &= (h(\ker D\bar{\tau}(w)) \cdot h)^{\perp} = \\ &= (\ker D\bar{\tau}(hw \cdot h))^{\perp}. \end{aligned} \tag{2.27}$$

This completes the proof of $G \times H$-equivariance of F, and of theorem 1.1. $\square$

We note that our orthogonality condition in (2.18) at $w \in \{\mathrm{id}\} \times V$ determines the lifted vector field F uniquely. We formalize this statement for $F(\mathrm{id}, v) = (\mathbf{a}(v), \varphi(v))$.

COROLLARY 2.1 *Let the assumptions of theorem 1.1 hold. Let $\langle \cdot, \cdot \rangle_{\mathrm{alg}(G)}$ denote an invariant scalar product on* $\mathrm{alg}(G)$ *under the adjoint action* $\mathrm{Ad}(h)$ *of $h \in H$, and let $(\cdot, \cdot)_V$ denote an H-invariant scalar product on the linear slice space V.*

Then the lifted vector field $F(\mathrm{id}, v) = (\mathbf{a}(v), \varphi(v))$ can be chosen such that

$$(\varphi(v), \eta v)_V = \langle \mathbf{a}(v), \eta \rangle_{\mathrm{alg}(G)}, \tag{2.28}$$

for any $v \in V, \eta \in \mathrm{alg}(H)$. The above conditions, together with the vector field f on the base U, determine F uniquely.

3 Equivariant periodic orbits in a slice

By theorem 1.1 we can discuss any local bifurcation from a relative equilibrium u_0 with isotropy H in the associated $G \times H$-equivariant skew product system

$$\begin{aligned} \dot{g} &= g\mathbf{a}(v), \\ \dot{v} &= \varphi(v). \end{aligned} \tag{3.1}$$

To interpret results in terms of $u = \bar{\tau}(g, v)$ in the tube U around $G \cdot u_0$, we just have to identify points $w = (g, v)$ on the same right H-orbit $w \cdot H \in G \times_H V$. In this section, we investigate some elementary consequences of our decomposition (3.1) in case the H-equivariant $\dot{v}$ equation possesses a periodic orbit. Such periodic orbits may arise by H-equivariant Hopf bifurcation from the H-invariant equilibrium $v = u_0$ of the $\dot{v}$ equation; for a detailed background using compactness of H see [GSS88] or [Fie88].

The spatio-temporal *symmetry* of any periodic solution $v(t)$ of $\dot{v} = \varphi(v)$, with minimal period normalized to 1, can be described by a triple (L, K, Θ) as follows. Let L denote the set of $h \in H$ mapping some point $v(t_1)$ to any point $v(t_2)$ on the periodic orbit. Denoting $\Theta(h) := t_2 - t_1$, equivariance of φ then implies

$$hv(t) = v(t + \Theta(h)), \tag{3.2}$$

for all real t. Moreover

$$\Theta : L \to S^1 := \mathbb{R}/\mathbb{Z} \tag{3.3}$$

is a homomorphism into the additively written circle group. Letting $K := \ker \Theta$, we have a normal subgroup of L, and $L/K \cong$ image(Θ). Note that the groups L, K, image(Θ) are closed. The kernel K is the isotropy of some, and hence all, $v(t)$ with $t \in \mathbb{R}$. Following [Fie88], we call $v(\cdot)$ a *discrete wave*, if image$(\Theta) = \mathbb{Z}_n = \{0, 1/n, \cdots, (n-1)/n\}$ is finite. A *rotating wave* has image$(\Theta) = S^1$.

The periodic solution $v(t)$ gives rise to solutions $g(t)$ of $\dot{g} = g\mathbf{a}(v)$. By left G-equivariance, any solution $g(t)$ with initial condition $g(0) = g_0$ is given by

$$g(t) = g_0 g_*(t), \tag{3.4}$$

where $g_*(t)$ denotes the fundamental solution

$$\begin{aligned} \dot{g}_*(t) &= g_*(t)\mathbf{a}(t) \\ g_*(0) &= \mathrm{id} \end{aligned} \tag{3.5}$$

with the abbreviation $\mathbf{a}(t) := \mathbf{a}(v(t))$.

Theorem 3.1 *Let $v(t)$ be a rotating wave solution of $\dot{v} = \varphi(v)$ in (3.1).*

Then there exist $\eta \in \mathrm{alg}(H), \mathbf{a}_0 \in \mathrm{alg}(G)$ such that

$$\begin{aligned} v(t) &= \exp(\eta t) v_0 \\ g_*(t) &= \exp((\mathbf{a}_0 + \eta)t) \exp(-\eta t). \end{aligned} \tag{3.6}$$

The projected solution $u(t) = \bar{\tau}(g_*(t), v(t))$ near the relative equilibrium u_0 is again a relative equilibrium and can be represented as

$$u(t) = \exp((\mathbf{a}_0 + \eta)t) u(0). \tag{3.7}$$

Proof: To construct η, just note that $v_0 := v(0)$ is a relative equilibrium to the action of H on V because $v(t)$ is a rotating wave. In particular

$$\dot{v}(0) \in T_{v_0}(Hv_0) = \mathrm{alg}(H) \cdot v_0. \tag{3.8}$$

Pick $\eta \in \mathrm{alg}(H)$ such that $\eta v_0 = \dot{v}(0) = \varphi(v_0)$. Let $v_*(t) := \exp(\eta t)v_0$. Then $v_*(0) = v_0$ and H-equivariance (1.20) of φ implies

$$\begin{array}{rcl} \dot{v}_*(t) & = & \exp(\eta t)\eta v_0 = \exp(\eta t)\varphi(v_0) = \\ & = & \varphi(\exp(\eta t)v_0) = \varphi(v_*(t)), \end{array} \tag{3.9}$$

for all t. Therefore $v(t) = v_*(t)$, for all t.

Let $\mathbf{a}_0 := \mathbf{a}(v(0))$ and define $g_*(t)$ as in (3.6). We have to show that $g_*(t)$ solves (3.5). Trivially $g_*(0) =$ id. Using H-equivariance (1.20) of $\mathbf{a}(v(t)) = \mathbf{a}(t)$ yields

$$\begin{array}{rl} \dot{g}_*(t) = & \exp((\mathbf{a}_0+\eta)t)\ (\mathbf{a}_0+\eta)\ \exp(-\eta t) - \\ & -\exp((\mathbf{a}_0+\eta)t)\ \eta\ \exp(-\eta t) = \\ = & \exp((\mathbf{a}_0+\eta)t)\exp(-\eta t)\exp(\eta t)\ \mathbf{a}(v(0))\ \exp(-\eta t) = \\ = & g_*(t)\ \mathbf{a}(\exp(\eta t)v(0)) = \\ = & g_*(t)\ \mathbf{a}(v(t)) = \\ = & g_*(t)\ \mathbf{a}(t). \end{array} \tag{3.10}$$

This proves (3.6). To prove (3.7), we remember that $\bar{\tau}$ is left G equivariant and collapses right H-orbits. Therefore (3.6) implies

$$\begin{array}{rl} u(t) & := \bar{\tau}(g_*(t), v(t)) = \bar{\tau}(\exp(\mathbf{a}_0+\eta)t)(\mathrm{id}, v_0)\cdot\exp(\eta t)) = \\ & = \exp((\mathbf{a}_0+\eta)t)\bar{\tau}(\mathrm{id}, v_0) = \exp((\mathbf{a}_0+\eta)t)u(0), \end{array}$$

and the theorem is proved. □

Note that the relative equilibrium $u(t)$ above can be stationary, periodic, meandering, or drifting, depending on the values of the infinitesimal generator $\mathbf{a}_0 + \eta \in \mathrm{alg}(G)$. In particular, the closure of the orbit $u(\cdot)$ can have large dimension, for example if G contains large dimensional tori. Although the motion of $u(\cdot)$ can then be quasiperiodic in time, the associated rotation numbers given by $a_0 + \eta$ vary smoothly with parameters, and phase locking does not occur.

Next let $v(t)$ be a discrete wave with symmetry $(L, K, \Theta), L/K \cong \mathbb{Z}_n$. We describe the spatio-temporal symmetry of the associated not necessarily periodic solution $u(t) = \bar{\tau}(g_*(t), v(t))$ by a triple $(\tilde{L}, \tilde{K}, \tilde{\Theta})$ similarly to the periodic case. Let $\tilde{L}$ denote the set of $g \in G$ such that $gu(t_1) = u(t_2)$, for some t_1, t_2. Letting $\tilde{\Theta}(g) := t_2 - t_1$, we obtain

$$gu(t) = u(t + \tilde{\Theta}(g)), \tag{3.11}$$

for all real t and $g \in \tilde{L}$, similarly to (3.2). Let $\Sigma := \mathbb{R}/p\mathbb{Z}$ if u is periodic with minimal period $p > 0$, and $\Sigma := \mathbb{R}$ if u is nonperiodic ($p = \infty$). Then

$$\tilde{\Theta} : \tilde{L} \to \Sigma \tag{3.12}$$

is a homomorphism with kernel, alias isotropy of any $u(t)$, denoted by $\tilde{K}$.

THEOREM 3.2 *Let v be a discrete wave solution of $\dot{v} = \varphi(v)$ in (3.1) with symmetry (L, K, Θ),* image $(\Theta) = \mathbb{Z}_n$, *and minimal period 1, as above. Let $g_*(t)$ denote the associated solution of (3.5), a nonautonomous, 1-periodic, G-equivariant equation.*

Then, for any $k \in \mathbb{Z}, h_0 \in K, h \in L, t \in \mathbb{R}$, we have

$$\begin{array}{rcl} g_*(t) & = & h_0 g_*(t) h_0^{-1} \\ g_*(t + \Theta(h) + k) & = & g_*(k) g_*(\Theta(h)) h g_*(t) h^{-1}. \end{array} \tag{3.13}$$

For the stroboscope map $g_*(1)$ of (3.5) we obtain

$$\begin{array}{rcl} g_*(1) & = & (g_*(1/n) h_*)^n h_*^{-n} \\ g_*(k) & = & g_*(1)^k \end{array} \tag{3.14}$$

where $h_* \in L$ can be chosen arbitrarily such that $\Theta(h_*) = 1/n$ generates $\mathrm{image}(\Theta) = \mathbb{Z}_n = \{0, 1/n, \cdots, (n-1)/n\}$. *In particular, the stroboscope maps $g_*(k)$ commute with $g_*(\Theta(h))h$, for all $k \in \mathbb{Z}$ and $h \in L$.*

For $g_*(\Theta(h)), \Theta(h) = k/n, k = 0, 1, \cdots, n-1$, we have the more explicit expression

$$g_*(k/n) = (g_*(1/n) h_*)^k \; h_*^{-k}. \tag{3.15}$$

The symmetry $(\tilde{L}, \tilde{K}, \tilde{\Theta})$ of the projected solution $u(t) = \bar{\tau}(g_*(t), v(t))$ with minimal "period" $0 < p \leq \infty$ satisfies

$$\begin{array}{rcl} \tilde{K} & = & K, \\ \tilde{L} & = & \{g_*(k) g_*(\Theta(h)) h \mid k \in \mathbb{Z}, h \in L\}, \text{ and} \\ \tilde{\Theta}(g_*(k) g_*(\Theta(h)) h) & = & \Theta(h) + k \qquad (\mathrm{mod}\, p), \end{array} \tag{3.16}$$

where we fix representatives $0 \leq \Theta(h) < 1$. In particular we obtain for $k := 0, h := h_*$

$$u(1/n) = g_*(1/n) h_* u(0). \tag{3.17}$$

PROOF: To prove (3.13), we first claim that

$$g_\sharp(t) := h^{-1} g_*(\Theta(h))^{-1} g_*(k)^{-1} g_*(t + \Theta(h) + k) h \tag{3.18}$$

solves the same initial value problem (3.5) as $g_*(t)$ does. Then $g_\sharp \equiv g_*$, of course. To prove the claim, differentiate (3.18) with an eye on H-equivariance (1.20) and 1-periodicity of $\mathbf{a}(t)$:

$$\begin{array}{rcl} \dot{g}_\sharp(t) & = & h^{-1} g_*(\Theta(h))^{-1} g_*(k)^{-1} \dot{g}_*(t + \Theta(h) + k) h = \\ & = & g_\sharp(t) h^{-1} \mathbf{a}(t + \Theta(h) + k) h = \\ & = & g_\sharp(t) h^{-1} \mathbf{a}(v(t + \Theta(h))) h = \\ & = & g_\sharp(t) \mathbf{a}(t). \end{array} \tag{3.19}$$

We now show that $g_\sharp(t)$ satisfies the same initial condition $g_\sharp(0) = \mathrm{id}$ as $g_*(t)$, for any choice of $h \in L$. Consider the special case $h = \mathrm{id}$, first. Then $\Theta(h) = 0$ and $g_\sharp(0) = \mathrm{id}$ is trivial. In particular $g_\sharp(t) = g_*(t)$, in that case, proving

$$g_*(t) = g_*(k)^{-1} g_*(t + k) \tag{3.20}$$

for all $t \in \mathbb{R}, k \in \mathbb{Z}$. Now (3.20) with $t := \Theta(h)$ implies $g_\sharp(0) = \mathrm{id}$, for all choices of $h \in L$. This proves $g_\sharp(t) = g_*(t)$ in (3.18).

Inserting $h := h_0 \in K = \ker \Theta$ and $k := 0$ into (3.18) with $g_\sharp = g_*,\ g_*(0) = \mathrm{id}$, we immediately see that $g_*(t)$ and h_0 commute. Together with (3.18), $g_\sharp = g_*$, this proves (3.13).

The choice $t = 1$ in (3.20) yields

$$g_*(1+k) = g_*(k)g_*(1), \tag{3.21}$$

whence $g_*(k) = g_*(1)^k$, for all $k \in \mathbb{Z}$. This also follows directly, because multiplication by $g_*(1)$ is the time 1 stroboscope map for the nonautonomous equation $\dot{g} = g\mathbf{a}(t)$ with time period 1.

Inserting $k := 0, h := h_*$ with $\Theta(h_*) = 1/n$ in (3.13) yields

$$g_*(t + 1/n) = g_*(1/n)h_*g_*(t)h_*^{-1}. \tag{3.22}$$

An n-fold iteration of (3.22), evaluated at $t = 0$, yields the expression for $g_*(1)$ in (3.14). Together with (3.21), this proves (3.14).

Similarly, a k-fold iteration of (3.22) for $k = 0, 1, \cdots, n-1$ proves (3.15). Since $h_*^{-n} \in K$ commutes with h_* and, by (3.13), with $g_*(1/n)$, the stroboscope map $g_*(1)$ in (3.14) also commutes with $g_*(1/n)h_*$. Since $g_*(1)$ also commutes with K, and because h_* generates H/K, the stroboscope $g_*(1)$ and its iterates $g_*(k)$ also commute with all $g_*(\Theta(h))h, h \in L$.

To prove (3.16), let $g \in \tilde{L}$. Then $gu(t) = u(t + \vartheta)$ for some real ϑ and all $t \in \mathbb{R}$. Upstairs, there exists $h \in H$ such that

$$\begin{aligned} (g_*(t+\vartheta), v(t+\vartheta)) &= g(g_*(t), v(t)) \cdot h = \\ &= (gg_*(t)h^{-1}, hv(t)) \end{aligned} \tag{3.23}$$

for some, and hence all, real t. Comparing the second components we see that $h \in L$ and there exists a unique $k \in \mathbb{Z}$ such that

$$\vartheta = \Theta(h) + k, \tag{3.24}$$

if we fix representatives $0 \le \Theta(h) < 1$. Comparing the first components, in view of (3.13), (3.24), we find

$$g_*(k)g_*(\Theta(h))h = g, \tag{3.25}$$

after cancellation of $g_*(t)h^{-1}$. Conversely, any such g lies in $\tilde{L}$, by (3.13), (3.23), (3.24). Letting $\tilde{\Theta}(g) = \vartheta \pmod p$, it only remains to prove $\tilde{K} = K$.

Note that $g \in \tilde{K} = \ker \tilde{\Theta}$ if, and only if, (3.23) holds with $\vartheta = 0$ and for some (hence all) t, say $t = 0$. Comparing components and using $g_*(0) = \mathrm{id}$, this is equivalent to $gh^{-1} = \mathrm{id}$ with $h \in K$. Hence $\tilde{K} = K$, and the proof is complete. □

The simple fact $\tilde{K} = K$, in our notation, implies that the isotropy groups occurring in the tube U are precisely the conjugates $gKg^{-1},\ g \in G$, of isotropy groups K occurring in the (linear) slice V. Concisely: the isotropy types in U and V coincide.

We emphasize that the spatio-temporal symmetry $\tilde{L}$ of $u(t)$, given in (3.16), is a group, and $\tilde{\Theta} : \tilde{L} \to \Sigma$ is a group homomorphism. For suitable H-equivariant choices of $\mathbf{a}(v)$, the element $g_*(1/n)$ can be thought of as an arbitrary element of the connected component G_0 of the identity in G. Indeed, H-equivariance does not impose any significant restriction on $\mathbf{a}(t), 0 \le t < 1/n$, thus leaving sufficient freedom to prescribe a path $g_*(t) \in G$ from $g_*(0) = \mathrm{id}$ to $g_*(1/n)$. However, the skew product consequences of the *interplay* of the various spatio-temporal symmetries (L, K, Θ) in equivariant Hopf bifurcation certainly deserve further investigation.

4 Basic facts on Euclidean groups

We collect some background material concerning $G = E(N)$ (or $SE(N)$), the (special) Euclidean groups on $\mathbb{R}^N$. In lemma 4.1 below, we identify the compact subgroups of G as translation conjugates of purely orthogonal groups. In lemma 4.2 this is applied to distinguish meandering from drifting solutions. We recall the semidirect product structure $(S)E(N) = (S)O(N) \times \mathbb{R}^N$ and the composition rule, coming from the standard affine action on $\mathbb{R}^N$; see (1.1)–(1.4).

For computations involving the Lie algebras $se(N)$ it is convenient to represent $(R, S) \in SE(N) = SO(N) \times \mathbb{R}^N$ isomorphically as an element in $SL(N+1)$,

$$(R, S) \mapsto \begin{pmatrix} R & S \\ 0 & 1 \end{pmatrix}, \tag{4.1}$$

in block matrix notation. With this identification, an element $(\mathbf{r}, \mathbf{s})$ of the Lie algebra $se(N)$ becomes the $(N+1) \times (N \times 1)$ matrix

$$(\mathbf{r}, \mathbf{s}) \mapsto \begin{pmatrix} \mathbf{r} & \mathbf{s} \\ 0 & 0 \end{pmatrix}. \tag{4.2}$$

In particular, conjugation, iterates, the exponential map exp, the adjoint representation Ad of $E(N)$ on $se(N)$, and the commutator $[\cdot, \cdot]$ are given by

$$\begin{array}{rcl} (R,S)(R',S')(R,S)^{-1} & = & (RR'R^{-1}, (\mathrm{id} - RR'R^{-1})S + RS'); \\ (R,S)^n & = & (R^n, (\mathrm{id} + R + \cdots + R^{n-1})S); \\ (\mathbf{r},\mathbf{s})^n & = & (\mathbf{r}^n, \mathbf{r}^{n-1}\mathbf{s}); \\ \exp\,(\mathbf{r},\mathbf{s}) & = & (\exp\,(\mathbf{r}), \mathbf{r}^{-1}(\exp\,(\mathbf{r}) - \mathrm{id})\mathbf{s}); \\ (R,S)(\mathbf{r},\mathbf{s}) & = & (R\mathbf{r}, R\mathbf{s}); \\ (R,S)(\mathbf{r},\mathbf{s})(R,S)^{-1} & = & (R\mathbf{r}R^{-1}, -R\mathbf{r}R^{-1}S + R\mathbf{s}); \\ [(\mathbf{r},\mathbf{s}),(\mathbf{r}',\mathbf{s}')] & = & ([\mathbf{r},\mathbf{r}'], \mathbf{r}\mathbf{s}' - \mathbf{r}'\mathbf{s}). \end{array} \tag{4.3}$$

The notation in (4.3) is concise, but somewhat tricky. The first/last two relations hold in the group/algebra, respectively. Similarly, $\exp(\mathbf{r}, \mathbf{s})$ is in the group. The expressions for $(\mathbf{r}, \mathbf{s})^n, (R, S)(\mathbf{r}, \mathbf{s})$ are neither in the group nor in the algebra, in general, and are evaluated in $SL(N+1)$ in the sense of (4.1), (4.2). Similarly, all equations of (4.3) are easily checked in $SL(N+1)$.

In the settings of theorems 1.1, 3.1, 3.2, the compact isotropy subgroup H of G was playing a central role. We determine the compact subgroups H of $E(N)$ next. We use the equivariant projection

$$p : E(N) = O(N) \times \mathbb{R}^N \to O(N) \tag{4.4}$$

onto the first component.

LEMMA 4.1 *Let H be a compact subgroup of $E(N)$. Then H is conjugate to its projection $p(H) \leq O(N)$ by a fixed translation $S_0 \in \mathbb{R}^N$:*

$$H \quad = \quad (\mathrm{id}, S_0)\, p(H)\, (\mathrm{id}, -S_0) \tag{4.5}$$

PROOF: We will first prove that there exists a map

$$\sigma : p(H) \to \mathbb{R}^N \tag{4.6}$$

such that H has the form

$$H = p(H)^\sigma := \{(R, \sigma(R)) \mid R \in p(H)\}. \tag{4.7}$$

In a second step, we identify a fixed $S_0 \in \mathbb{R}^N$ such that

$$\sigma(R) = (\mathrm{id} - R)S_0, \tag{4.8}$$

for all $R \in p(H)$. Then (4.5) is proved.

To construct σ, let $(R, S), (R, S') \in H$ possess the same projection $R \in p(H)$. Then, for any integer n,

$$H \ni ((R, S)(R, S')^{-1})^n = (\mathrm{id}, n(S - S')). \tag{4.9}$$

Since H is compact, this implies $S' = S$ and $\sigma := S$ is well-defined. This proves (4.6), (4.7).

For the second step note that σ is at least continuous. Indeed, H is compact and the bijection $p : H \to p(H)$ is continuous, with inverse determined by σ. Therefore p is a homeomorphism, and σ is continuous.

Multiplying $(R, S), (R', S')$ in H yields the functional equation

$$\sigma(RR') = \sigma(R) + R\sigma(R'). \tag{4.10}$$

Note continuous dependence on R'. We integrate (4.10) over R' with respect to the left invariant Haar measure on the compact Lie group $p(H)$. With the abbreviation

$$S_0 := \int_{p(H)} \sigma(R')dR', \tag{4.11}$$

we obtain

$$\begin{aligned} \sigma(R) &= \int \sigma(R)dR' = \int \sigma(RR')dR' - \int R\sigma(R')dR' = \\ &= (\mathrm{id} - R)S_0. \end{aligned} \tag{4.12}$$

This proves the lemma. □

The lemma holds, more generally, for any compact subgroup H of the general affine group $GL(N) \times R^N$. The proof is the same, and the compact group $p(H) \leq GL(N)$ may in fact be assumed to act orthogonally.

Using the notation of section 3, we now consider a periodic solution $v(t)$ of $\dot{v} = \varphi(v)$ in the skew product, with period 1, and with associated fundamental solution

$g_*(t)$ of (3.5). We derive a criterion to decide whether the projected solution $u(t) = \bar{\tau}(g_*(t), v(t))$ is meandering or drifting, in the sense of definition 1.2.

Lemma 4.2 *Let $G = SE(N)$ or $E(N)$. Consider v of period 1, and g_*, u as above. Assume $u(t)$ is neither stationary nor periodic. Let*

$$(R_*, S_*) := g_*(1). \tag{4.13}$$

Then $u(t)$ is meandering, if S_* is orthogonal to the fix space of the rotation $R_* \in SO(N)$ in $\mathbb{R}^N$, that is,

$$S_* \perp \ker(\mathrm{id} - R_*) =: (\mathbb{R}^N)^{R_*}. \tag{4.14}$$

If, on the other hand, (4.14) does not hold, then $u(t)$ is drifting.

Proof: By definition 1.2, the nonstationary, nonperiodic solution $u(t)$ is meandering if the orbit $g_*(t), t \in \mathbb{R}$, is relatively compact, and drifting otherwise. By theorem 3.2 and the differential equation (3.5) for $g_*(t)$, this orbit is relatively compact if, and only if,

$$\begin{aligned} H' &:= \operatorname{clos}\{g_*(k) \mid k \in \mathbb{Z}\} \\ &= \operatorname{clos}\{g_*(1)^k \mid k \in \mathbb{Z}\} \end{aligned} \tag{4.15}$$

is a compact subgroup of $SE(2)$. (Note here that theorem 3.2 also applies to rotating waves $v(t)$, viewed as discrete waves with arbitrary $n \in \mathbb{N}$. By lemma 4.1, the group H' is compact if, and only if, it can be conjugated to its projection $p(H') \subseteq (S)O(N)$, by a pure translation $S_0 \in \mathbb{R}^N$. This is possible if, and only if, the translation component of

$$(\mathrm{id}, -S_0)(R_*, S_*)(\mathrm{id}, S_0) = (R_*, -S_0 + S_* + R_* S_0) \tag{4.16}$$

vanishes. Using orthogonality of R_*, this is equivalent to

$$S_* \in \operatorname{image}(\mathrm{id} - R_*) = \ker(\mathrm{id} - R_*)^\perp, \tag{4.17}$$

proving claim (4.14), and the lemma. □

We note a dichotomy with respect to dimension N, here, which was also observed by [AM96]. For even N, we have $(\mathbb{R}^N)^{R_*} = \{0\}$, for generic rotations R_*, and hence generic meandering. For odd N, in contrast, $\dim(\mathbb{R}^N)^{R_*} = 1$, generically, which implies generic drifting.

If the 1-periodic solution $v(t) \in V$ possesses spatio-temporal symmetry (L, K, Θ) with non-trivial pointwise isotropy K, we obtain a particularly simple criterion excluding drifts.

Lemma 4.3 *Let $G = SE(N)$ or $E(N)$, consider v, u, g_* as above, and let $g_*(1) = (R_*, S_*)$. Assume the compact isotropy group K of $v(t)$ to be contained in $O(N)$, after conjugation by a translation as in lemma 4.1.*

Then the translation component S_* of the stroboscope map $g_*(1)$ is fixed under K, that is

$$S_* \in (\mathbb{R}^N)^K. \tag{4.18}$$

In particular, drifting is excluded if

$$(\mathbb{R}^N)^K \perp (\mathbb{R}^N)^{R_*}. \tag{4.19}$$

Most trivially, of course, condition (4.19) holds if $(\mathbb{R}^N)^K = \{0\}$ *or* $(\mathbb{R}^N)^{R_*} = \{0\}$.
PROOF: Lemma 4.2 and (4.18) imply claim (4.19). To prove (4.18), we let $h_0 \in K \leq O(N)$. Since h_0 and $g_*(1)$ commute, by theorem 3.2, (3.13), this implies

$$\begin{aligned} (R_*, S_*) &= g_*(1) = h_0 g_*(1) h_0^{-1} = \\ &= (h_0 R_* h_0^{-1}, h_0 S_*). \end{aligned} \tag{4.20}$$

Therefore $h_0 S_* = S_*$, and the lemma is proved. □

The projected solution $u(t)$ satisfies

$$\begin{aligned} u(k) &= \bar{\tau}(g_*(k), v(k)) = (g_*(1))^k \bar{\tau}(\mathrm{id}, v(0)) = \\ &= g_*(1)^k u(0) \end{aligned} \tag{4.21}$$

for all stroboscope times $k \in \mathbb{Z}$. Let $g_*(1)^k = (R_*^k, S_*^k)$. Aside from a compact part, due to R_*^k, and possibly the isotropy H of u_0, the displacement of $u(0)$ is therefore given by the translation component S_*^k of the k-fold iterated stroboscope $g_*(1)^k$. From (4.3), we recall $S_*^k = (\mathrm{id} + R_* + \cdots + R_*^{k-1})S_*$ and $R_*^k = (R_*)^k$.

To analyze S_*^k, we consider the meandering case $S_* \perp \ker(\mathrm{id} - R_*)$ next, for the stroboscope map $g_*(1) = (R_*, S_*)$. Let $(\mathrm{id} - R_*)^\dagger$ denote the pseudo-inverse of $(\mathrm{id} - R_*)$, that is, the isomorphism inverting $(\mathrm{id} - R_*)$ within the R_*-invariant subspace $(\ker(\mathrm{id} - R_*))^\perp =$ image$(\mathrm{id} - R_*)$. Define

$$S_\dagger := (\mathrm{id} - R_*)^\dagger S_*. \tag{4.22}$$

LEMMA 4.4 *As in the above setting, let* $S_* \perp \ker(\mathrm{id} - R_*)$. *Then* $g_*(1)^k, k \in \mathbb{Z}$, *are all conjugate to the rotations* $(R_*^k, 0)$ *around the origin, by the fixed translation* $S_\dagger$:

$$\begin{aligned} g_*(1)^k &= (\mathrm{id}, S_\dagger)(R_*^k, 0)(\mathrm{id}, -S_\dagger) \\ &= (R_*^k, S_\dagger - R_*^k S_\dagger). \end{aligned} \tag{4.23}$$

In particular, the translation components $S_*^k \in \mathbb{R}^N$ *of* $(g_*(1))^k$ *all lie on a sphere around* $S_\dagger \in \mathbb{R}^N$ *with radius* $|S_\dagger|_2$.
PROOF: By (4.3), applied to $(R_*^k, S_*^k) = (g_*(1))^k$, $k > 0$, and geometric summation, we have

$$\begin{aligned} S_*^k &= (\mathrm{id} + R_* + \cdots + R_*^{k-1})S_* = \\ &= (\mathrm{id} - R_*^k)(\mathrm{id} - R_*)^\dagger S_* = (\mathrm{id} - R_*^k)S_\dagger \\ &= S_\dagger - R_*^k S_\dagger. \end{aligned} \tag{4.24}$$

In case $k < 0$, the same formula holds, by $(g_*(1))^k = ((g_*(1))^{-1})^{-1}$ and (1.4). This proves (4.23) and, by orthogonality of R_*^k, the lemma. □

The radius $|S_\dagger|_2$ defined in (4.22) and lemma 4.4 relates to the "radius" of a meandering solution $u(t) = g_*(t)v(t)$ as follows. Let $u_0(t) = (\exp(\mathbf{r}_0 t), 0)u_0$ be a primary rotating wave solution, as in the introduction (1.5), (1.6). Then $u_0(t)$ rotates around its core point centered at zero. For $v(0)$ near u_0, we can consider zero also as the core point of $u(0) = \mathrm{id}\ v(0)$. Then S_*^k, the translation component of $g_*(k) = g_*(1)^k$, is the core position of $u(k) = g_*(k)v(0)$, by 1-periodicity of $v(\cdot)$. Since S_*^k all lie on a sphere around $S_\dagger$ with radius $|S_\dagger|_2$, we can call the Euclidean length $|S_\dagger|_2$

the *stroboscope radius* of $u(t)$. In section 5, (5.7) we will see how $|S_\dagger|_2 \to \infty$, when a planar meandering spiral passes through a drift resonance, for which $S_* \neq 0$ and $R_* = \mathrm{id}$.

We caution our reader that our notion of a stroboscope radius requires $u_0(t)$ to rotate around the origin. Moreover, the precise value of $|S_\dagger|_2$ depends on our choice of $t = 0$ as a reference point within the period of v. Indeed, other choices lead to expressions

$$\tilde{S}_\dagger = R_*(t)^{-1}(S_\dagger - P_* S_*(t)), \tag{4.25}$$

$0 \leq t \leq 1$, replacing $S_\dagger$, with correspondingly modified radii $|\tilde{S}_\dagger|_2$. Here P_* projects onto $\ker(id - R_*)$, orthogonally. Note that (4.25) has period 1 in t, by definition (4.22) of $S_\dagger$. Bounded modifications as in (4.25), however, do not affect the asymptotics of $|S_\dagger|_2 \to \infty$, when passage through a drift resonance occurs.

□

5 The planar case E(2): meandering and drifting multi-armed spirals

First rigorous results on meandering and drifting one-armed spirals in the plane were obtained by [Wul96], using a Lyapunov-Schmidt procedure in scales of Banach spaces. First formal results on meandering and drifting *multi-armed* spirals in the plane were obtained by [GLM96], using a formal center bundle reduction in the spirit of [Kru90]. Using the rigorous center manifold reduction due to [SSW96a], [SSW96b], the skew product structure developed in the present paper applies. We recover results of [GLM96], and investigate the behavior of meander radii at drift resonance.

Throughout this section, $G = E(2)$, and H is a compact subgroup which we may consider to be a subgroup of $O(2)$, after conjugation by a fixed translation, without loss of generality. As in section 3, we consider H-equivariant Hopf bifurcation for $\dot{v} = \varphi(\lambda, v)$ in the slice $v \in V$ of our skew product (3.1). Let (L, K, Θ) denote the spatio-temporal symmetry of our periodic solution $v(t)$, with minimal period normalized to 1. We also normalize the primary relative equilibrium u_0 to become $v = 0$, without loss of generality. The case of a rigidly rotating "primary" spiral wave with n identical arms, in the setting of the introduction, now corresponds to a rotating wave u_0 with $H = \mathbb{Z}_n \leq SO(2)$.

We begin with a simple criterion excluding drifting solutions $u(t) := \bar{\tau}(g_*(t), v(t))$ for general $H \leq O(2)$.

Corollary 5.1 *In the above planar setting, assume the isotropy group K of $v(t)$ contains some nontrivial rotation, that is, $K \leq O(2)$ is neither trivial nor generated by a single reflection.*

Then $u(t)$ cannot drift, in the sense of definition 1.2.

Proof: Suppose $K \leq O(2)$ contains some nontrivial rotation. Then K fixes only the origin, in $\mathbb{R}^2$, that is $(\mathbb{R}^2)^K = \{0\}$. By lemma 4.3, this excludes drifting. □

We look at meandering and drifting for spatio-temporal symmetries (H, K, Θ) of $v(\cdot)$ next. Throughout, we identify $\mathbb{R}^2 = \mathbb{C}$ and write $(R, S) \in SE(2)$ in complex notation:

$$R = e^{2\pi i \alpha}, \quad \alpha \in \mathbb{R}/\mathbb{Z}, \quad S \in \mathbb{C}. \tag{5.1}$$

We consider solutions $v(\cdot)$ with spatio-temporal symmetry (H, K, Θ) given by

$$\begin{array}{rcl} H &=& \mathbb{Z}_n = \{e^{2\pi i k/n} \mid k = 0, \ldots, n-1\}, \quad n \geq 2, \\ K &=& \{1\}, \\ \Theta(e^{2\pi i k/n}) &=& mk/n \in S^1 = \mathbb{R}/\mathbb{Z}, \quad k = 0, \ldots, n-1. \end{array} \tag{5.2}$$

We require $K = \{1\}$, to give drifting a chance. Note that this is equivalent to requiring the integer $m \in \{1, \ldots, n-1\}$ to be relatively prime to n.

Meandering, meander radii, drifting, and drift resonance will follow from theorem 3.2 and lemma 4.4. We will express all these effects in terms of the fractional stroboscope map

$$g_*(1/n) = \left(\exp(2\pi i \alpha_{1/n}), S_{1/n}\right). \tag{5.3}$$

Also, we have to choose $h_* \in H$ such that $\Theta(h_*) = 1/n$ generates image(Θ). Of course, we have to choose

$$\begin{array}{rcl} h_* &=& \exp(2\pi i m'/n), \text{ where} \\ m'm &\equiv& 1 \pmod{n}. \end{array} \tag{5.4}$$

In other words, m' is the unique multiplicative inverse of $m, \bmod n$.

Corollary 5.2 *With the above notation, the stroboscope map $g_*(1)$ is given explicitly by*

$$g^*(1) = \left(\exp(2\pi i n \alpha_{1/n}), \left(\sum_{k=0}^{n-1} \exp(2\pi i k(\alpha_{1/n} + m'/n))\right) S_{1/n}\right). \tag{5.5}$$

The solution $u(t) = \bar{\tau}(g_*(t), v(t))$ satisfies

$$u\,(1/n) = (\exp(2\pi i(\alpha_{1/n} + m'/n)), S_{1/n}) u(0). \tag{5.6}$$

In particular, the solution $u(t)$ is

(i)	*periodic,*	*if*	$S_{1/n} = 0$ and $\alpha_{1/n} \in \mathbb{Q}$;
(ii)	*periodic,*	*if*	$\alpha_{1/n} + m'/n \notin \mathbb{Z}$ and $\alpha_{1/n} \in \mathbb{Q}$;
(iii)	*meandering,*	*if*	$(\alpha_{1/n} + m'/n) \notin \mathbb{Z}$ and $\alpha_{1/n} \notin \mathbb{Q}$;
(iv)	*drifting,*	*if*	$\alpha_{1/n} + m'/n \in \mathbb{Z}$ and $S_{1/n} \neq 0$.

In case (iii), the meandering stroboscope radius r is given explicitly by

$$r = \frac{1}{2}\, |\sin((\alpha_{1/n} + m'/n)\pi)|^{-1} \cdot\ |S_{1/n}|_2. \tag{5.7}$$

Proof: By theorem 3.2, (3.14), we compute the stroboscope map $g_*(1)$ as

$$\begin{array}{l} (\exp(2\pi i \alpha_*), S_*) := g_*(1) = (g_*(1/n) h_*)^n h_*^{-n} = \\ = (\exp(2\pi i(\alpha_{1/n} + m'/n)), S_{1/n})^n. \end{array} \tag{5.8}$$

In particular (4.3) implies

$$\begin{array}{rcl} \alpha_* &=& n\alpha_{1/n} \qquad \pmod{1} \\ S_* &=& \left(\sum_{k=0}^{n-1} \exp(2\pi i k(\alpha_{1/n} + m'/n))\right) S_{1/n}. \end{array} \tag{5.9}$$

This proves (5.5) and case (i).

We now have to distinguish two cases. If $\alpha_{1/n} + m'/n$ is integer, then

$$\begin{array}{rcll} \alpha_* & = & 0 & \pmod 1, \\ S_* & = & nS_{1/n}, & \end{array} \tag{5.10}$$

proving case (iv). If, on the other hand, $\alpha_{1/n}+m'/n \notin \mathbb{Z}$, then we can easily compute

$$S_* = \frac{\exp(2\pi i\alpha_*) - 1}{\exp(2\pi i(\alpha_{1/n} + m'/n)) - 1} \quad S_{1/n}. \tag{5.11}$$

In view of formula (4.3) for the iterates $g_*(n) = (g_*(1))^n = (\exp(2\pi i\alpha_*), S_*)^n$, this proves cases (ii), (iii).

In case (iii), we can apply lemma 4.4 to compute the meandering radius r, because $\ker(\mathrm{id} - R_*) = \{0\}$ for $\alpha_* = n\alpha_{1/n} \notin \mathbb{Z}$. Therefore (5.11) implies

$$\begin{array}{rcl} r = |S_\dagger|_2 & = & |(\mathrm{id} - R_*)^{-1} S_*|_2 = \\ & = & |S_{1/n}|_2 / |\exp(2\pi i(\alpha_{1/n} + m'/n)) - 1| = \\ & = & \frac{1}{2}\, |S_{1/n}|_2 / |\sin((\alpha_{1/n} + m'/n)\pi)|. \end{array} \tag{5.12}$$

This proves the corollary. □

In a Hopf bifurcation situation, it is easy to derive expansions for the various cases of the previous corollary. Indeed, consider a primary n-armed spiral $u_0(t) = \exp(i\omega_{\mathrm{rot}} t) u_0(0)$ with isotropy $H = \mathbb{Z}_n$ and minimal period $T_{\mathrm{rot}} = 2\pi/(n\omega_{\mathrm{rot}})$. Assume an additional pair $\pm 2\pi i$ of imaginary eigenvalues of the linearization (in rotating coordinates). Then we can parameterize

$$v(t) = \epsilon \mathrm{e}^{2\pi i t} + O(\epsilon^2) \tag{5.13}$$

at parameter $\lambda = \lambda_0 + \lambda_2 \epsilon^2 + O(\epsilon^3)$. The equation for $g_*(t)$ becomes

$$\dot{g}_* \quad = \quad g_*(\mathbf{a}_0 + \mathbf{a}_1 v(t) + \cdots), \tag{5.14}$$

where $\mathbf{a}_0 = \mathbf{a}(v = 0)$ and $\mathbf{a}_1 = D\mathbf{a}(v = 0)$. For simplicity of presentation, we focus on the rotational component $R_*(t) = \exp(2\pi i\alpha(t))$ of $g_*(t)$. Inserting the v-expansion (5.13) we obtain

$$2\pi\dot{\alpha}(t) = \omega_{\mathrm{rot}} + \cdots, \quad \alpha(0) = 0, \tag{5.15}$$

omitting time dependent terms of order ϵ. Note that, indeed, ω_{rot} is the rotation frequency of the rotating spiral $u_0(t)$. Solving (5.15), up to terms of order ϵ, we get for $g_*(1/n) = (\alpha_{1/n}, S_{1/n})$

$$\alpha_{1/n} = \alpha(1/n) = \frac{\omega_{\mathrm{rot}}}{2\pi n} + \cdots. \tag{5.16}$$

Letting $2\pi = \omega_{\mathrm{Hopf}}$ denote the (normalized) frequency of the nontrivial Hopf eigenvalues, the transition to the drift case (iv) occurs, for example, at

$$n\alpha_{1/n} + m' = \frac{\omega_{\mathrm{rot}}}{\omega_{\mathrm{Hopf}}} + m' \equiv 0 \pmod n. \tag{5.17}$$

From (5.7) we see how the meandering stroboscope radius blows up, at this resonance, provided $S_{1/n} \neq 0$.

To analyze $S_{1/n}$ in more detail, we write the differential equation for the component $S(t)$ of $g_*(t)$ as

$$\dot{S}(t) = \mathrm{e}^{2\pi i\alpha(t)} \cdot \xi(v(t)). \tag{5.18}$$

In view of the v-expansion (5.13), we can restrict our attention to the case $v \in \mathbb{C}$. Note that the spatio-temporal symmetry

$$v(t + \Theta(h)) = h \cdot v(t) \tag{5.19}$$

then forces $h \in H = \mathbb{Z}_n \subseteq \mathbb{C}$ to act on v as complex multiplication by h^m, to be consistent with (5.2) and (5.13). Writing the H-action in this complex notation, equivariance condition (1.20) becomes

$$\xi(h^m v) = h\xi(v). \tag{5.20}$$

Expanding, as far as necessary, by

$$\xi(v) = \sum_{k,l=0}^{\infty} \xi_{kl}\, v^k\, \bar{v}^l, \tag{5.21}$$

we see that $\xi_{kl} = 0$, unless

$$(k - l)m \equiv 1 \qquad (\mathrm{mod}\ n). \tag{5.22}$$

Requiring m coprime to n, still, this yields

$$k \equiv l + m' \qquad (\mathrm{mod}\ n) \tag{5.23}$$

with the mod n multiplicative inverse m' of m. The terms of leading order are $\xi_{m',0}v^{m'}$, if $0 < m' \leq n/2$, and $\xi_{0,n-m'}\bar{v}^{n-m'}$, in case $n/2 \leq m' < n$. Integrating the $\dot{S}$ equation, up to higher order in ϵ, yields

$$S(1/n) = \epsilon^{m'} \frac{\xi_{m',0}}{\omega_{\mathrm{rot}} + 2\pi m'} \left(\mathrm{e}^{i(\omega_{\mathrm{rot}}+2\pi m')/n} - 1\right) \neq 0. \tag{5.24}$$

for $\epsilon, \xi_{m',0} \neq 0$, in case $0 < m' < n/2$. The case $n/2 < m' < n$ reads

$$S(1/n) = \epsilon^{n-m'} \frac{\xi_{0,n-m'}}{\omega_{\mathrm{rot}} + 2\pi(m' - n)} \left(\mathrm{e}^{i(\omega_{\mathrm{rot}}+2\pi m')/n} - 1\right) \neq 0. \tag{5.25}$$

for $\epsilon, \xi_{0,n-m'} \neq 0$. For $m' = n/2$, the coefficients of $\xi_{m',0}$ and $\xi_{0,n-m'}$ add. Most notably, we see a stroboscopic radius of meandering r proportional to higher powers of ϵ, in these cases; see (5.7). A similar calculation for $H = \{\mathrm{id}\}, n = 1$, yields r proportional to ϵ.

6 Meandering and drifting in three dimensions: twisted scroll rings

Let $G = SE(3)$, in this section. We first consider a primary wave $u_0(t)$ with trivial isotropy $H = \{\mathrm{id}\}$. At the end of this section, we comment on the case $H = \mathbb{Z}_n$. Pictorially, we think of u_0 as a hypothetical one parameter family of one-armed spirals with a core filament aligned along a unit circle parallel to the (x, y)-plane. The spiral patterns occur, locally, in the bundle of normal planes to the core circle. Such patterns have been called *scroll waves* by [Win73]. Moreover, assume the spirals to possess a phase difference along the family of normal planes. For simplicity, we assume that phase difference to equal the angle difference of the core points on the unit circle (rather than equaling an integer multiple of that angle.) While that pattern rotates, horizontally, around the vertical z-axis, as a rotating wave, it also propagates, vertically, along the z-axis, at constant speed. We call such a hypothetical pattern (if it exists) a *twisted scroll ring* [PW85]. The so inclined reader may also visualize smoke rings, with an inner rotating structure. For another recent example involving rigid body motion (of submarines) with $SE(3)$ symmetry see [LM96]. More mathematically, we require

$$u_0(t) = \exp(\mathbf{a}_0 t) u_0(0), \tag{6.1}$$

where u_0 has trivial isotropy H, and $\mathbf{a}_0 = (\mathbf{r}_0, \mathbf{s}_0)$ in the Lie algebra of $SE(3)$ has the special form

$$\mathbf{r}_0 = \begin{pmatrix} i\omega_0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{s}_0 \ = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{6.2}$$

We use complex notation in the horizontal (x, y)-plane, here, writing $\mathbb{R}^3 = \mathbb{C} \times \mathbb{R}$. We assume $\omega_0 \neq 0$ for the horizontal rotation frequency. Technically speaking, we might call $u_0(t)$ a drifting and rotating relative equilibrium. Lemma 4.2 explains why we choose the translation $\mathbf{s}_0$ to be vertical to the rotation plane.

Because the isotropy H is trivial, the skew product

$$\begin{array}{rcl} \dot{g}_* & = & g_* \mathbf{a}(v), \quad g_*(0) = id, \\ \dot{v} & = & \varphi(v), \end{array} \tag{6.3}$$

describes the flow in a neighborhood U of $G \cdot u_0$. We consider a family of periodic solutions $v = v(\epsilon, t)$ of period normalized to 1, bifurcating from the trivial solution $v \equiv u_0$. The parameter λ, so necessary for such a Hopf bifurcation, is suppressed. Instead, we represent dependence of $\mathbf{a}(v)$ on $v = v(\epsilon, t)$ by a differentiable function

$$\mathbf{a}(\epsilon, t) := \mathbf{a}(v(\epsilon, t)) \tag{6.4}$$

in the Lie algebra, directly. Note that

$$\mathbf{a}_0 := \mathbf{a}(0, t) = \mathbf{a}(u_0) \tag{6.5}$$

does not depend on time, while $\mathbf{a}(\epsilon, \cdot)$ has (normalized) period 1 for $\epsilon > 0$.

As in any differential equation, we can differentiate the solution $g_* = g_*(\epsilon, t)$ with respect to ϵ. Writing

$$\begin{array}{rcl} \gamma(\epsilon, t) & := & (\partial_\epsilon g_*) g_*^{-1} \\ \eta(\epsilon, t) & := & g_*^{-1} \partial_\epsilon g_* \end{array} \tag{6.6}$$

with $\gamma, \eta \in \mathrm{alg}(G)$, the differential equations for γ, η, respectively, are

$$\begin{array}{rcl} \dot{\gamma} & = & g_*(\partial_\epsilon \mathbf{a}) g_*^{-1}, \\ \dot{\eta} & = & [\eta, \mathbf{a}] + \partial_\epsilon \mathbf{a}, \end{array} \tag{6.7}$$

with initial conditions $\gamma = \eta = 0$ at $t = 0$ and $\dot{} = \partial_t$. For example, at $\epsilon = 0$ and $t = 1$, the derivative of the stroboscope map g_* with respect to ϵ becomes

$$\partial_\epsilon g_* = \gamma \cdot g_* = \int_0^1 \exp(\mathbf{a}_0 t') \partial_\epsilon \mathbf{a}(t') \exp(-\mathbf{a}_0 t') dt' \; g_*, \tag{6.8}$$

because $g_*(0, t) = \exp(\mathbf{a}_0 t)$.

What are the effects of this ϵ-expansion on the dynamics, alias on the iterates of the stroboscope map $g_*(\epsilon, 1) = (R(\epsilon), S(\epsilon))$? At $\epsilon = 0$, we have

$$\begin{array}{rcl} R(0) & = & \left(\begin{array}{cc} \exp(i\omega_0) & 0 \\ 0 & 1 \end{array} \right), \\ S(0) & = & \left(\begin{array}{c} 0 \\ 1 \end{array} \right) \in \mathbb{C} \times \mathbb{R}. \end{array} \tag{6.9}$$

For small positive ϵ, by (6.8), we get a rotation axis of $R(\epsilon)$ near the z-axis, tilted by an angle proportionally to ϵ. Conjugating by a small rotation around a horizontal axis orthogonal to that angle, we can assume

$$R(\epsilon) \quad = \quad \left(\begin{array}{cc} \exp(i\omega) & 0 \\ 0 & 1 \end{array} \right), \tag{6.10}$$

with $\omega = \omega(\epsilon)$ near ω_0. Conjugating by yet another rotation around the z-axis, afterwards, we can assume

$$S(\epsilon) \quad = \quad \left(\begin{array}{c} \sigma(\epsilon) \\ 1 + s(\epsilon) \end{array} \right), \tag{6.11}$$

with small complex σ and small real s. Now we can iterate the stroboscope map $g_*(\epsilon, 1) = (R, S)$. Using (3.14) and (4.3),

$$g_*(\epsilon, n) = (R_n, S_n) = (R, S)^n = (R^n, \sum_{k=0}^{n-1} R^k S). \tag{6.12}$$

With (6.10) we obtain the rotation

$$R_n = \left(\begin{array}{cc} \exp(i\omega n) & 0 \\ 0 & 1 \end{array} \right). \tag{6.13}$$

Similarly, the translation $S_n = (\sigma_n, n + sn)$ is given by

$$\sigma_n = (\sum_{k=0}^{n-1} e^{i\omega k}) \sigma. \tag{6.14}$$

Summarizing, the propagation speed of our original twisted scroll ring u_0 experiences periodic fluctuations, due to $v(t)$. The period near 1 has been scaled to 1, here. Lighted with a stroboscope at (normalized) integer times $t = n$, we observe identical shapes of the twisted scroll ring. It propagates along the (slightly tilted) z-axis at a slightly modified average speed $1 + s$. This oscillating propagation is a three-dimensional analogue of Hopf bifurcation from a traveling wave in one space dimension; for the latter see [Pos92]. In a plane perpendicular to the vertical propagation direction, our scroll ring performs a planar meandering motion of stroboscopic radius

$$r = \frac{1}{2} |\sin(\pi\omega_{\text{rot}}/\omega_{\text{Hopf}})|^{-1} \cdot |\sigma|, \tag{6.15}$$

as has been investigated in section 5. (We have returned to the notation $\omega_{\text{rot}} = \omega$, $\omega_{\text{Hopf}} \approx 2\pi$ used there). Typically, $|\sigma|$ will be of order ϵ. Note the horizontal drift resonance which occurs at integer values

$$\omega_{\text{rot}}/\omega_{\text{Hopf}} \in \mathbb{Z}. \tag{6.16}$$

At these values, the meandering propagation along a spiral around the z-axis becomes a slow sidewards drift, away from the z-axis.

Additional isotropies $H = \mathbb{Z}_n$, commuting with the primary rotation $\exp(\mathbf{r}_0 t)$ of $u_0(t)$ in (6.1), (6.2), can be incorporated. Note that H rotates around the vertical z-axis. For the horizontal planar meandering, the results of section 5 will reappear. Specifically, let (H, K, Θ) be the spatio-temporal symmetry of a bifurcating periodic solution $v(t)$ in the skew product. According to lemma 4.3, nontrivial rotations in K will force translations S_* in the stroboscope map $g_*(1) = (R_*, S_*)$ to point along the z-axis. Likewise, R_* near $\exp \mathbf{r}_0$ will rotate around the z-axis, unless $R_* = \text{id}$. Indeed $R_* \in SO(3) \setminus \{\text{id}\}$ commutes with K, by (3.13) and lemma 4.3, and hence R_* and K fix the same axis of rotation. Therefore horizontal meandering is impossible, if K contains a nontrivial rotation. Pure drifts $g_*(1) = (\text{id}, S_*)$ can only point along the z-axis.

If $K = \{\text{id}\}$ is trivial, transverse meandering perpendicular to the direction of propagation becomes possible. Indeed, let $g_*(1/n) = (R_{1/n}, S_{1/n})$. Again we conjugate the axis of $R_{1/n}$ to be vertical, so that

$$R_{1/n} = \begin{pmatrix} \exp(i\alpha_{1/n}) & 0 \\ 0 & 1 \end{pmatrix}. \tag{6.17}$$

Then $S_{1/n}$ possesses a rather irrelevant vertical component, which only modifies the vertical propagation speed. The important horizontal component, however, produces periodicity, meandering, and drifting phenomena transversely to the propagation direction. Note how the ϵ-expansions (5.24), (5.25) force the transverse drifting to be of small radius, or the transverse drifting to be slow.

Arrows by American Indians and other early, even neolithic civilizations are a practical visualization of some of the results discussed here. In fact, elastic vibrations and interaction with the air flow could lead to destabilization of the straight flight path. However, the feathers can provide an isotropy K, if they prevent rotation around the axis of the arrow. This isotropy, in turn, prevents transverse drifting and fixes the direction of propagation to be, quite literally, "straight as an arrow". Even in the case of a rotating feathered arrow, transverse deviations caused by symmetry breaking bifurcations from the straight path will be slow, due to (5.24), (5.25).

REFERENCES

[Bar94] D. Barkley. Euclidean symmetry and the dynamics of rotating spiral waves. *Phys. Rev. Lett.*, 72, 164–167, 1994.

[BE93] M. Braune and H. Engel. Compound rotation of spiral waves in a light-sensitive Belousov-Zhabotinsky medium. *Chem. Phys. Lett.*, 204, 257–264, 1993.

[BHN96] V. A. Biktashev, A. V. Holden, and E. V. Nikolaev. Spiral wave meander and symmetry of the plane. Preprint, University of Leeds, 1996.

[Bre72] G. E. Bredon. *Introduction to Compact Transformation Groups.* Academic Press, New York, 1972.

[BtD85] T. Bröcker and T. tom Dieck. *Representations of Compact Lie Groups.* Springer-Verlag, New York, 1985.

[Con78] C. Conley. *Isolated Invariant Sets and the Morse Index.* Conf. Board Math. Sci. 38, AMS, Providence, 1978.

[Die72] J. Dieudonné. *Treatise on Analysis* 3. Academic Press, New York, 1972.

[Fie88] B. Fiedler. *Global Bifurcation of Periodic Solutions with Symmetry.* Lect. Notes Math. 1309, Springer-Verlag, Berlin, 1988.

[Fie95] B. Fiedler. Meandering Spirals. Lecture at Workshop *Dynamics and Symmetry*, D. Chillingworth (org.), Isaac Newton Institute, Cambridge, 1995.

[GLM96] M. Golubitsky, V. LeBlanc, and I. Melbourne. Meandering of the spiral tip – an alternative approach. Preliminary version, Houston, 1996.

[GSS88] M. Golubitsky, I. Stewart, and D. Schaeffer. *Singularities and Groups in Bifurcation Theory* 2. Springer-Verlag, New York, 1988.

[Hel62] S. Helgason. *Differential Geometry and Symmetric Spaces.* Academic Press, New York, 1962.

[JSW89] W. Jahnke, W.E. Skaggs, and A.T. Winfree. Chemical vortex dynamics in the Belousov-Zhabotinskii reaction and in the two-variable Oregonator model. *J. Chem. Phys.*, 93, 740–749, 1989.

[Kru90] M. Krupa. Bifurcations of relative equilibria. *SIAM J. Math. Analysis*, 21, 1453–1486, 1990.

[LM96] N. E. Leonard and J. E. Marsden. Stability and drift of underwater vehicle dynamics: Mechanical systems with rigid motion symmetry. Preprint, 1996.

[NvORE93] S. Nettesheim, A. von Oertzen, H.H. Rotermund, and G. Ertl. Reaction diffusion patterns in the catalytic CO-oxidation on Pt(110) – front propagation and spiral waves. *J. Chem. Phys.*, 98, 9977–9985, 1993.

[Pal61] R. S. Palais. On the existence of slices for actions of non-compact Lie groups. *Ann. of Math.*, 73, 295–323, 1961.

[PW85] A.V. Panfilov and A. T. Winfree. Dynamical simulations of twisted scroll rings in three dimensional excitable media. *Physica D*, 17, 323–330, 1985.

[Pos92] G. Pospiech. *Eigenschaften, Existenz und Stabilität von travelling wave Lösungen zu einem System von Reaktions-Diffusions-Gleichungen.* Dissertation, Heidelberg, 1992.

[Ran82] D. Rand. Dynamics and symmetry: Predictions for modulated waves in rotating fluids. *Arch. Rat. Mech. Analysis*, 79, 1–38, 1982.

[Ren82] M. Renardy. Bifurcation from rotating waves. *Arch. Rat. Mech. Analysis*, 79, 49–84, 1982.

[SSW96a] B. Sandstede, A. Scheel, and C. Wulff. Center-manifold reduction for spiral waves. C.R. Acad. Sc., Série I, to appear.

[SSW96b] B. Sandstede, A. Scheel, and C. Wulff. Dynamics of spiral waves in unbounded domains using center-manifold reductions. In preparation, 1996.

[tD91] T. tom Dieck. *Topologie.* Walter de Gruyter, Berlin, 1991.

[UNUM93] Zs. Ungvarai-Nagy, J. Ungvarai, and S.C. Müller. Complexity in spiral wave dynamics. *Chaos*, 3(1), 15–19, 1993.

[Win73] A. T. Winfree. Scroll-shaped waves of chemical activity in three dimensions. *Science*, 181, 937–939, 1973.

[Wul96] C. Wulff. *Theory of meandering and drifting spiral waves in reaction-diffusion systems.* Dissertation, Berlin, 1996.

[Reference added in proof:]

[AM96] P. Ashwin and I. Melbourne. Noncompact drift for relative equilibria and relative periodic orbits. Preprint, University of Houston, 1996.

Bernold Fiedler
Institut für Mathematik I
Freie Universität Berlin
Arnimallee 2-6
D-14195 Berlin, Germany
fiedler@fu-berlin.de

Björn Sandstede
Div. of Applied Mathematics
Brown University
Providence, RI 02912, USA

Arnd Scheel
Institut für Mathematik I
Freie Universität Berlin
Arnimallee 2-6
D-14195 Berlin, Germany

Claudia Wulff
Institut für Mathematik I
Freie Universität Berlin
Arnimallee 2-6
D-14195 Berlin, Germany